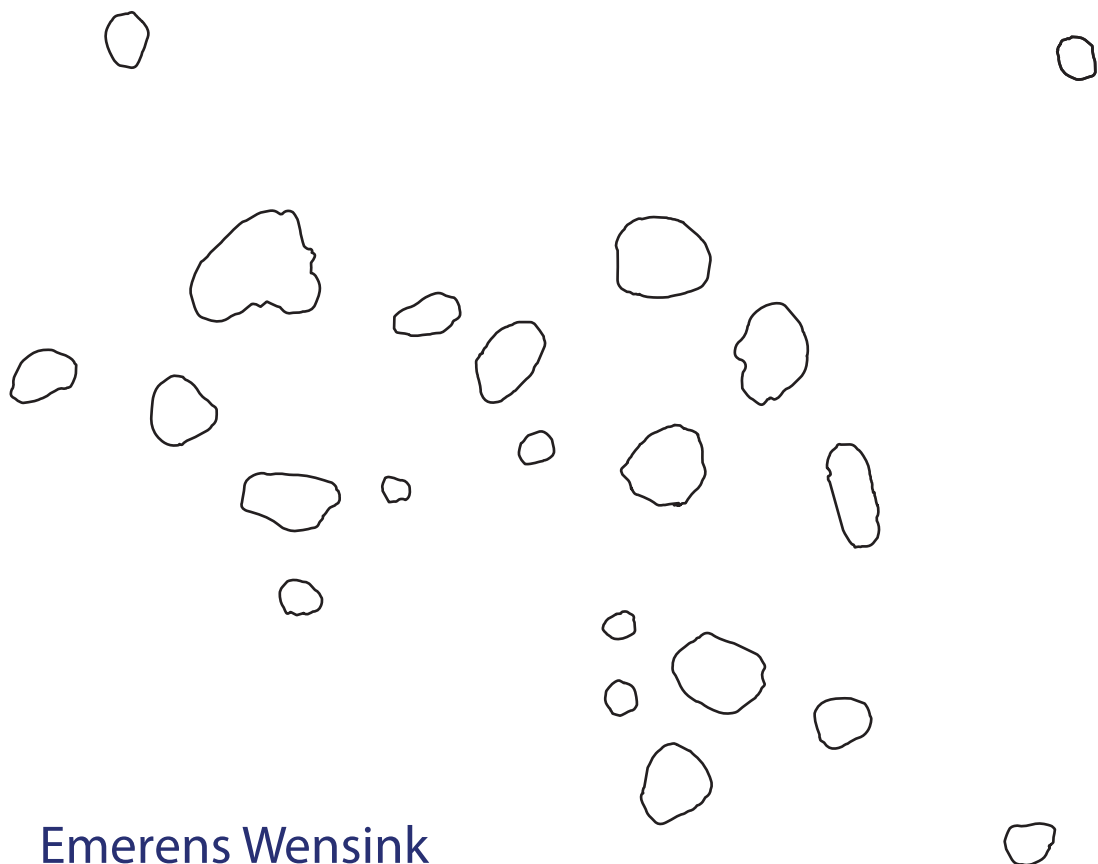


# TOWARDS PERSONALIZED TREATMENT FOR PATIENTS WITH METASTATIC COLORECTAL CANCER



Emerens Wensink



ISBN: 978-94-6458-660-2  
Cover design: Publiss | [www.publiss.nl](http://www.publiss.nl)  
Inside cover: Emerens Wensink  
Lay-out: Publiss | [www.publiss.nl](http://www.publiss.nl)  
Print: Ridderprint | [www.ridderprint.nl](http://www.ridderprint.nl)

© Copyright 2022: Grada Emerentia Wensink, the Netherlands.

Financial support for the publication of this thesis was kindly provided by ABN AMRO, ChipSoft, Hubrecht Organoid Technology, Integraal Kankercentrum Nederland (IKNL), Nordic Pharma B.V. and Servier Nederland Farma.

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

# **Towards personalized treatment for patients with metastatic colorectal cancer**

**Op weg naar gepersonaliseerde behandeling van patiënten met  
uitgezaaide dikkedarmkanker**  
(met een samenvatting in het Nederlands)

## **Proefschrift**

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

donderdag 22 december 2022 des middags te 2.15 uur

door

**Grada Emerentia Wensink**

geboren op 7 mei 1986  
te Lelystad

**Promotoren:**

Prof. dr. M. Koopman

Prof. dr. O.W. Kranenburg

**Copromotor:**

Dr. J.M.L. Roodhart

**Beoordelingscommissie:**

Prof. dr. G. Kops

Prof. dr. P.O. Witteveen

Prof. dr. M.J.L. Ligtenberg

Prof. dr. C.H. van Gils (voorzitter)

Dr. M. Chalabi

# TABLE OF CONTENTS

General introduction		9
<b>Part I</b>	<b>Optimizing treatment for patients with deficient mismatch repair colorectal cancer</b>	<b>23</b>
Chapter 1	Survival of patients with deficient mismatch repair metastatic colorectal cancer in the pre-immunotherapy era	25
Chapter 2	A review of the sensitivity of metastatic colorectal cancer patients with deficient mismatch repair to standard-of-care chemotherapy and monoclonal antibodies, with recommendations for future research	53
Chapter 3	Mismatch repair status in patient-derived colorectal cancer organoids does not affect intrinsic tumour cell sensitivity to systemic therapy	85
<b>Part II</b>	<b>Patient-derived organoids as a predictive biomarker to guide cancer treatment</b>	<b>111</b>
Chapter 4	Patient-derived organoids as a predictive biomarker for treatment response in cancer patients	113
Chapter 5	Colorectal cancer organoids reflect patient response to systemic treatment	161
<b>Part III</b>	<b>Prediction models for patients with metastatic colorectal cancer</b>	<b>199</b>
Chapter 6	External validation of two established clinical risk scores predicting outcome after local treatment of colorectal liver metastases in a nationwide cohort	201
Chapter 7	Predicting early extrahepatic recurrence after local treatment of colorectal liver metastases	239
Chapter 8	External validation of the Colon Life nomogram for predicting 12-week mortality in Dutch metastatic colorectal cancer patients treated with trifluridine-tipiracil in daily practice	287

Chapter 9	Summary (English)	317
Chapter 10	General discussion & future perspectives	329
Appendices	Summary in Dutch (Nederlandse samenvatting)	349
	Authors & affiliations	363
	Review committee	373
	List of publications	377
	Acknowledgements (Dankwoord)	381
	About the author	387





# GENERAL INTRODUCTION

---

In the Netherlands, 13,000 people are diagnosed yearly with colorectal cancer (CRC)<sup>1</sup>. Upon diagnosis, 20-24% have metastatic disease while approximately 20% develop metastatic disease after the initial diagnosis of localized disease<sup>2,3</sup>. Metastases are most frequently located in the lymph nodes, liver, lung and peritoneum<sup>4</sup>. There have been many improvements in the treatment of metastatic colorectal cancer (mCRC) patients, including new types of targeted treatment, immunotherapy and advances in the local treatment of metastases. Despite these advances, patients with mCRC continue to have a poor prognosis<sup>2,5</sup>.

In a population-based study of Dutch patients with mCRC, the median overall survival (OS) was 12 months and remained unchanged from 2008 to 2016<sup>5</sup>. These results differ markedly from the results seen in patients enrolled in clinical trials for first line systemic treatment, where the median OS has improved and reached 30 months<sup>6-8</sup>. The difference in survival between trial-based and population-based patients re-enforces the need to ensure that study results can be translated to the general population of mCRC patients. Since only a minority of patients participate in clinical trials, study results should be examined in ‘real-life’ patients reflecting the total population of patients with mCRC<sup>5</sup>. This thesis will focus on improving the treatment of patients with mCRC using personalized treatment.

## **TREATMENT FOR PATIENTS WITH METASTATIC COLORECTAL CANCER**

Treatment options for patients with mCRC consist of systemic treatment, including chemotherapy, targeted treatment and immunotherapy, and local treatment, such as resection of metastases. Depending on the extent of disease, treatment options for patients with mCRC can be broadly described as having two aims<sup>9</sup>. For patients with (potentially) resectable metastases, treatment has the intention to cure<sup>9</sup>. For example, in patients with metastases confined to the liver, local treatment of these metastases through resection or tumour ablation can result in 5-year OS rates of up to 55%<sup>10-12</sup>. However, in patients with extensive, permanently unresectable disease, treatment is aimed to improve survival and quality of life<sup>9</sup>.

Systemic treatment for mCRC patients commonly consists of chemotherapy combined with targeted treatment. An overview of standard-of-care treatments is shown in **Table 1**. For first and second line treatment, the majority of patients receive a chemotherapy

combination regimen consisting of a fluoropyrimidine (5-fluorouracil or capecitabine) with oxaliplatin and/or irinotecan and an angiogenesis inhibitor (anti-vascular endothelial growth factor, VEGF; bevacizumab)<sup>13,14</sup>. Depending on the tumour's mutational profile and primary tumour location ('sidedness'), additional targeted agents may be available, including an anti-epidermal growth factor receptor (EGFR) antibody or a *BRAF* inhibitor<sup>6,13,15</sup>. For patients with refractory disease, treatment may include trifluridine/tipiracil (TAS-102)<sup>16</sup>. For specific subgroups of CRC, other treatment types are available.

**Table 1.** Commonly used treatments for metastatic colorectal cancer in the Netherlands

Drug name	Mechanism of action	Typical use
<b>Chemotherapy</b>		
Capecitabine, oral	Pyrimidine analog; interrupts DNA synthesis	Used as a single agent or in combination (oral fluorouracil)
Fluorouracil	Pyrimidine analog; interrupts DNA synthesis	In combination with leucovorin
Irinotecan	Topoisomerase I inhibitor; interrupts the breaking and rejoining of DNA strands during replication	Used as a single agent or in combination
Leucovorin	Folic acid analog; interrupts DNA synthesis	In combination with fluorouracil
Oxaliplatin	Alkylating agent; causes DNA breaks	Used only in combination, not as a single agent
Trifluridine/tipiracil	Nucleic acid analogue/ thymidine phosphorylase inhibitor; interrupts DNA synthesis	Used as a single agent
<b>Targeted treatment</b>		
Bevacizumab	Humanized monoclonal antibody to VEGF; interrupts growth of blood vessels	In combination with fluorouracil/leucovorin, oxaliplatin, irinotecan
Cetuximab	Recombinant chimeric monoclonal antibody to <i>EGFR</i> ; interrupts or stops cell growth	Used only for <i>KRAS</i> / <i>NRAS</i> wildtype tumours
Encorafenib	<i>BRAF</i> inhibitor; interrupts or slows down cell growth	Used only for <i>BRAF</i> <sup>V600E</sup> -variant tumours in combination with cetuximab
Panitumumab	Humanized monoclonal antibody to <i>EGFR</i> ; interrupts or slows down cell growth	Used only for <i>KRAS</i> / <i>NRAS</i> wildtype tumours
<b>Immune checkpoint inhibitors</b>		
Ipilimumab*	Humanized monoclonal antibody against CTLA-4; activates T-cell-mediated immune response	Used only for dMMR tumours in combination with nivolumab

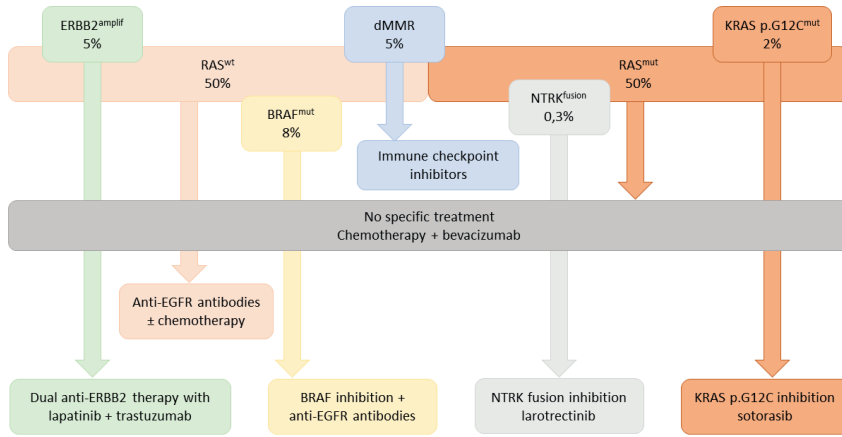
Drug name	Mechanism of action	Typical use
Nivolumab*	Humanized monoclonal antibody against PD-1 receptor; activates T-cell-mediated immune response	Used only for dMMR tumours
Pembrolizumab	Humanized monoclonal antibody against PD-1 receptor; activates T-cell-mediated immune response	Used only for dMMR tumours; approved in the first or subsequent line of therapy

Adapted from Biller *et al.*<sup>13</sup> to reflect treatments available in the Netherlands for mCRC. These treatments are European Medicines Agency (EMA) approved and have been included in the CieBOM advisory by the Dutch Society of Medical Oncology. \* Nivolumab and ipilimumab are not available in daily clinical practice (outside of clinical trials) since financial compensation for mCRC is not yet available in the Netherlands.

*Abbreviations:* BRAF (B-Raf proto-oncogene), CieBOM (scientific commission of the Dutch Society of Medical Oncology), CTLA-4 (cytotoxic T-lymphocyte-associated antigen 4), dMMR (deficient mismatch repair), EGFR (epithelial growth factor receptor), PD-1 (programmed cell death-1), VEGF (vascular endothelial growth factor).

## PERSONALIZED TREATMENT IN METASTATIC COLORECTAL CANCER

Colorectal cancer is a heterogeneous disease, with a complex cascade of molecular alterations developing during carcinogenesis<sup>17,18</sup>. Certain molecular alterations and CRC subtypes can be used to guide treatment, enabling personalized treatment or precision treatment (**Figure 1**). CRC tumours may have oncogenic mutations in driver genes, such as RAS and BRAF, which affect the prognosis of patients and response to targeted treatments<sup>17</sup>. These molecular alterations can function as predictive biomarkers, since they identify patients who are more likely to respond, or not respond, to a given treatment<sup>19</sup>. Predictive biomarkers in CRC include selecting patients with BRAF<sup>V600E</sup> mutant tumours for BRAF inhibition plus anti-EGFR combination treatment, ERBB2 amplifications for anti-ERBB2 therapy, KRAS p.G12C mutant tumours for sotorasib, NTRK fusion for larotrectinib treatment and excluding patients with RAS/BRAF mutant tumours from anti-EGFR targeted treatment<sup>15,20–23</sup>.

**Figure 1.** Guiding metastatic CRC treatment based on the molecular landscape of the tumourFig 1. Adapted from Punt *et al.*<sup>17</sup>

An important CRC subtype which impacts treatment is deficient mismatch repair (dMMR) CRC. Patients with dMMR CRC have an impaired DNA mismatch repair pathway, resulting in a high mutational load and microsatellite instability (MSI-H)<sup>24</sup>. A deficient mCRC tumour is rare in mCRC patients (5%)<sup>25</sup>. However, these patients have a poor prognosis in the metastatic setting and have a lower response rate to chemotherapy and targeted treatment<sup>25,26</sup>. The mechanism behind the poor response rate of these patients is unknown. Fortunately, patients with dMMR mCRC tumours show a long-lasting response to immune checkpoint inhibitors (ICI) in the neo-adjuvant, first-line and later-line setting<sup>27–31</sup>. Aside from ICI, it is unclear which specific systemic treatment types offer benefit for patients with dMMR CRC tumours.

Despite advancements in new treatment options and better selection of patients who are most likely to benefit, survival in mCRC has not improved for the general population<sup>5</sup>. A large obstacle in enabling precision medicine for CRC patients is the inability to predict to which treatment a patient will respond, since for the most commonly given treatment – chemotherapy – no biomarkers exist<sup>19</sup>. Patients still often receive a one-size fits all approach. Additionally, despite the use of molecular profiling in selecting patients for targeted treatment, patients still have a poor response rate while receiving targeted treatment<sup>15,32</sup>. As such, patients may be exposed to toxic treatment, without clinical benefit, but with the associated side effects and high treatment costs. There is an urgent need to better tailor treatment to individual patients with mCRC.

---

A novel approach for implementing personalized treatment for oncology patients is the use of patient-derived organoids (PDOs)<sup>33–35</sup>. PDOs are three-dimensional self-organizing structures comprised of epithelial cells, mimicking its corresponding tumour<sup>36,37</sup>. Various drugs, including standard-of-care chemotherapy and targeted treatment, can be screened on PDOs to assess if a patient is likely to respond to a given treatment type<sup>34,35,38,39</sup>. As well, PDOs can be used to identify new treatments for individual patients and new treatment combinations for CRC subtypes<sup>40,41</sup>. Although promising, these results need to be validated in larger studies and should assess if PDOs can predict response for all common treatment types and can be used for clinical decision-making. Currently, it is unclear if implementing PDOs to guide treatment in CRC is feasible and will result in benefit for patients.

Lastly, prediction models can help estimate the risk that a specific event, e.g. recurrence after resection of liver metastases, will occur<sup>42</sup>. A high quality prediction model can help inform decision-making by identifying patients who are most likely to benefit from a given treatment. How well a prediction model is able to predict the risk for individual patients in the general population can be assessed through external validation of the risk model in an external cohort of patients. Real-life patient cohorts, including the Prospective Dutch Colorectal cancer (PLCRC) cohort<sup>43</sup> and national database registries, such as the Netherlands Cancer Registry (NCR), are ideally suited to assess the performance of prediction models and create improved prediction models.

## AIMS AND OUTLINE OF THIS THESIS

A holy grail in oncology research is the ability to predict to which type of systemic treatment an individual patient will respond. Despite advances in prognostic and predictive biomarkers for the treatment of mCRC patients, there remains an urgent need to further improve personalized treatment. The research in this thesis will examine different strategies which can be used to enable personalized treatment for patients with mCRC.

**Part I** of this thesis examines the role of microsatellite-instability, resulting from tumours with a dMMR complex, as a prognostic and predictive biomarker in the treatment of mCRC. In 2018, when the research of this thesis started, the first clinical study results regarding ICI in the treatment of CRC patients with dMMR tumours were emerging<sup>29,31,44</sup>. Due to the low incidence rate of dMMR tumours among mCRC patients<sup>25</sup>, studies

regarding the survival of these patients while receiving chemotherapy and targeted treatment were lacking, which complicated the interpretation of the single-arm ICI trials. In **Chapter 1**, we report the OS of patients with dMMR mCRC tumours while receiving first line, second line and third line chemotherapy and targeted treatment. Both trial-based patients, from the CAIRO, CAIRO2 and CAIRO3 clinical phase III trials<sup>45–47</sup>, and real-life patients, from the Netherlands Cancer Registry were included in the study.

Due to the worse prognosis of patients with dMMR mCRC tumours while receiving non-immunotherapy systemic treatment, we performed an evidence-based review to assess the sensitivity of these patients to standard-of-care chemotherapy and monoclonal antibodies. The results of this review are shown in **Chapter 2**, which summarizes clinical trials reporting treatment efficacy of different types of chemotherapy and monoclonal antibodies in dMMR mCRC patients. As well, we review the biological rationale behind a potential differential benefit of these treatments in dMMR mCRC patients and provide recommendations for future research.

Furthermore, in **Chapter 3**, we show results from drug screens on a panel of CRC organoids derived from patients with dMMR CRC tumours and proficient mismatch repair (pMMR) CRC tumours to assess whether dMMR causes therapy resistance in a tumour cell-intrinsic manner. We also report results regarding the role of activating mutations in the *KRAS* and *BRAF* oncogenes in therapy resistance.

**Part II** of this thesis assesses patient-derived organoid (PDOs) as a predictive biomarker for oncology treatment. In 2018, results were published from a pioneering study, which examined whether PDOs can be used to predict treatment response in oncology patients<sup>38</sup>. By 2021, 17 studies spanning multiple cancer types had been published. In **Chapter 4**, we review the results, including pooled results, from these 17 studies which examined individualized tumour response testing using PDOs. The review reports regarding three properties which predictive biomarkers must fulfill: analytical validity, clinical validity and clinical utility, while also providing perspective for future research.

Although promising, the results for PDOs as a predictive marker for CRC patients should be confirmed in other patient cohorts. **Chapter 5** shows the results of a study which examines if organoids can be used to predict treatment response for standard systemic treatment in mCRC patients. This study provides preliminary results to guide our prospective **OPTIC** clinical trial (**O**rganoids to **P**redict **T**herapy **R**esponse **I**n **C**olorectal **C**ancer<sup>48</sup>). In the OPTIC trial, we examine if PDOs can predict a patient's response to treatment and assess

---

feasibility aspects. The OPTIC trial obtains tissue for PDO establishment prior to treatment initiation for 160 patients with mCRC receiving standard-of-care systemic treatment and the patient receives systemic treatment through their oncologist. Drug screens are performed *ex vivo* on the PDOs and compared with the observed clinical response of the patient's treatment. For the study reported in **Chapter 5**, organoids were derived from 18 CRC patients, from the HUB-Cancer Biobank and prospective OPTIC clinical study, who were treated for mCRC with systemic treatment. We investigate the strength of association of PDO drug screen results with patient response and we assess if organoids can be used to predict treatment response for all standard-of-care systemic chemotherapies and targeted treatment in multiple treatment lines.

**Part III** of this thesis assesses prediction models for patients with mCRC. The use of prediction models is assessed in two different populations of mCRC patients: patients undergoing local treatment for colorectal cancer liver metastases (CRLM, **Chapter 6-7**) and in mCRC patients receiving trifluridine/tipiracil treatment (**Chapter 8**). In **Chapter 6** we show the results of an external validation of two existing prediction models in a large cohort of real-life patients who received local treatment for CRLM from the Netherlands Cancer Registry. Subsequently, we report the results of a new prediction model which describes the risk of developing an early extrahepatic recurrence after receiving local treatment for CRLM in the same cohort of real-life patients with CRLM in **Chapter 7**.

Trifluridine/tipiracil treatment increases the OS of patients with refractory mCRC compared to placebo, however, not all patients benefit from treatment<sup>16</sup>. Potential strategies to personalize treatment include selecting patients based on existing prediction models and thymidine kinase (TK1) expression<sup>49-51</sup>. In **Chapter 8**, we externally validate an existing prediction model, the Colon Life nomogram<sup>50</sup>, which predicts 12-week mortality in patients receiving trifluridine/tipiracil treatment for refractory mCRC. The external validation cohort consists of mCRC patients from the QUALITAS study who received trifluridine/tipiracil in daily practice<sup>52</sup>. The QUALITAS study demonstrated that quality of life summary score (QoL-SS, EORTC QLQ-C30) at start of treatment was independently associated with shorter OS<sup>52</sup>. Thus, in **Chapter 8**, we also explore the additional prognostic value of QoL and TK1 expression as two novel prognostic markers for patients receiving FTD/TPI treatment.

Lastly, a summary of the results from this thesis and a general discussion and future perspectives are reported.

## REFERENCES

1. IKNL. NKR-cijfers [Internet]. Available from: <https://iknl.nl/nkr-cijfers>
2. Brouwer N.P.M., Bos A.C.R.K., Lemmens V.E.P.P., Tanis P.J., Hugen N., Nagtegaal I.D., et al. An overview of 25 years of incidence, treatment and outcome of colorectal cancer patients. Results from the Netherlands Cancer Registry. *Int J Cancer*. (2018);**143**(11):2758–66.
3. Elferink M.A.G., de Jong K.P., Klaase J.M., Siemerink E.J., de Wilt J.H.W. Metachronous metastases from colorectal cancer: a population-based study in North-East Netherlands. *Int J Colorectal Dis*. (2015);**30**(2):205–12.
4. Meyer Y., Olthof P.B., Grünhagen D.J., de Hingh I., de Wilt J.H.W., Verhoef C., et al. Treatment of metachronous colorectal cancer metastases in the Netherlands: A population-based study. *Eur J Surg Oncol*. (2021);**4**:S0748-7983(21)00951-3.
5. Hamers P.A.H., Elferink M.A.G., Stellato R.K., Punt C.J.A., May A.M., Koopman M., et al. Informing metastatic colorectal cancer patients by quantifying multiple scenarios for survival time based on real-life data. *Int J Cancer*. (2021);**148**(2):296–306.
6. Venook A.P., Niedzwiecki D., Lenz H.J., Innocenti F., Fruth B., Meyerhardt J.A., et al. Effect of first-line chemotherapy combined with cetuximab or bevacizumab on overall survival in patients with KRAS wild-type advanced or metastatic colorectal cancer a randomized clinical trial. *JAMA*. (2017);**317**(23):2392–401.
7. Cremolini C., Loupakis F., Antoniotti C., Lupi C., Sensi E., Lonardi S., et al. FOLFOXIRI plus bevacizumab versus FOLFIRI plus bevacizumab as first-line treatment of patients with metastatic colorectal cancer: updated overall survival and molecular subgroup analyses of the open-label, phase 3 TRIBE study. *Lancet Oncol*. (2015);**16**(13):1306–15.
8. Heinemann V., Weikersthal L.F. Von, Decker T., Kiani A., Vehling-kaiser U., Scholz M., et al. FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): A randomised, open-label, phase 3 trial. *Lancet Oncol*. (2014);**15**(10):1065–75.
9. Van Cutsem E., Cervantes A., Adam R., Sobrero A., Van Krieken J.H., Aderka D., et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol*. (2016);**27**(8):1386–422.
10. Rees M., Tekkis P.P., Welsh F.K., O'Rourke T., John T.G. Evaluation of long-term survival after hepatic resection for metastatic colorectal cancer: a multifactorial model of 929 patients. *Ann Surg*. (2008);**247**(1):125–35.
11. Kanas G.P., Taylor A., Primrose J.N., Langeberg W.J., Kelsh M.A., Mowat F.S., et al. Survival after liver resection in metastatic colorectal cancer: review and meta-analysis of prognostic factors. *Clin Epidemiol*. (2012);**4**:283–301.
12. Andreou A., Aloia T.A., Brouquet A., Dickson P. V, Zimmitti G., Maru D.M., et al. Margin status remains an important determinant of survival after surgical resection of colorectal liver metastases in the era of modern chemotherapy. *Ann Surg*. (2013);**257**(6):1079–88.
13. Biller L.H., Schrag D. Diagnosis and treatment of metastatic colorectal cancer: A review. *JAMA*. (2021);**325**(7):669–85.
14. Pietrantonio F., Morano F., Corallo S., Miceli R., Lonardi S., Raimondi A., et al. Maintenance Therapy with Panitumumab Alone vs Panitumumab Plus Fluorouracil-Leucovorin in Patients with RAS Wild-Type Metastatic Colorectal Cancer: A Phase 2 Randomized Clinical Trial. *JAMA Oncol*. (2019);**5**(9):1268–75.

- 
15. Kopetz S., Grothey A., Yaeger R., Van Cutsem E., Desai J., Yoshino T., et al. Encorafenib, Binimetinib, and Cetuximab in BRAF V600E-Mutated Colorectal Cancer. *NEJM*. (2019);**381**(17):1632–43.
  16. Mayer R.J., Van Cutsem E., Falcone A., Yoshino T., Garcia-Carbonero R., Mizunuma N., et al. Randomized Trial of TAS-102 for Refractory Metastatic Colorectal Cancer. *NEJM*. (2015);**372**(20):1909–19.
  17. Punt C.J.A., Koopman M., Vermeulen L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat Rev Clin Oncol*. (2017);**14**(4):235–46.
  18. The Cancer Atlas Genome Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. (2012);**487**(7407):330–7.
  19. Sawyers C.L. The cancer biomarker problem. *Nature*. (2008);**452**(7187):548–52.
  20. Van Cutsem E., Köhne C.H., Hitre E., Zaluski J., Chien C.R.C., Makhson A., et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *NEJM*. (2009);**360**(14):1408–17.
  21. Drilon A., Laetsch T.W., Kummar S., DuBois S.G., Lassen U.N., Demitri G.D., et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *NEJM*. (2018);(378):731–9.
  22. Hong D.S., Fakih M.G., Strickler J.H., Desai J., Durm G.A., Shapiro G.I., et al. KRAS G12C Inhibition with Sotorasib in Advanced Solid Tumors. *N Engl J Med*. (2020);**383**(13):1207–17.
  23. Meric-Bernstam F., Hurwitz H., Singh R.K.P., McWilliams R.R., Fakih M., VanderWalde A., et al. Pertuzumab and trastuzumab for HER2-amplified metastatic colorectal cancer: an updated report from MyPathway, a multicentre, open-label, phase 2a multiple basket study. *Lancet Oncol*. (2019);**20**(4):518–30.
  24. Gelsomino F., Barbolini M., Spallanzani A., Pugliese G., Cascinu S. The evolving role of microsatellite instability in colorectal cancer: A review. *Cancer Treat Rev*. (2016);**51**:19–26.
  25. Koopman M., Kortman G.A.M., Mekenkamp L., Ligtenberg M.J.L., Hoogerbrugge N., Antonini N.F., et al. Deficient mismatch repair system in patients with sporadic advanced colorectal cancer. *Br J Cancer*. (2009);**100**(2):266–73.
  26. Aasebø K., Dragomir A., Sundström M., Mezheyeuski A., Edqvist P.H., Eide G.E., et al. Consequences of a high incidence of microsatellite instability and BRAF-mutated tumors: A population-based cohort of metastatic colorectal cancer patients. *Cancer Med*. (2019);**8**(7):3623–35.
  27. Chalabi M., Fanchi L.F., Dijkstra K.K., Van den Berg J.G., Aalbers A.G., Sikorska K., et al. Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat Med*. (2020);**26**:566–76.
  28. André T., Shiu K.-K., Kim T.W., Jensen B.V., Jensen L.H., Punt C., et al. Pembrolizumab in Microsatellite-Instability–High Advanced Colorectal Cancer. *NEJM*. (2020);**383**(23):2207–18.
  29. Overman M.J., McDermott R., Leach J.L., Lonardi S., Lenz H., Morse M.A., et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *Lancet*. (2017);**18**(9):1182–91.
  30. Lenz H.-J., Lonardi S., Zagonel V., Van Cutsem E., Limon M.L., Yeung K., et al. Nivolumab plus low-dose ipilimumab as first-line therapy in microsatellite instability-high/DNA mismatch repair deficient metastatic colorectal cancer: Clinical update. *JCO*. (2020);**38**(Suppl 4):abstr 11.

31. Le D.T., Uram J.N., Wang H., Bartlett B.R., Kemberling H., Eyring A.D., et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *NEJM*. (2015);**372**(26):2509–20.
32. Kim T.W., Elme A., Park J.O., Udrea A.A., Kim S.Y., Ahn J.B., et al. Final Analysis of Outcomes and RAS/BRAF Status in a Randomized Phase 3 Study of Panitumumab and Best Supportive Care in Chemorefractory Wild Type KRAS Metastatic Colorectal Cancer. *Clin Colorectal Cancer*. (2018);**17**(3):206–14.
33. Van De Wetering M., Francies H.E., Francis J.M., Bounova G., Iorio F., Pronk A., et al. Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell*. (2015);**161**(4):933–45.
34. Ooft S.N., Weeber F., Dijkstra K.K., McLean C.M., Kaing S., van Werkhoven E., et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Sci Transl Med*. (2019);**11**(513):eaay2574.
35. Ganesh K., Wu C., O'Rourke K.P., Szeglin B.C., Zheng Y., Sauvé C.E.G., et al. A rectal cancer organoid platform to study individual responses to chemoradiation. *Nat Med*. (2019);**25**(10):1607–14.
36. De Souza N. Organoids. *Nat Methods*. (2018);**15**(1):23.
37. Li M., Izpisua Belmonte J.C. Organoids - Preclinical models of human disease. *NEJM*. (2019);**380**(6):569–79.
38. Vlachogiannis G., Hedayat S., Vatsiou A., Jamin Y., Fernández-mateos J., Khan K., et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science*. (2018);**359**(6378):920–6.
39. Yao Y., Xu X., Yang L., Zhu J., Wan J., Shen L., et al. Patient-Derived Organoids Predict Chemoradiation Responses of Locally Advanced Rectal Cancer. *Cell Stem Cell*. (2020);**26**(1):17–26.
40. Narasimhan V., Wright J.A., Churchill M., Wang T., Rosati R., Lannagan T.R.M., et al. Medium-throughput Drug Screening of Patient-derived Organoids from Colorectal Peritoneal Metastases to Direct Personalized Therapy. *Clin Cancer Res*. (2020);**26**(14):3662–70.
41. Verissimo C.S., Overmeer M., Ponsioen B., Drost J., Verissimo C.S., Mertens S., et al. Targeting mutant RAS in patient-derived colorectal cancer organoids by combinatorial drug screening. *Elife*. (2016);**15**(5):e18489.
42. Collins G.S., Reitsma J.B., Altman D.G., Moons K.G.M. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med*. (2015);**162**(1):55–63.
43. Burbach J.P.M., Kurk S.A., Coebergh van den Braak R.R.J., Dik V.K., May A.M., Meijer G.A., et al. Prospective Dutch colorectal cancer cohort: an infrastructure for long-term observational, prognostic, predictive and (randomized) intervention research. *Acta Oncol (Madr)*. (2016);**55**(11):1273–80.
44. Overman M.J., Lonardi S., Yeung K., Wong M., Lenz H., Gelsomino F., et al. Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair–Deficient/Microsatellite Instability–High Metastatic Colorectal Cancer. *JCO*. (2018);**36**(8):773–9.
45. Koopman M., Antonini N.F., Douma J., Wals J., Honkoop A.H., Erdkamp F.L.G., et al. Sequential versus combination chemotherapy with capecitabine, irinotecan, and oxaliplatin in advanced colorectal cancer (CAIRO): a phase III randomised controlled trial. *Lancet*. (2007);**370**:135–42.

- 
46. Tol J., Koopman M., Rodenburg C.J., Cats A., Creemers G.J., Schrama J.G., et al. A randomised phase III study on capecitabine, oxaliplatin and bevacizumab with or without cetuximab in first-line advanced colorectal cancer, the CAIRO2 study of the Dutch Colorectal Cancer Group (DCCG). An interim analysis of toxicity. *Ann Oncol.* (2008);**19**(4):734–8.
  47. Simkens L.H.J., Van Tinteren H., May A., Ten Tije A.J., Creemers G.J.M., Loosveld O.J.L., et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): A phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet.* (2015);**385**(9980):1843–52.
  48. Netherlands Trial Register [Internet]. Available from: <https://www.trialregister.nl/trial/7305>
  49. Kuboki Y., Nishina T., Shinozaki E., Yamazaki K., Shitara K., Okamoto W., et al. TAS-102 plus bevacizumab for patients with metastatic colorectal cancer refractory to standard therapies (C-TASK FORCE): an investigator-initiated, open-label, single-arm, multicentre, phase 1/2 study. *Lancet Oncol.* (2017);**18**(9):1172–81.
  50. Pietrantonio F., Miceli R., Rimassa L., Lonardi S., Aprile G., Mennitto A., et al. Estimating 12-week death probability in patients with refractory metastatic colorectal cancer: The Colon Life nomogram. *Ann Oncol.* (2017);**28**(3):555–61.
  51. Fernández Montes A., Carmona-Bayonas A., Jimenez-Fonseca P., Vázquez Rivera F., Martínez Lago N., Covela Rúa M., et al. Prediction of survival in patients with advanced, refractory colorectal cancer in treatment with trifluridine/tipiracil: real-world vs clinical trial data. *Sci Rep.* (2021);**11**(1):1–10.
  52. Hamers P.A.H., Vink G.R., Elferink M.A.G., Stellato R.K., Dijksterhuis W.P.M., Punt C.J.A., et al. Quality of life and survival of metastatic colorectal cancer patients treated with trifluridine-tipiracil (QUALITAS). *Clin Colorectal Cancer.* (2022);**21**(2):154–66.





A large, stylized, light blue and purple letter 'I' graphic that serves as a background element for the title.

# **OPTIMIZING TREATMENT FOR PATIENTS WITH DEFICIENT MISMATCH REPAIR COLORECTAL CANCER**

**PART I**

A vertical strip of colorful, abstract, watercolor-like background on the right side of the page, featuring shades of pink, purple, and blue.



# **SURVIVAL OF PATIENTS WITH DEFICIENT MISMATCH REPAIR METASTATIC COLORECTAL CANCER IN THE PRE-IMMUNOTHERAPY ERA**

*British Journal of Cancer*, 2020, 124(2), 399-406.

Emerens Wensink  
Marloes Elferink  
Anne May  
Linda Mol  
Patricia Hamers  
Sandra Bakker  
Geert-Jan Creemers  
Jan Willem de Groot  
Gerty de Klerk  
Brigitte Haberkorn  
Annebeth Haringhuizen  
Ronald Hoekstra  
Cornelis Hunting  
Emile Kerver  
Danielle Mathijssen-van Stein  
Marco Polée  
Johannes Pruijt  
Patricia Quarles van Ufford-Mannesse  
Sandra Radema  
Ronald Rietbroek  
Lieke Simkens  
Bea Tanis  
Daan ten Bokkel Huinink  
Manuel Tjin-A-Ton  
Cathrien Tromp-van Driel  
Monique Troost  
Agnes van de Wouw  
Franchette van den Berkmortel  
Anke van der Pas  
Ankie van der Velden  
Marjan van Dijk  
Joyce van Dodewaard-de Jong  
Edith van Druten  
Theo van Voorthuizen  
Gerrit Jan Veldhuis  
Henk Verheul  
Hanneke Vestjens  
Jeroen Vincent  
Onno Kranenburg  
Cornelis Punt  
Geraldine Vink  
Jeanine Roodhart  
Miriam Koopman

## **CHAPTER 1**

## ABSTRACT

**Background:** Metastatic colorectal cancer patients with deficient mismatch repair (dMMR mCRC) benefit from immunotherapy. Interpretation of the single-arm immunotherapy trials is complicated by insignificant survival data during systemic non-immunotherapy. We present survival data on a large, comprehensive cohort of dMMR mCRC patients, treated with or without systemic non-immunotherapy.

**Methods:** 281 dMMR mCRC patients ( $n=54$  from three prospective phase 3 CAIRO trials;  $n=227$  from the Netherlands Cancer Registry). Overall survival was analysed from diagnosis of mCRC (OS), from initiation of first line (OS1) and second line (OS2) systemic treatment. Cox regression analysis examined prognostic factors. As comparison for OS 2746 MMR proficient mCRC patients were identified.

**Results:** Of 281 dMMR patients, 62% received first line and 26% second line treatment. Median OS was 16.0 months (13.8–19.6) with anti-tumour therapy and 2.5 months (1.8–3.5) in untreated patients. OS1 was 12.8 months (10.7–15.2) and OS2 6.2 months (5.4–8.9) in treated dMMR patients. Treated dMMR patients had a 7.6 month shorter median OS than pMMR patients.

**Conclusion:** Available data from immunotherapy trials lack a control arm with standard systemic treatment. Given the poor outcome compared to the immunotherapy results, our data strongly suggests a survival benefit of immunotherapy in dMMR mCRC patients.

## BACKGROUND

Approximately 5% of metastatic colorectal cancer (mCRC) patients have a tumour with deficient DNA mismatch repair (dMMR), also referred to as microsatellite instability<sup>1</sup>. dMMR arises through germline mutations or epigenetic methylation and inactivation of the MMR pathway, resulting in insertions or deletions in tandem repetitive sequences in DNA, a hypermutated genome, and a strong immune infiltrate<sup>2</sup>. Sporadic dMMR tumours frequently harbour a *BRAFV600E* mutation and sporadic dMMR patients have a worse prognosis compared to Lynch syndrome patients<sup>2,3</sup>.

A dMMR tumour status has prognostic and therapeutic implications for patients, with dMMR having a favourable impact on prognosis in early-stage CRC, but resulting in a worse prognosis in mCRC compared to proficient mismatch repair (pMMR) tumours<sup>1,4,5</sup>. Immune checkpoint inhibitor (ICI) trials have shown a durable response in pre-treated dMMR mCRC patients<sup>6-8</sup>. dMMR tumours are highly sensitive to ICI due to the high mutational load, immune infiltrate and immune checkpoint signalling<sup>6</sup>. Overman demonstrated a one-year OS rate of 85% in dMMR mCRC patients, who were refractory to at least one systemic treatment line, upon treatment with nivolumab/ipilimumab<sup>8</sup>. ICI have been approved by the Food and Drug Administration (FDA) in dMMR mCRC patients beyond first line treatment. However, the European Medicines Agency (EMA) has not approved ICI based on the lack of a standard control arm without immunotherapy in the ICI trials. For a better interpretation of published immunotherapy results, survival data beyond first line in dMMR mCRC patients who did not receive immunotherapy are needed.

Due to the low incidence of dMMR among mCRC patients, data on survival in dMMR mCRC patients receiving standard systemic treatments are scarce<sup>1,3,9-12</sup>. Published data of survival beyond first line treatment in a cohort of dMMR mCRC patients not receiving immunotherapy is highly needed. We report on the survival and factors affecting survival of a large cohort of dMMR mCRC patients.

## MATERIALS AND METHODS

### Study population

We analysed dMMR mCRC patients in two populations: population-based patients registered in the Netherlands Cancer Registry (NCR) managed by the Netherlands Comprehensive Cancer Organisation (IKNL) and trial-based patients from three

prospective phase 3 first line clinical trials (CAIRO<sup>13</sup>, CAIRO2<sup>14</sup> and CAIRO3<sup>15</sup>). For trial-based patients, patient inclusion criteria, informed consent and study protocols for the trials were published previously<sup>13-15</sup>. The privacy rights for patients were maintained.

Inclusion criteria for the current analysis were histologically proven mCRC with a dMMR tumour. Patients who received immunotherapy during the course of their disease were excluded. Clinical data of all newly diagnosed cancer patients in the Netherlands are registered in the NCR. dMMR mCRC patients with an incidence date between January 1<sup>st</sup>, 2015-December 31<sup>st</sup>, 2017 were included, since mismatch repair status is registered in the NCR for patients with an incidence date only after 2015. Figure 1 is a flow diagram of the patients identified in the NCR cohort. dMMR status was known when determined in routine clinical practice during the period of data registration. NCR-derived patients include all synchronous and some metachronous mCRC patients with known dMMR status due to the data collection procedure of the NCR. For the same registration period, the overall survival and treatment status for pMMR mCRC NCR-derived patients was collected.

**Figure 1.** Flow diagram of Netherlands Cancer Registry (NCR) patients

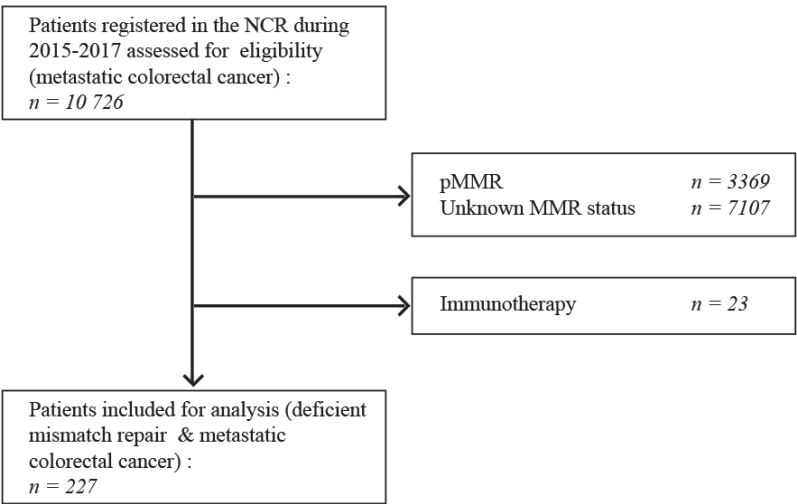


Fig 1. All NCR patients with histologically proven metastatic colorectal cancer (mCRC) were assessed for eligibility. Patients were excluded if mismatch repair (MMR) status was unknown or proficient MMR (pMMR) and if patients received immunotherapy during treatment. The final population-based cohort consists of 227 patients with deficient mismatch repair (dMMR) mCRC.

## Data collection

For population-based patients, pseudonymized clinical data on demographic characteristics, tumour characteristics and treatment information (type, response) were obtained from the NCR. Vital status for NCR patients was obtained using a yearly coupling with the municipal population registry to the cancer registry on February 1<sup>st</sup>, 2019. For NCR pMMR patients, clinical variables were obtained from the first registration period. For patients in the CAIRO trials, clinical data was available and follow-up information was updated up to October, 2019. Data registration was completed for all dMMR patients, population-based and trial-based, by ensuring that no more than 30 days of clinical data was lacking prior to the vital status coupling. If data registration could not be completed, vital status was censored to the last known date of clinical data ( $n=13$ ).

Any *BRAF* mutation detected was included in the definition of *BRAF*-mutant status. Sidedness of the primary tumour was defined as right-sided (coecum-transverse colon), left-sided (splenic flexure-sigmoid) and rectosigmoid/rectal. Anti-tumour therapy was defined as systemic treatment (excluding adjuvant therapy) or local treatment (surgical resection of metastases, radiofrequency ablation (RFA), microwave ablation (MWA), HIPEC (Hyperthermic Intraperitoneal Chemotherapy) or PIPAC (Pressurized Intra Peritoneal Aerosol Chemotherapy)). Anti-tumour therapy was categorized as follows: no anti-tumour therapy, local treatment only, local and systemic treatment and systemic treatment only.

## Deficient mismatch repair

In the CAIRO trials dMMR status was determined according to the study protocol<sup>1,16,17</sup>. In the NCR cohort, dMMR was determined according to Dutch guidelines in accredited laboratories, using immunohistochemistry and/or polymerase chain reaction.

## Study parameters

Overall survival was defined as the interval from diagnosis of metastatic disease until death of any cause or date of last follow-up if alive (OS). In patients receiving systemic treatment, OS was measured from each treatment line initiation, resulting in OS1 in patients receiving first line systemic treatment, and similarly OS2 and OS3 in patients receiving second line or third line systemic treatment starting from second line or third line initiation, respectively.

Survival rates and patient characteristics were obtained from published ICI trials<sup>7,8</sup>. In order to bring our results in perspective with ICI trials without a control arm, our results were reported alongside the most comparable ICI trials which analyzed survival from second line.

## Statistical analysis

Baseline characteristics of the patients were analysed for the whole cohort and relevant differences between population-based and trial-based groups were described. Kaplan-Meier curves and 9-month and 12-month survival rate estimates were obtained for OS, OS1, OS2 and OS3. Subgroup analyses between population-based and trial-based patients were performed using the log-rank test.

Cox regression univariate and multivariable analysis was performed in patients receiving first line treatment for OS. Ten pre-selected prognostic factors were selected: age at diagnosis of metastatic disease, gender, trial participation, *BRAF* mutation, primary tumour sidedness, metastatic sites location, stage at diagnosis, number of treatment lines given, primary tumour resection and metastasectomy<sup>3,10,18</sup>. An unadjusted median overall survival (from diagnosis metastatic disease) for each level of the covariates was obtained by performing a log-rank test in patients receiving first line systemic therapy. Multiple imputation by chained equations was used for covariates with missing data<sup>19</sup>. From the complete dataset of variables, predictor variables with a correlation >0.20 with the missing variables and <15% missing values were selected to use alongside the ten covariates and Cox regression outcome variables in multiple imputation. Patients with missing data were compared to patients with complete-cases (Supplementary Table 1). Univariate hazard ratios for each covariate were obtained using Cox regression. A stratified Cox proportional hazards multivariable model was obtained using the pre-selected ten covariates for OS; stratified for the number of treatment lines received since this covariate violated the proportional hazards assumption. Regression analysis was performed on each imputed dataset and combined using Rubin's rules.

All analyses were performed in R (version 3.5.1, “survival”, “survminer”, “mice” and “lattice” packages<sup>20</sup>).

## RESULTS

### Patient characteristics

The cohort comprises 281 patients: 227 population-based (NCR) and 54 trial-based patients. The characteristics of the cohort are described in Table 1.

**Table 1.** Patient characteristics

		<b>dMMR Cohort</b> <i>n=281</i>
<b>Patient type</b>	Trial-based	54 (19.2)
	Population-based	227 (80.8)
<b>Age (%)</b>	≤ 55 years	59 (21.1)
	56-65 years	57 (20.4)
	66-75 years	103 (36.9)
	> 75 years	60 (21.5)
<b>Female (%)</b>		159 (56.6)
<b>BRAF mutational status (%)</b>	Wildtype	68 (45.3)
	Mutation	82 (54.7)
	Unknown	131
<b>RAS mutational status (%)</b>	Wildtype	104 (81.2)
	Mutation	24 (18.8)
	Unknown	153
<b>Stage (%)</b>	I	2 (0.7)
	II	28 (10.0)
	III	88 (31.5)
	IV	161 (57.7)
<b>Sidedness (%)</b>	Right-sided	202 (72.9)
	Left-sided	55 (19.9)
	Rectosigmoid/Rectum	20 (7.2)
<b>Synchronous metastatic pattern (%)</b>		194 (69.0)
<b>Metastatic localization (%)</b>	Liver-only	61 (21.7)
	Extra-hepatic	133 (47.3)
	Peritoneal	87 (31.0)
<b>Number of metastatic sites (%)</b>	1	166 (59.3)
	2	71 (25.4)
	3	33 (11.8)
	≥4	10 (3.6)
<b>Primary tumour resection (%)</b>		219 (77.9)
<b>Metastasectomy (%)</b>		64 (22.8)
<b>Local treatment metastases (%)</b>	RFA	6 (2.3)
	MWA	1 (0.4)
	HIPEC	27 (9.6)

		<b>dMMR Cohort</b> <i>n</i> =281
<b>Anti-tumour therapy (%)</b>	No treatment	72 (25.6)
	Local treatment	36 (12.8)
	Local and systemic treatment	38 (13.5)
	Systemic treatment	135 (48.0)
<b>Adjuvant chemotherapy (%)</b>		47 (16.7)
<b>Systemic therapy</b>	First line treatment (%)	173 (61.6)
	Second line treatment (%)	72 (25.6)
	Third line treatment (%)	21 (7.5)
	Fourth line treatment (%)	3 (1.1)
<b>WHO PS (at start first line, % of first line)</b>	Score 0	57 (52.2)
	Score 1	45 (41.3)
	Score >=2	7 (6.4)
	Unknown	64

Characteristics of patients at diagnosis of metastatic disease with treatment information during the course of disease. Trial-based patients were obtained from the CAIRO (n=19), CAIRO2 (n=31) and CAIRO3 (n=4) phase III randomized controlled trials. WHO PS percentages relative to amount of people receiving first line treatment. Sidedness of the primary tumour was defined as right-sided (coecum-transverse colon), left-sided (splenic flexure-sigmoid) and rectosigmoid/rectal. Local treatment was defined as metastasectomy or local metastatic treatment (RFA, MWA or HIPEC/PIPAC) and systemic treatment as all systemic therapy given for metastatic disease (excluding adjuvant therapy). Missing values are not shown if missing frequency was less than 5%.

*Abbreviations:* WHO PS (World Health Organisation Performance Score), HIPEC (Hyperthermic intraperitoneal chemotherapy), MWA (microwave ablation), *n* (count), PIPAC (Pressurized Intra-Peritoneal Aerosol Chemotherapy), RFA (radiofrequency ablation).

Of the 281 patients, 57% were female, 73% had a right-sided tumour. Age had a bimodal distribution around 50 and 70 years. Of patients with a known *BRAF* mutation status, 55% (*n*=82/150) had a *BRAF* mutation. A primary tumour resection and metastasectomy was performed in 78% and 23% of patients, respectively. Of patients with a known WHO performance score at start of first line treatment, 93% (*n*=102/109) had a WHO performance score of 0-1, with an unknown performance score in 64 patients.

In our cohort, 26% (*n*=72) of patients received no anti-tumour treatment, 13% (*n*=36) received local treatment, 14% (*n*=38) received local and systemic treatment and 48% (*n*=135) received only systemic treatment (Table 1). Sixty-two percent of patients received first line, 26% second line and 8% third line systemic non-immunotherapy. The type of treatment regimens and agents used per treatment line are described in Supplementary Table 2.

Comparing patient characteristics between population-based (2015-2018) and trial-based patients (2002-2011): population-based patients were older at diagnosis of metastatic disease (age above 75 years in 24% versus 10%), had less primary tumour resections (74% versus 96%) and more often resection of metastases (26% versus 9%), as shown in Supplementary Table 3. All trial-based patients were recruited from intervention trials with first line systemic therapy, which is reflected in the proportion of patients receiving systemic therapy (100% trial-based versus 52% population-based).

## Follow-up

For the population-based and trial-based cohort, the median follow-up period from diagnosis of metastatic disease was 8.3 months (inter-quartile range [IQR] 3.1–17.6) and 16.8 months (9.5 – 22.8), respectively. At the end of the follow-up period, 70.1% of patients were deceased, which was an indication that follow-up was adequate. Follow-up was completed for 264 patients (94%).

## OS during treatment

The Kaplan-Meier survival curves from diagnosis of metastatic disease and from start of each therapy line are demonstrated in Figure 2. We describe survival for patients who received anti-tumour therapy (systemic treatment, excluding adjuvant therapy, surgical resection of metastases and/or local treatment of metastases (including HIPEC or PIPAC)) versus patients who did not receive anti-tumour therapy. For patients who received anti-tumour therapy, median OS was 16.0 months (95% Confidence Interval [C.I.] 13.8–19.6;  $n=207$ ), whereas OS was 2.5 months (95% C.I. 1.8–3.5;  $n=72$ ) in patients without anti-tumour therapy. Furthermore, examining survival per type of treatment received, median OS was longer in patients receiving local and systemic therapy (median OS 29.9 months, 95% C.I. 17.9-not reached;  $n=38$ ) compared to patients receiving only systemic therapy (13.9 months, 95% C.I. 11.4–16.5;  $n=133$ ), as shown in Supplementary Table 4. Compared to the other treatment categories, patients receiving local and systemic treatment more often had primary rectal tumours (19% compared to <7% in other categories), *BRAF* wildtype tumours (65% compared to <49% in other categories) and were younger ( $\leq 55$  years in 42% versus <23% in other categories), as shown in Supplementary Table 5. Median OS increased in patients who received more systemic treatment lines (Supplementary Table 4).

**Figure 2.** OS, OS1 and OS2 in mCRC patients with dMMR

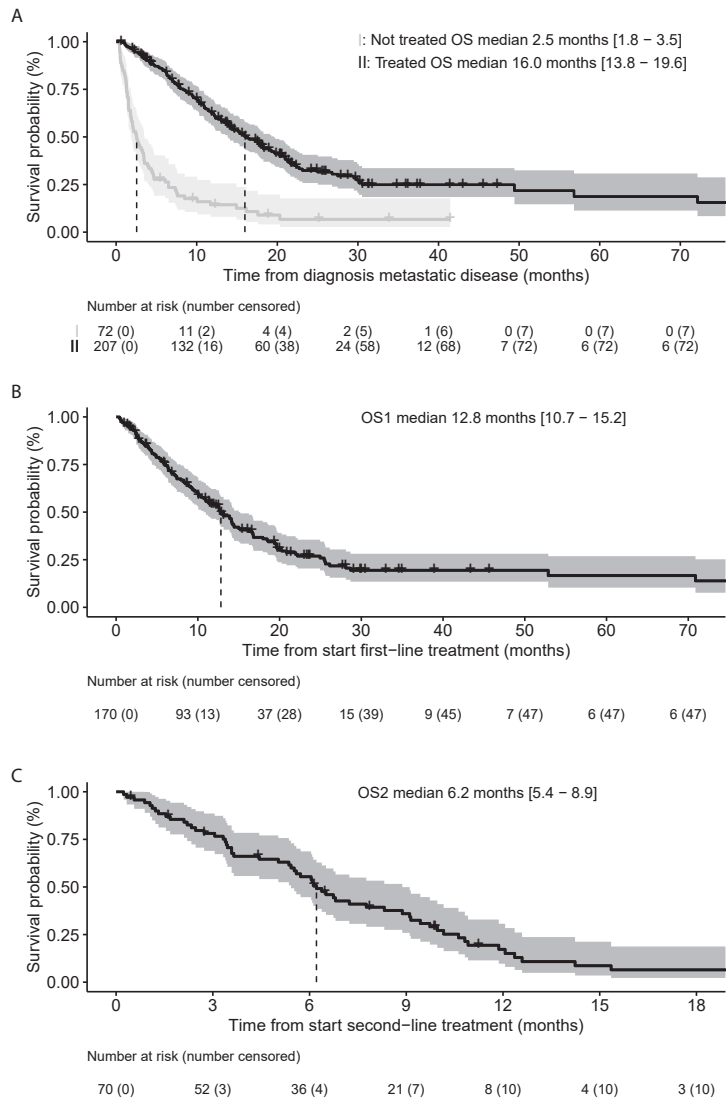


Fig 2. Predicted overall survival (OS) using Kaplan-Meier curves with confidence intervals, examined for the cohort as a whole and in patients who received systemic treatment. A) OS from metastatic disease in the unselected dMMR cohort ( $n=279$ ), stratified by having received anti-tumour therapy ( $n=207$ ) versus no anti-tumour therapy ( $n=72$ ), B) Overall survival from first line systemic therapy initiation (OS1) in first line patients ( $n=170$ ), C) Overall survival from second line systemic therapy initiation (OS2) in second line patients ( $n=70$ ). The risk tables display the number of patients at risk and the number of censored patients. Median survival is indicated with a vertical dashed line in each plot and is indicated with 95% Confidence Interval lower and upper ranges in the legend.

In order to examine the OS from initiation of each treatment line, comprising only patients receiving the given treatment line, we examined the OS1, OS2 and OS3. The median overall survival from first line systemic therapy initiation (OS1) was 12.8 months (95% C.I. 10.7–15.2;  $n=170$ ), from second line systemic therapy initiation (OS2) 6.2 months (95% C.I. 5.4–8.9;  $n=70$ ) and from third line systemic therapy initiation (OS3) 3.6 months (95% C.I. 2.7– not reached;  $n=19$ ).

From first line systemic therapy initiation (OS1), estimated 9-month and 12-month survival rates were 63.6% (95% C.I. 56.6–71.5) and 53.8% (95% C.I. 46.5–62.1). Similarly, from second line initiation (OS2), 9-month and 12-month survival rates were 35.9% (95% C.I. 25.9–49.9) and 17.2% (95% C.I. 9.7–30.7). In patients who received second line systemic treatment with a known WHO performance score  $\leq 1$ , similar to the inclusion criteria of the CheckMate 142 trials ( $n=24$ )<sup>7,8</sup>, predicted 9-month and 12-month survival rates from start of second line systemic therapy were 43.6% (95% C.I. 27.4–69.4) and 17.4% (95% C.I. 7.2–42.4), respectively. Supplementary Table 6 demonstrates the characteristics of our second line patients alongside the nivolumab and nivolumab/ipilimumab CheckMate 142 trial cohorts.

We compared the survival between population-based and trial-based dMMR patients, examining differences in OS among treated patients, OS1 and OS2. There was no significant difference in log-rank comparison of population-based versus trial-based patients in OS, OS1 or OS2. The median OS was 16.0 months (95% C.I. 13.0–22.1;  $n=155$ ) and 16.8 months (95% C.I. 13.5–21.0;  $n=52$ ;  $p=0.27$ ) in population-based and trial-based patients, respectively. Similarly, the median OS1 was 12.6 months (95% C.I. 10.1–15.0;  $n=118$  population-based) and 13.5 months (95% C.I. 9.1–19.6;  $n=52$  trial-based;  $p=0.59$ ). The median OS2 was 6.1 months (95% C.I. 4.4–8.9) in 43 population-based patients and 6.7 months (95% C.I. 5.0–10.2;  $p=0.58$ ) in 27 trial-based patients.

For population-based patients, we compared the median OS between patients with dMMR tumours versus pMMR tumours. For trial-based patients this was previously published, showing that the median OS was shorter in dMMR versus pMMR trial-based patients<sup>1,17</sup>. For population-based patients, the median OS was significantly shorter in patients with dMMR tumours upon receiving anti-tumour therapy than pMMR tumours (Table 2). The median OS was 7.6 months shorter, with a median OS of 16.0 months (95% C.I. 13.0–22.1;  $n=155$ ) for dMMR tumours compared to 23.6 months (95% C.I. 22.6–24.6;  $n=2746$ ;  $p<0.005$ ) for pMMR tumours in patients

receiving treatment. In untreated patients, median OS was 2.5 months (95% C.I. 1.8–3.5;  $n=72$ ) for dMMR tumours versus 3.9 months (95% C.I. 3.4–4.8;  $n=610$ ;  $p = 0.005$ ) for pMMR tumours.

**Table 2.** Overall survival in population-based metastatic colorectal cancer patients

Anti-tumour therapy	Mismatch repair status	Median OS (months)	<i>p</i> -value
<b>Treated</b>	dMMR ( $n = 155$ )	16.0 [13.0 – 22.1]	<b>&lt;0.005</b>
	pMMR ( $n = 2746$ )	23.6 [22.6 – 24.6]	
<b>Untreated</b>	dMMR ( $n = 72$ )	2.5 [1.8 – 3.5]	<b>0.005</b>
	pMMR ( $n = 610$ )	3.9 [3.4 – 4.8]	

Population-based patients with metastatic colorectal cancer and known mismatch repair status, registered between 2015 and 2017 in the Netherlands Cancer Registry. OS was measured from diagnosis of metastatic disease until vital status coupling and is indicated with the 95% confidence interval. Log-rank test was performed for untreated patients (dMMR versus pMMR) and for treated patients (dMMR versus pMMR). Anti-tumour therapy was defined as systemic treatment (with exception of adjuvant chemotherapy), metastasectomy or local metastatic treatment (RFA, MWA or HIPEC/PIPAC).

*Abbreviations:* dMMR (deficient mismatch repair), HIPEC (hyperthermic intraperitoneal chemotherapy), MWA (microwave ablation), OS (overall survival), PIPAC (pressurized intraperitoneal aerosol chemotherapy), pMMR (proficient mismatch repair), RFA (radio-frequency ablation).

## Prognostic variables associated with overall survival in patients receiving first line systemic treatment

In univariate analysis, metastasectomy and sidedness were significantly associated with OS from diagnosis of metastatic disease in patients receiving first line systemic treatment (Table 3). *BRAF*-mutational status had a higher risk for shorter OS, albeit non-significant, in univariate analysis (unadjusted hazard ratio [HR] 1.51 (95% C.I. 1.00–2.28);  $p=0.052$ ) as shown in Table 3.

**Table 3.** Univariate and multivariable Cox regression models for overall survival in first line treated patients

Variable	Level	Univariate		Multivariable	
		n	OS in months [95% C.I.] <sup>KM</sup>	Hazard ratio (95% C.I.) <sup>CPH</sup>	p-value
First line patients					
Age	≤ 65 years	83	18.0 [13.9 - 22.4]		
	> 65 years	88	13.8 [10.9 - 17.6]	1.409 [0.979 - 2.027]	0.065
Gender	Male	86	15.9 [12.3 - 21.0]		
	Female	85	14.6 [11.8 - 18.0]	1.238 [0.861 - 1.782]	0.249
BRAF	Wildtype	63	19.6 [14.6 - 22.7]		
	Mutation	59	11.5 [8.7 - 17.9]	1.509 [0.998 - 2.282]	0.052
Metastatic sites	Hepatic	39	15.9 [12.2 - 23.1]		
	Extra-hepatic	85	14.6 [11.1 - 21.0]	1.012 [0.651 - 1.574]	0.957
	Peritoneal	47	15.2 [12.3 - 21.6]	1.017 [0.601 - 1.721]	0.950
				0.814 [0.447 - 1.484]	0.502
Metastectomy	No	136	14.1 [11.5 - 17.2]		
	Yes	35	29.5 [17.9 - NR]	0.454 [0.271 - 0.761]	<b>0.003**</b>
Primary tumour resection	No	45	14.1 [11.1 - 17.6]		
	Yes	126	16.5 [13.5 - 21.0]	0.664 [0.437 - 1.007]	0.054
Primary tumour location	Left-sided/ rectosigmoid/ rectum	53	21.6 [16.5 - 72.1]		
	Right-sided	114	13.8 [10.9 - 17.2]	1.663 [1.112 - 2.486]	<b>0.013**</b>
Stage at diagnosis	Stage I/II	16	14.6 [10.3 - NR]		
	Stage III	39	14.6 [7.1 - 20.8]	1.293 [0.655 - 2.550]	0.459
	Stage IV	113	16.0 [13.1 - 19.8]	0.980 [0.532 - 1.805]	0.948
	Population-based	119	15.2 [12.3 - 18.3]		
Cohort	Trial-based	52	16.8 [13.5 - 21.0]	1.096 [0.752 - 1.598]	0.633
	<2	100	13.8 [9.9 - 19.6]		
Treatment lines received	<2	71	16.0 [14.1 - 21.0]	1.121 [0.780 - 1.610]	0.538
	>2				

Multivariable results were calculated stratified per number of treatment lines received (which had violated the assumption of proportional hazards) using the imputed dataset. An unadjusted median overall survival from diagnosis metastatic disease was obtained from a Kaplan Meier curve stratified for the given variable using the original dataset. The counts (n) reflect the counts of the patients used in the survival analysis (non-imputed dataset) which may differ from the total cohort due to missing data in the variable or outcome. \*\* Indicates statistically significant hazard ratio's (p-value <0.05).

Abbreviations: 95% C.I. (95% confidence interval), CPH (Cox Proportionate Hazards Model), HR (hazard ratio), KM (Kaplan Meier), NR (not reached), OS (overall survival).

The final multivariable model was a stratified Cox regression model for the number of treatment lines received ( $\leq 2$  and  $> 2$ ) since the variable violated the proportional hazards assumption. In the stratified Cox regression model for number of treatment lines received, metastasectomy is significantly associated with a longer survival (hazard ratio [HR] 0.49 (95% C.I. 0.26–0.90);  $p < 0.05$ ) and right-sided tumour location is significantly associated with a shorter survival (HR 1.71 (95% C.I. 1.04–2.79);  $p < 0.05$ ) as shown in Table 3. In patients receiving first line systemic treatment, the unadjusted median OS was 29.5 months (95% C.I. 17.9– not reached;  $n = 35$ ) and 14.1 months (95% C.I. 11.5–17.2;  $n = 136$ ) for patients with and without a metastasectomy, respectively (Table 3). Similarly, the unadjusted median OS was 13.8 months (95% C.I. 10.9–17.2;  $n = 114$ ) versus 21.6 months (95% C.I. 16.5–27.1;  $n = 53$ ) in patients receiving first line systemic treatment with a right-sided versus left-sided primary tumour location. *BRAF*-mutational status had a higher risk for shorter OS, albeit non-significant, in multivariable analysis (HR 1.61 (95% C.I. 0.94–2.77);  $p = 0.086$ ), with an unadjusted median OS of 11.5 months (95% C.I. 8.7–17.9;  $n = 59$ ) in patients receiving first line systemic treatment with a *BRAF*-mutant tumour versus 19.6 months (95% C.I. 14.6–22.7;  $n = 63$ ) in patients with a *BRAF*-wildtype tumour. The other covariates were not significantly associated with survival in the final multivariable model.

## DISCUSSION

We present survival data of a large, comprehensive cohort of dMMR mCRC patients, not treated with immunotherapy. Our cohort offers a unique insight into the survival of dMMR mCRC patients while receiving systemic non-immunotherapy in first-, second- and third line treatment. The OS in our dMMR mCRC cohort for all patients and patients receiving first line treatment is comparable to previously reported survival data in dMMR mCRC patients without immunotherapy, including two population-based dMMR mCRC cohorts with a similar percentage of patients receiving systemic therapy<sup>5,9,10,12,21</sup>. However the OS in our dMMR mCRC patients is shorter than the median OS in three other publications, which ranged from 26–39 months<sup>3,11,22</sup>. The difference may be due to the patient characteristics in the cohorts, with cohorts including patients receiving immunotherapy<sup>22</sup>, a high proportion (44–63%) of Lynch syndrome (*BRAF*-wildtype) patients<sup>3,22</sup>, and a high proportion (57%) of patients who underwent a metastasectomy<sup>11</sup>. The median OS from initiation of second line treatment (6.2 months) in our cohort is drastically shorter than the recently reported median OS in second line patients (21.6 months) in a cohort of dMMR mCRC patients

receiving systemic non-immunotherapy and immunotherapy<sup>22</sup>. The difference may be due to our cohort comprising only patients receiving systemic non-immunotherapy, while the Tougeron *et al.* dMMR mCRC patient cohort included patients receiving immunotherapy and a high proportion (44%) of Lynch syndrome patients<sup>22</sup>.

In our population-based patients, the median OS during treatment was significantly shorter in dMMR compared to pMMR mCRC patients. This supports previous studies reporting a worse survival in mCRC patients with dMMR<sup>1,5,12</sup> and is in contrast to studies showing a non-significant, null or opposite effect on survival<sup>9,11,21,23</sup>. The conflicting results may lie in heterogeneity of the population cohort being studied, with studies finding a null or opposite effect on survival having included patients with a low percentage of *BRAF* mutations<sup>9</sup>, only metachronous disease<sup>23</sup> or younger patients<sup>11</sup>. Our population reflects a clinically relevant cross-sectional population of Dutch patients who received MMR testing, indicating that patients with mCRC and known dMMR have a worse prognosis compared to pMMR patients.

In patients treated with at least one line of systemic treatment, we observed a significant association between metastasectomy with better survival and right-sided primary tumour location at diagnosis ('sidedness') with worse survival. In unselected mCRC patients and dMMR mCRC patients, metastasectomy is a known prognostic factor for OS<sup>3,10,24</sup>. Sidedness is an important prognostic factor in mCRC patients. However, sidedness has not yet been shown to be associated with OS in patients with dMMR tumours<sup>3,24</sup>. Our results indicate that in dMMR mCRC patients, right-sidedness is associated with worse survival. Patients receiving first line systemic treatment with a *BRAF* mutation had a higher risk for shorter survival in multivariable analysis, although non-significant, which is reflected with a 8 month difference in the unadjusted median OS in patients with a *BRAF*-mutant versus *BRAF*-wildtype tumour. Studies have demonstrated that *BRAF* mutational status was prognostic within dMMR mCRC patients<sup>10,23</sup>, however, this is not consistently shown<sup>5,22</sup>. Although we did not identify a significant association for *BRAF* mutation with survival in patients receiving first line systemic treatment, our results suggest that patients with a *BRAF* mutation do have a higher risk for shorter survival.

In addition to the prognostic factors which we examined in dMMR mCRC patients receiving first line systemic treatment, other factors may contribute to the worse prognosis seen in dMMR mCRC patients. Population-based dMMR mCRC patients

have a lower response rate to first line systemic non-immunotherapy compared to pMMR mCRC patients (5% versus 44%, respectively) and are also less likely to receive systemic therapy compared to pMMR mCRC patients (47% versus 73%,  $p<0.001$ )<sup>5,21</sup>. Thus, the worse prognosis in dMMR mCRC patients is likely driven by several factors, potentially including primary tumour sidedness, *BRAF* mutational status, the response rate to systemic therapy, the ability to receive a metastatic resection, and other less well known factors such as the *PD-L1* gene expression level, reflecting immune evasion<sup>25</sup>. Additionally, although dMMR mCRC patients are often analysed as one entity, different subgroups with different prognosis should be identified to compare survival results between studies including Lynch syndrome (often *BRAF*-wildtype), sporadic *BRAF* mutated dMMR tumours and sporadic *BRAF*-wildtype dMMR tumours<sup>3,9,10</sup>. Although *BRAF* status and MMR status was known, due to unavailable data regarding MLH1 methylation and Lynch syndrome status we were unable to distinguish between sporadic versus Lynch origin.

We are aware of several limitations. Although we were able to include a broad range of relevant variables, we cannot exclude confounding from unmeasured variables. Secondly, our retrospective study design may have resulted in a selection of the population, since MMR status was not determined in all patients in daily practice. Lastly, comparison of our data with other studies may be confounded by differences in patient characteristics.

Our study is unique in providing survival data on a large cohort of population-based and trial-based dMMR mCRC patients in the pre-immunotherapy era. Our survival data of dMMR mCRC patients beyond first line treatment may be compared with for instance the CheckMate 142 trial results, which examined nivolumab and nivolumab/ipilimumab treatment in dMMR mCRC patients beyond first line treatment<sup>7,8</sup>. The 9-month and 12-month survival rates of patients receiving second line treatment in our cohort of 35.9% and 17.2%, respectively, are lower than the published 9-month survival rate in the nivolumab/ipilimumab arm of 87%, the 12-month survival rates in the nivolumab arm of 73% and the nivolumab/ipilimumab arm of 85%<sup>7,8</sup>. The cohorts are comparable in key patient and tumour characteristics, although the immunotherapy cohorts were more heavily pre-treated (40-54% receiving  $\geq 3$  treatment lines in the CheckMate 142 trials compared to 29% in our cohort) and patients more often having *BRAF*-wildtype status. Both characteristics may reflect a patient selection in the CheckMate 142 trials with a less aggressive clinical course

compared to our cohort. Still, even the heavily treated patients in our cohort (who had received  $\geq 3$  treatment lines) had a median OS of only 18 months. However, as we had no access to individual patient data of the other cohorts, a direct comparison between the cohorts was not possible. A comparison with the phase II pembrolizumab trial was not possible due to the low number of patients in our cohort who received third line treatment<sup>6</sup>. The CheckMate 142 results for patients receiving nivolumab/ipilimumab in first line setting suggest a benefit of immunotherapy compared to our systemic non-immunotherapy first line cohort, with a median 12-month OS rate of 83% versus 54%, respectively<sup>26</sup>. This is supported by the Keynote-177 phase III randomized controlled trial results, which show that dMMR mCRC patients have a PFS benefit when receiving first line pembrolizumab versus first line systemic therapy (mFOLFOX6 or FOLFIRI combined with bevacizumab or cetuximab), with a median PFS of 16.5 months versus 8.2 months, respectively (HR 0.60, 95% C.I. 0.45-0.80,  $p=0.0002$ )<sup>27</sup>.

In conclusion, we present survival data on a large, comprehensive cohort of dMMR mCRC patients, treated with or without systemic non-immunotherapy. Currently, available data from immunotherapy trials lack a control arm with standard systemic treatment. We demonstrate a poor prognostic value for dMMR in mCRC patients. Given the poor outcome in our data set compared to the results of immunotherapy in dMMR mCRC patients, our data strongly supports a survival benefit of immunotherapy in dMMR mCRC patients.

## ACKNOWLEDGEMENTS

Part of the results were included in an abstract which was accepted for an oral presentation at the 2019 AACR Annual Meeting<sup>28</sup>. The following oncologists and research nurses were involved in the study: M. Bax (Maxima Medical Center), M. Beusink (Amsterdam University Medical Center), C. Bresser-de Ruyter (Rode Kruis Hospital Beverwijk), S. Brouwer (Hospital Rijnstate), A. Cats (MD, Stichting Antoni van Leeuwenhoek Hospital), M. de Buck (Zorgsaam), H. de Graaf (MD, Medical Center Leeuwarden, Leeuwarden), M. Deelen (Diaconessenhuis Hospital Utrecht), M. Fellingier (Ziekenhuisgroep Twente), N. Golsteijn (Hospital Rivierenland), S.M. Hiddema (Groene Hart Hospital), D.F.S. Kehrer (MD, IJsselland Hospital), M. Laven (Catharina Hospital), H. Polderdijk (Haga Hospital), F. Ramakers (Zuyderland Medical Center), C.J. Rienks-Bosma (Medical Center Leeuwarden), R. Roukema (Antonius Hospital Sneek), S. Ruijgrok (IJsselland Hospital), J. Schellekens-van Bronswijk (Bravis Hospital), T. Simon (LangeLand Hospital), S. Sloof (Gelre Hospital), D.J. Storm (Spaarne Hospital), E. Valenteijn (Zaans Medical Center), S. van Broekhoven (Franciscus Gasthuis & Vlietland), A.J. van de Vendel (Hospital Gelderse Vallei), E.J.E. van Gestel-Wink (Meander Medical Center), M.-L. van Groesen (OLVG), I. van Rooij-Tieleman (VieCuri Medical Center), M. Vercoulen (Elkerliek Hospital), M.J. Weterman (Amsterdam University Medical Center, University of Amsterdam). The authors thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry as well as IKNL staff for scientific advice.

## REFERENCES

1. Venderbosch S., Nagtegaal I.D., Maughan T.S., Smith C.G., Cheadle J.P., Fisher D., et al. Mismatch repair status and BRAF mutation status in metastatic colorectal cancer patients: A pooled analysis of the CAIRO, CAIRO2, COIN, and FOCUS studies. *Clin Cancer Res.* (2014);**20**(20):5322–30.
2. Gelsomino F., Barbolini M., Spallanzani A., Pugliese G., Cascinu S. The evolving role of microsatellite instability in colorectal cancer: A review. *Cancer Treat Rev.* (2016);**51**:19–26.
3. Cohen R., Buhard O., Cervera P., Hain E., Dumont S., Andre T. Clinical and molecular characterisation of hereditary and sporadic metastatic colorectal cancers harbouring microsatellite instability/DNA mismatch repair deficiency. *Eur J Cancer.* (2017);**86**:266–74.
4. Kim C.G., Ahn J.B., Jung M., Beom S.H., Kim C., Kim J.H., et al. Effects of microsatellite instability on recurrence patterns and outcomes in colorectal cancers. *Br J Cancer.* (2016);**115**(1):25–33.
5. Aasebø K., Dragomir A., Sundström M., Mezheyski A., Edqvist P.H., Eide G.E., et al. Consequences of a high incidence of microsatellite instability and BRAF-mutated tumors: A population-based cohort of metastatic colorectal cancer patients. *Cancer Med.* (2019);**8**(7):3623–35.
6. Le D.T., Uram J.N., Wang H., Bartlett B.R., Kemberling H., Eyring A.D., et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *NEJM.* (2015);**372**(26):2509–20.
7. Overman M.J., McDermott R., Leach J.L., Lonardi S., Lenz H., Morse M.A., et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *Lancet.* (2017);**18**(9):1182–91.
8. Overman M.J., Lonardi S., Yeung K., Wong M., Lenz H., Gelsomino F., et al. Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. *JCO.* (2018);**36**(8):773–9.
9. Alex A.K., Siqueira S., Coudry R., Santos J., Alves M., Hoff P.M., et al. Response to Chemotherapy and Prognosis in Metastatic Colorectal Cancer With DNA Deficient Mismatch Repair. *Clin Colorectal Cancer.* (2017);**16**(3):228–39.
10. Goldstein J., Tran B., Ensor J., Gibbs P., Wong H.L., Wong S.F., et al. Multicenter retrospective analysis of metastatic colorectal cancer (CRC) with high-level microsatellite instability (MSI-H). *Ann Oncol.* (2014);**25**:1032–8.
11. Jin Z., Sanhueza C.T., Johnson B., Nagorney David M., Larson D.W., Mara K.C., Harmsen W.C., et al. Outcome of Mismatch Repair-Deficient Metastatic Colorectal Cancer: The Mayo Clinic Experience. *Oncologist.* (2018);**23**(9):1083–91.
12. Tran B., Kopetz S., Tie J., Gibbs P., Jiang Z. Impact of BRAF Mutation and Microsatellite Instability on the Pattern of Metastatic Spread and Prognosis in Metastatic Colorectal Cancer. *Cancer.* (2011);**117**:4623–32.
13. Koopman M., Antonini N.F., Douma J., Wals J., Honkoop A.H., Erdkamp F.L.G., et al. Sequential versus combination chemotherapy with capecitabine, irinotecan, and oxaliplatin in advanced colorectal cancer (CAIRO): a phase III randomised controlled trial. *Lancet.* (2007);**370**:135–42.

14. Tol J., Koopman M., Rodenburg C.J., Cats A., Creemers G.J., Schrama J.G., et al. A randomised phase III study on capecitabine, oxaliplatin and bevacizumab with or without cetuximab in first-line advanced colorectal cancer, the CAIRO2 study of the Dutch Colorectal Cancer Group (DCCG). An interim analysis of toxicity. *Ann Oncol.* (2008);**19**(4):734–8.
15. Simkens L.H.J., Van Tinteren H., May A., Ten Tije A.J., Creemers G.J.M., Loosveld O.J.L., et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): A phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet.* (2015);**385**(9980):1843–52.
16. Koopman M., Kortman G.A.M., Mekenkamp L., Ligtenberg M.J.L., Hoogerbrugge N., Antonini N.F., et al. Deficient mismatch repair system in patients with sporadic advanced colorectal cancer. *Br J Cancer.* (2009);**100**(2):266–73.
17. Goey K.K.H., Elias S.G., van Tinteren H., Laclé M.M., Willems S.M., Offerhaus G.J.A., et al. Maintenance treatment with capecitabine and bevacizumab versus observation in metastatic colorectal cancer: Updated results and molecular subgroup analyses of the phase 3 CAIRO3 study. *Ann Oncol.* (2017);**28**(9):2128–34.
18. Sjoquist K.M., Renfro L.A., Simes R.J., Tebbutt N.C., Clarke S., Seymour M.T., et al. Personalizing Survival Predictions in Advanced Colorectal Cancer: The ARCAD Nomogram Project. *JNCI.* (2018);**110**:638–48.
19. van Buuren S., Groothuis-Oudshoorn K., Vink G., Doove L., Jolani S., Schouten R., et al. mice: Multivariate Imputation by Chained Equations in R. R package. *J Stat Softw.* (2011);**45**(3):1–67.
20. Team R.C. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2018. p. <https://www.R-project.org/>.
21. Chong L.C., Townsend A.R., Young J., Roy A., Piantadosi C., Hardingham J.E., et al. Outcomes for Metastatic Colorectal Cancer Based on Microsatellite Instability: Results from the South Australian Metastatic Colorectal Cancer Registry. *Target Oncol.* (2019);**14**(1):85–91.
22. Tougeron D., Sueur B., Zaanan A., de la Fouchardi re C., Sefrioui D., Lecomte T., et al. Prognosis and chemosensitivity of deficient MMR phenotype in patients with metastatic colorectal cancer: An AGEO retrospective multicenter study. *Int J Cancer.* (2020);**147**:285–96.
23. Taieb J., Shi Q., Pederson L., Alberts S., Wolmark N., Cutsem E. Van, et al. Prognosis of microsatellite instability and / or mismatch repair deficiency stage III colon cancer patients after disease recurrence following adjuvant treatment: results of an ACCENT pooled analysis of seven studies. *Ann Oncol.* (2019);**30**:1466–71.
24. Boeckx N., Janssens K., Van Camp G., Rasschaert M., Papadimitriou K., Peeters M., et al. The predictive value of primary tumor location in patients with metastatic colorectal cancer: A systematic review. *Crit Rev Oncol Hematol.* (2018);**121**:1–10.
25. Marisa L., Svrcek M., Collura A., Becht E., Cervera P., Wanherdrick K., et al. The Balance Between Cytotoxic T-cell Lymphocytes and Immune Checkpoint Expression in the Prognosis of Colon Tumors. *JNCI.* (2018);**110**(1):68–77.
26. Lenz H.-J., Lonardi S., Zagonel V., Van Cutsem E., Limon M.L., Yeung K., et al. Nivolumab plus low-dose ipilimumab as first-line therapy in microsatellite instability-high/DNA mismatch repair deficient metastatic colorectal cancer: Clinical update. *JCO.* (2020);**38**(Suppl 4):abstr 11.

27. André T., Shiu K.-K., Kim T.W., Jensen B.V., Jensen L.H., Punt C.J.A., et al. Pembrolizumab versus chemotherapy for microsatellite instability-high/mismatch repair deficient metastatic colorectal cancer: The phase 3 KEYNOTE-177 study. *JCO*. (2020);**38**(suppl):abstr LBA4.
28. Wensink G.E., Elferink M.A., May A.M., Mol L., Hamers P.A., Punt C.J., et al. Survival of patients with microsatellite instable (MSI) metastatic colorectal cancer (mCRC) upon systemic non-immunotherapy [abstract]. Proceedings of the American Association for Cancer Research Annual Meeting 2019; 2019 Mar 29-Apr 3; Atlanta, GA, Philad. *Cancer Res*. (2019);**79**(13 Suppl):Abstract nr 4467.

## SUPPLEMENTARY MATERIAL

**Supplementary Table 1.** Characteristics of incomplete and complete cases in first line treated patients

	Original dataset <i>n</i> =171	Complete cases <i>n</i> =122	Incomplete cases <i>n</i> =49
Age >65 years (%)	88 (51.5)	61 (50.0)	27 (55.1)
Trial-based (%)	52 (30.4)	50 (41.0)	2 (4.1)
Female (%)	85 (49.7)	59 (48.4)	26 (53.1)
<i>BRAF</i> (%)			
<i>BRAF</i> mutation	59 (34.5)	59 (48.4)	
Unknown	49 (28.7)		49 (100.0)
Stage (%)			
I/II	17 (9.9)	12 (9.8)	5 (10.2)
III	39 (22.8)	28 (23.0)	11 (22.4)
IV	113 (66.1)	81 (66.4)	32 (65.3)
Unknown	2 (1.2)	1 (0.8)	1 (2.0)
Sidedness (%)			
Left-sided	53 (31.0)	34 (27.9)	19 (38.8)
Right-sided	114 (66.7)	85 (69.7)	29 (59.2)
Unknown	4 (2.3)	3 (2.5)	1 (2.0)
Metastatic localization (%)			
Extra-hepatic	85 (49.7)	64 (52.5)	21 (42.9)
Liver-only	39 (22.8)	31 (25.4)	8 (16.3)
Peritoneal	47 (27.5)	27 (22.1)	20 (40.8)
Primary tumour resection (%)	126 (73.7)	90 (73.8)	36 (73.5)
Metastasectomy (%)	35 (20.5)	21 (17.2)	14 (28.6)
Adjuvant chemotherapy (%)	21 (12.3)	15 (12.3)	6 (12.2)
2 or more treatment lines (%)	71 (41.5)	60 (49.2)	11 (22.4)

Characteristics of first line patients in different cohorts are reported: A) all first line patients with known survival status from the original dataset (*n*=171), B) subgroup of patients with complete *BRAF* cases (*n*=122), and C) subgroup of patients with incomplete *BRAF* cases (*n*=49).

**Supplementary Table 2.** Systemic therapy regimen for patients

	<b>Cohort</b> <i>n=281</i>
<b>First line regimen</b> (% of cohort)	173 (61.6)
Monotherapy (% of first line)	35 (20.2)
Doublet	132 (76.3)
Triplet	5 (2.9)
Chemotherapeutic agents (% of first line)	
Capecitabine	34 (19.7)
CAPOX/FOLFOX	113 (65.3)
Irinotecan/CAPIRI/FOLFIRI	20 (11.6)
FOLFOXIRI	5 (2.9)
Other	1 (0.6)
Targeted agents (% of first line)	
Anti-VEGF	89 (51.4)
Anti-EGFR	3 (1.7)
Combined anti-VEGF + EGFR	13 (7.5)
<b>Second line regimen</b> (% of cohort)	72 (25.6)
Monotherapy (% of second line)	50 (69.4)
Doublet	21 (29.2)
Chemotherapeutic agents (% of second line)	
Capecitabine/UFT	2 (2.8)
CAPOX/FOLFOX	5 (6.9)
Irinotecan /CAPIRI/FOLFIRI	55 (76.4)
FOLFOXIRI	1 (1.4)
Other	1 (1.4)
Targeted agents (% of second line)	
Anti-VEGF	6 (8.3)
Anti-EGFR	11 (15.3)
<b>Third line regimen</b> (% of cohort)	21 (7.5)
Monotherapy (% of third line)	14 (66.7)
Doublet	5 (23.8)
Chemotherapeutic agents (% of third line)	
Capecitabine/UFT	1 (4.8)
Oxaliplatin/CAPOX/FOLFOX	6 (28.6)
Irinotecan/CAPIRI/FOLFIRI	1 (4.8)
Other	3 (14.3)
Targeted agents (% of third line)	
Anti-VEGF	1 (4.8)
Anti-EGFR	8 (38.1)

Characteristics of patient treatment regimens for the whole follow-up period, reported for all patients receiving systemic therapy for palliative disease, including trial and population-based patients.

*Abbreviations:* CAPIRI (capecitabine and irinotecan), CAPOX (capecitabine and oxaliplatin), EGFR (epidermal growth factor receptor), FOLFIRI (5-fluoro-uracil and irinotecan), FOLFOX (5-fluoro-uracil and oxaliplatin), FOLFOXIRI (5-fluoro-uracil, oxaliplatin and irinotecan), UFT (tegafur/uracil), and VEGF (vascular endothelial growth factor).

**Supplementary Table 3.** Characteristics of population- and trial-based patients

		Population-based <i>n</i> =227	Trial-based <i>n</i> =54	<i>p</i> -value
<b>Age (%)</b>	≤ 55 years	47 (20.7)	12 (23.1)	0.120
	56-65 years	43 (18.9)	14 (26.9)	
	66-75 years	82 (36.1)	21 (40.4)	
	> 75 years	55 (24.2)	5 (9.6)	
<b>Female (%)</b>		134 (59.0)	25 (46.3)	0.123
<b>BRAF mutational status (%)</b>	Wildtype	42 (42.0)	26 (52.0)	0.324
	Mutation	58 (58.0)	24 (48.0)	
	Unknown	127	4	
<b>RAS mutational status (%)</b>	Wildtype	61 (78.2)	43 (86.0)	0.384
	Mutation	17 (21.8)	7 (14.0)	
	Unknown	149	4	
<b>Stage (%)</b>	I	1 (0.4)	1 (1.9)	0.141
	II	19 (8.4)	9 (17.0)	
	III	73 (32.3)	15 (28.3)	
	IV	133 (58.8)	28 (52.8)	
<b>Sidedness (%)</b>	Right-sided	160 (71.4)	42 (79.2)	0.565
	Left-sided	47 (21.0)	8 (15.1)	
	Rectosigmoid/ Rectum	17 (7.6)	3 (5.7)	
<b>Synchronous metastatic pattern (%)</b>	Synchronous	158 (69.6)	36 (66.7)	0.798
<b>Metastatic localization (%)</b>	Liver-only	46 (20.3)	15 (27.8)	<0.001**
	Extra-hepatic	97 (42.7)	36 (66.7)	
	Peritoneal	84 (37.0)	3 (5.6)	
<b>Number of metastatic sites (%)</b>	1	142 (62.6)	24 (45.3)	0.105
	2	53 (23.3)	18 (34.0)	
	3	24 (10.6)	9 (17.0)	
	≥4	8 (3.5)	2 (3.8)	
<b>Primary tumour resection (%)</b>		167 (73.6)	52 (96.3)	<0.001**
<b>Metastasectomy (%)</b>		59 (26.0)	5 (9.3)	0.014**
<b>Local treatment metastases (%)</b>	RFA	5 (2.2)	1 (1.9)	0.012**
	MWA	1 (0.4)	0 (0.0)	
	HIPEC	27 (11.9)	0 (0.0)	
<b>Anti-tumour therapy (%)</b>	No treatment	72 (31.7)	0 (0.0)	<0.001**
	Local treatment	36 (15.9)	0 (0.0)	
	Local and systemic treatment	33 (14.5)	5 (9.3)	
	Systemic treatment	86 (37.9)	49 (90.7)	

		Population-based <i>n</i> =227	Trial-based <i>n</i> =54	<i>p</i> -value
<b>Adjuvant chemotherapy (%)</b>		39 (17.2)	8 (14.8)	0.829
<b>Systemic therapy</b>	1 <sup>st</sup> line (%)	119 (52.4)	54 (100.0)	<b>&lt;0.001**</b>
	2 <sup>nd</sup> line (%)	43 (18.9)	29 (53.7)	<b>&lt;0.001**</b>
	3 <sup>rd</sup> line (%)	12 (5.3)	9 (16.7)	<b>0.010**</b>
	4 <sup>th</sup> line (%)	3 (1.3)	0 (0.0)	1.000

Characteristics of patients at diagnosis of metastatic disease with treatment information during the course of disease. Trial-based patients were obtained from the CAIRO (n=19), CAIRO2 (n=31) and CAIRO3 (n=4) phase III randomized controlled trials. \*\*Indicates statistically significant hazard ratios (p-value <0.05). Sidedness of the primary tumour was defined as right-sided (coecum-transverse colon), left-sided (splenic flexure-sigmoid) and rectosigmoid/rectal. Local treatment was defined as metastasectomy or local metastatic treatment (RFA, MWA or HIPEC/PIPAC) and systemic therapy as all systemic treatment given for metastatic disease (excluding adjuvant therapy). Missing values are not shown if missing frequency was less than 5%. *p*-values are shown for comparison between population-based and trial-based patients (default test is the Chi-squared test, except for variables with <5 events per cell, where a Fisher's Exact test was performed).

**Abbreviations:** HIPEC (Hyperthermic intraperitoneal chemotherapy), MWA (microwave ablation), *n* (count), PIPAC (Pressurized Intra-Peritoneal Aerosol Chemotherapy), RFA (radiofrequency ablation).

**Supplementary Table 4.** Overall survival per treatment group and treatment lines received

		Overall Survival (median months with 95% C.I.)
All patients	<i>n</i> = 279	11.8 [10.1 - 14.6]
No tumour-directed therapy	<i>n</i> = 72	2.5 [1.8 - 3.5]
Tumour-directed therapy	<i>n</i> = 207	16.0 [13.8 - 19.6]
Received local treatment	<i>n</i> = 36	NR [11.8 - NR]
Received local and systemic treatment	<i>n</i> = 38	29.9 [17.9 - NR]
Received systemic treatment	<i>n</i> = 133	13.9 [11.4 - 16.5]
Received first line systemic therapy	<i>n</i> = 171	15.3 [13.1 - 18.3]
Received second line systemic therapy	<i>n</i> = 70	16.0 [14.1 - 21.0]
Received third line systemic therapy	<i>n</i> = 21	18.0 [13.1 - 29.9]

Overall survival from diagnosis of metastatic disease until death or last follow-up if alive for patients with known survival data (*n*=2 missing in cohort). Tumour-directed treatment in metastatic setting was defined as systemic treatment, metastasectomy or local metastatic treatment (radio-frequency ablation (RFA), microwave ablation (MWA), hyperthermic intraperitoneal chemotherapy (HIPEC) or pressurized intraperitoneal aerosol chemotherapy (PIPAC)). Local treatment was defined as metastasectomy or local metastatic treatment (RFA, MWA or HIPEC/PIPAC). Systemic therapy is defined as all systemic treatment given for metastatic disease, excluding adjuvant systemic therapy.

**Abbreviations:** C.I. (Confidence Interval), NR (not reached).

**Supplementary Table 5.** Characteristics of patients per type of anti-tumour treatment received

		<b>No treatment</b> <i>n</i> =72	<b>Local treatment</b> <i>n</i> =36	<b>Local &amp; systemic treatment</b> <i>n</i> =38	<b>Systemic treatment</b> <i>n</i> =135
<b>Age (%)</b>	≤ 55 years	7 (9.7)	6 (16.7)	16 (42.1)	30 (22.6)
	56-65 years	11 (15.3)	9 (25.0)	9 (23.7)	28 (21.1)
	66-75 years	28 (38.9)	14 (38.9)	11 (28.9)	50 (37.6)
	> 75 years	26 (36.1)	7 (19.4)	2 (5.3)	25 (18.8)
<b>Female (%)</b>		49 (68.1)	24 (66.7)	19 (50.0)	67 (49.6)
<b><i>BRAF</i> mutational status (%)</b>	Wildtype	4 (20.0)	1 (12.5)	15 (65.2)	48 (48.5)
	Mutation	16 (80.0)	7 (87.5)	8 (34.8)	51 (51.5)
	Unknown	52	28	15	36
<b><i>RAS</i> mutational status (%)</b>	Wildtype	8 (80.0)	2 (66.7)	14 (70.0)	80 (84.2)
	Mutation	2 (20.0)	1 (33.3)	6 (30.0)	15 (15.8)
	Unknown	62	33	18	40
<b>Stage (%)</b>	I	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.7)
	II	7 (9.7)	5 (13.9)	2 (5.6)	14 (10.4)
	III	34 (47.2)	15 (41.7)	6 (16.7)	33 (24.4)
	IV	30 (41.7)	16 (44.4)	28 (77.8)	87 (64.4)
<b>Sidedness (%)</b>	Right-sided	57 (79.2)	29 (80.6)	21 (56.8)	95 (72.0)
	Left-sided	11 (15.3)	7 (19.4)	9 (24.3)	28 (21.2)
	Rectosigmoid/ Rectum	4 (5.6)	0 (0.0)	7 (18.9)	9 (6.8)
<b>Synchronous metastatic pattern (%)</b>		46 (63.9)	17 (47.2)	31 (81.6)	100 (74.1)
<b>Metastatic localization (%)</b>	Liver-only	13 (18.1)	9 (25.0)	13 (34.2)	26 (19.3)
	Extra-hepatic	35 (48.6)	12 (33.3)	12 (31.6)	74 (54.8)
	Peritoneal	24 (33.3)	15 (41.7)	13 (34.2)	35 (25.9)
<b>Number of metastatic sites (%)</b>	1	44 (61.1)	30 (83.3)	29 (76.3)	63 (47.0)
	2	18 (25.0)	4 (11.1)	5 (13.2)	44 (32.8)
	3	7 (9.7)	1 (2.8)	4 (10.5)	21 (15.7)
	≥4	3 (4.2)	1 (2.8)	0 (0.0)	6 (4.5)

Characteristics of patients at diagnosis of metastatic disease with treatment information during the course of disease. WHO PS, percentages relative to amount of people receiving first line treatment. Sidedness of the primary tumour was defined as right-sided (coecum-transverse colon), left-sided (splenic flexure-sigmoid) and rectosigmoid/rectal. Local treatment was defined as metastasectomy or local metastatic treatment (RFA, MWA, HIPEC or PIPAC). Missing values are not shown if missing frequency was less than 5%.

*Abbreviations:* HIPEC (hyperthermic intraperitoneal chemotherapy), MWA (microwave ablation), PIPAC (pressurized intraperitoneal aerosol chemotherapy), RFA (radio-frequency ablation), WHO PS (World Health Organisation Performance Score).

**Supplementary Table 6.** Comparison of non-immunotherapy versus immunotherapy cohorts

	<b>Current cohort: Second line patients <i>n</i>=72</b>	<b>CheckMate 142: Nivolumab arm <i>n</i>=74</b>	<b>CheckMate 142: Nivolumab/ Ipilimumab arm <i>n</i>=119</b>
<b>Age</b> (median [range])	63 [54-71]	53 [44-64]	58 [21-88]
<b>Metastatic disease diagnosis</b> (%)			
2002-2005	13 (18)		
2006-2010	14 (19)		
2011-2015	10 (14)	(100) 2014 - 2016	(100) 2015-2016
2016-2018	34 (47)		
Unknown	1 (1)		
<b>Female</b> (%)	31 (43)	30 (41)	49 (41)
<b>BRAF</b> (%)			
<i>BRAF</i> wildtype	39 (54)	55 (74)	74 (63)
<i>BRAF</i> mutation	21 (29)	12 (16)	29 (24)
Unknown	12 (17)	7 (9)	15 (13)
<b>WHO Performance score</b> (2 <sup>nd</sup> line; %)			
Score 0	11 (15)	32 (43)	54 (45)
Score 1	15 (21)	42 (57)	65 (55)
Unknown	46 (64)		
<b>Sidedness</b> (%)			
Right-sided	41 (57)	Unknown	65 (55)
Left-sided	20 (28)		45 (38)
Rectosigmoid/Rectum	9 (12)		6 (5)
Colon NOS	2 (3)		3 (3)
<b>Stage</b> (%)			
I	0 (0)	15 (20)	0 (0)
II	5 (7)	0 (0)	14 (12)
III	10 (14)	26 (35)	52 (44)
IV	56 (78)	33 (45)	53 (45)
Unknown	1 (1)		
<b>Trial participation</b> (%)	29 (40)	74 (100)	119 (100)
<b>Treatment lines given</b> (%)			
0 lines	-	1 (1)	1 (1)
1 line	-	11 (15)	27 (23)
2 lines	51 (71)	22 (30)	43 (36)
≥ 3 lines	21 (29)	40 (54)	48 (40)

Characteristics of the current cohort second line patients are displayed alongside the baseline characteristics of the CheckMate 142 nivolumab and nivolumab/ipilimumab cohorts<sup>7,8</sup>.

*Abbreviations:* colon NOS (colon not otherwise specified), WHO (World Health Organisation).



# A REVIEW OF THE SENSITIVITY OF METASTATIC COLORECTAL CANCER PATIENTS WITH DEFICIENT MISMATCH REPAIR TO STANDARD-OF-CARE CHEMOTHERAPY AND MONOCLONAL ANTIBODIES, WITH RECOMMENDATIONS FOR FUTURE RESEARCH

*Cancer Treatment Reviews*, 2021, 95, 102174.

Emerens Wensink  
Marinde Bond  
Emre Küçükköse  
Anne May  
Geraldine Vink  
Miriam Koopman  
Onno Kranenburg<sup>‡</sup>  
Jeanine Roodhart<sup>‡</sup>

<sup>‡</sup> Both authors share last authorship.

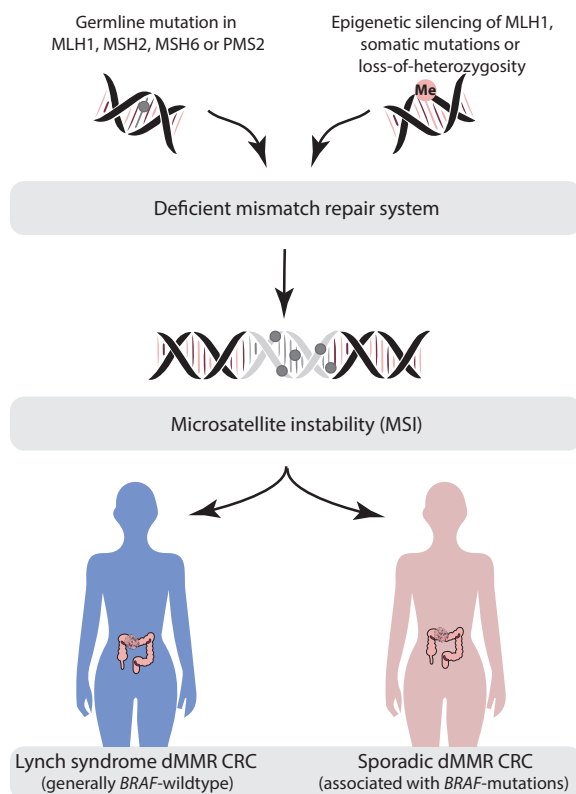
## ABSTRACT

In 5% of metastatic colorectal cancer (mCRC) patients, tumours display a deficient mismatch repair (dMMR) system. Immunotherapy is beneficial in dMMR mCRC patients and has recently been approved by the Food and Drug Administration for patients with unresectable or metastatic dMMR CRC. Although dMMR and proficient MMR (pMMR) CRC tumours are biologically distinct, they are commonly treated with the same chemotherapy and monoclonal antibodies. This includes dMMR mCRC patients who did not respond to immunotherapy (20-30%). However, it is unclear if these treatments are equally beneficial in dMMR mCRC. Of note, dMMR mCRC patients have a worse prognosis compared to pMMR, which may in part be caused by a lower response to treatment. To avoid unnecessary exposure to ineffective treatments and their associated toxicity, it is important to identify which systemic treatments are most beneficial in dMMR mCRC patients, thus improving their outcome. Indeed, future treatment strategies are likely to involve combinations of immunotherapy, chemotherapy and monoclonal antibodies. In this evidence-based review, we summarize clinical trials reporting treatment efficacy of different types of chemotherapy and monoclonal antibodies in dMMR mCRC patients. We also review the biological rationale behind a potential differential benefit of chemotherapy with or without monoclonal antibodies in dMMR mCRC patients. A barrier in the interpretation of preclinical results is the choice of model systems. They largely comprise traditional models, including cell lines and xenografts, rather than more representative models, such as patient-derived organoids. We provide concrete recommendations for clinical investigators and fundamental researchers to accelerate research regarding which systemic therapy is most effective in dMMR mCRC patients.

## INTRODUCTION

In 5% of metastatic colorectal cancer (mCRC) patients, tumours display a deficient mismatch repair (dMMR) system, resulting in accumulation of point-mutations and a microsatellite instability (MSI) phenotype<sup>1,2</sup>. dMMR arises through hereditary or sporadic inactivation of the mismatch repair (MMR) system, resulting in Lynch syndrome or sporadic dMMR CRC, respectively (Figure 1)<sup>2</sup>. The majority of Lynch syndrome patients inherit a germline mutation in one of the MMR genes, while in sporadic dMMR patients, inactivation is most frequently through epigenetic silencing, but can also occur through somatic mutations or loss-of-heterozygosity in the MMR system<sup>2-5</sup>.

**Figure 1.** Mechanisms of carcinogenesis for Lynch syndrome colorectal cancer (CRC) and sporadic deficient mismatch repair (dMMR) CRC



In contrast to early stage CRC, studies have reported a worse survival in patients with dMMR mCRC compared to proficient mismatch repair (pMMR) mCRC<sup>1,6-8</sup>. Accordingly, stage I – III dMMR CRC patients were less likely to have recurrent disease compared to

pMMR CRC patients, however, if disease did recur, OS was worse for dMMR CRC patients compared to pMMR<sup>9</sup>. Likewise, dMMR mCRC population-based patients receiving treatment had a median overall survival (OS) of 16.0 months compared to 23.6 months in pMMR<sup>8</sup>. In a population-based study, the response rate to first line chemotherapy with or without monoclonal antibodies was lower in dMMR mCRC patients compared to pMMR mCRC patients (5% versus 44%, respectively)<sup>7</sup>, suggesting that dMMR mCRC patients may benefit less from chemotherapy with or without monoclonal antibodies. This is supported by the recent KEYNOTE-177 results, where dMMR mCRC patients in the control arm (chemotherapy with/without targeted treatment) showed a lower response rate than expected in first line mCRC patients of 33.1% and progression-free survival (PFS) of 8.2 months<sup>10</sup>. The worse prognosis in dMMR mCRC patients is likely driven by several factors, including *BRAF* mutational status<sup>11,12</sup>, the ability to receive a metastatic resection<sup>8,13–15</sup> and other lesser known factors, such as the programmed cell death ligand 1 (*PD-L1*) gene expression level<sup>16</sup>. To avoid unnecessary exposure to ineffective treatments and their associated toxicity, it is important to identify which systemic treatments are most beneficial in dMMR mCRC patients, thus improving their outcome.

Recent publications report durable responses to immunotherapy in dMMR mCRC patients across all lines of treatment<sup>10,17–19</sup>. Immunotherapy has been approved for dMMR patients with metastatic or unresectable CRC by regulatory authorities (including the Food and Drug Administration (FDA) and the European Medicines Agency (EMA)). Therefore, the standard-of-care is increasingly becoming immunotherapy. However, chemotherapy with or without monoclonal antibodies will still be relevant in patients with dMMR mCRC. In the first line KEYNOTE-177 trial 29.4% of patients in the pembrolizumab arm had primary progressive disease compared to 12.3% in the chemotherapy control arm<sup>10</sup>. For patients in whom rapid downstaging of disease is desired, e.g. to allow resection of disease or based on poor prognostic features, first line chemotherapy with monoclonal antibodies may still be preferable over immunotherapy in selected cases to decrease the chance of having primary progression upon treatment. Furthermore, 12-month PFS rates ranged from 50 to 71% in the immune checkpoint inhibitor trials<sup>10,18–20</sup>, with a significant portion of patients requiring an alternative treatment during the course of their disease. Finally, potentially a subgroup of dMMR mCRC patients may benefit from combination chemotherapy and monoclonal treatment with immunotherapy and it is important to find the most effective combination treatment. With the exception of immunotherapy, dMMR mCRC patients receive the same chemotherapy with or without monoclonal antibodies as other mCRC patients. Importantly, it is unclear if these treatments are equally beneficial in dMMR mCRC or

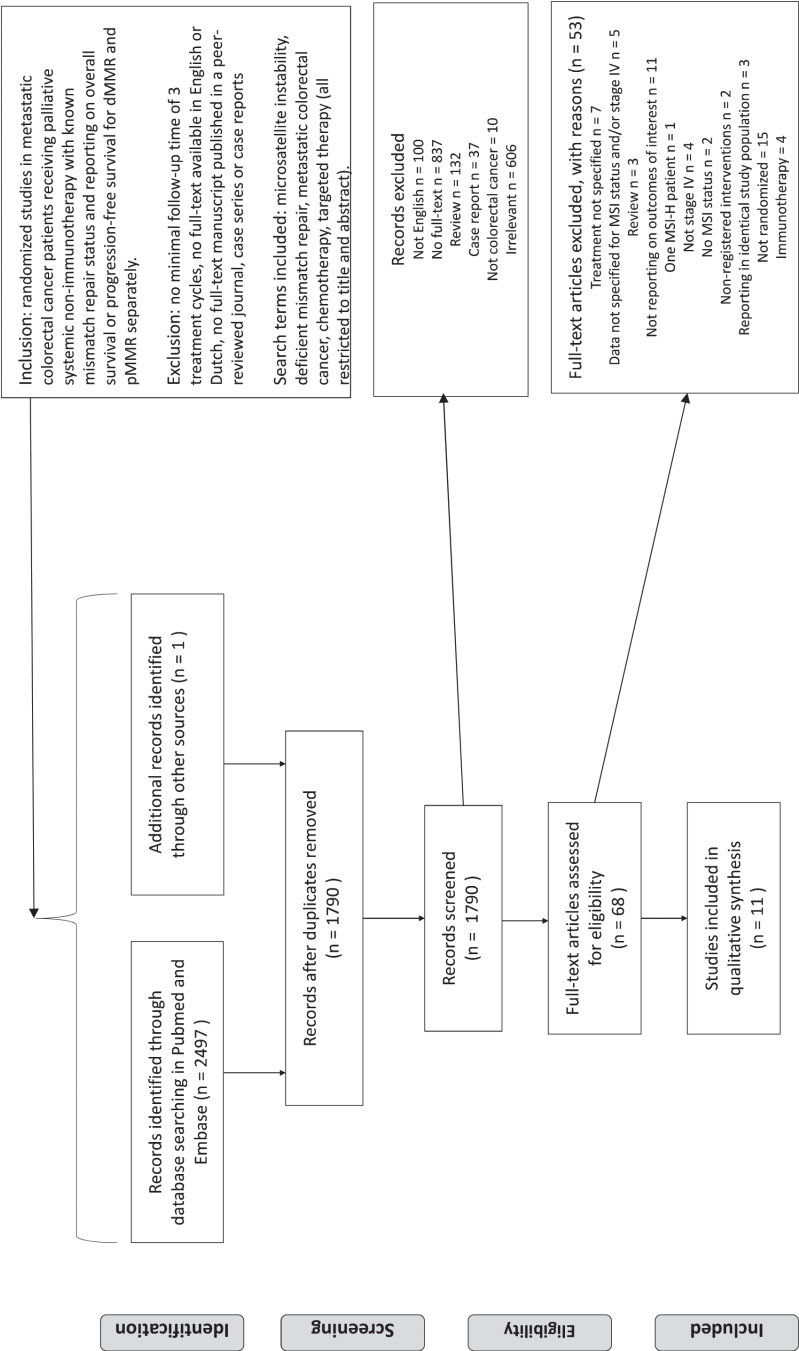
if a specific type of chemotherapy or monoclonal antibody may offer increased benefit. Thus, we focused on chemotherapy and/or monoclonal antibodies in this review since dMMR mCRC patients will still be treated with these agents either before or after immunotherapy, or in combination with receiving immunotherapy.

A published meta-analysis from 2009 and international CRC guidelines do not provide clarity regarding the choice and efficacy of chemotherapy agents for dMMR mCRC<sup>21,22</sup>. Clinical trials are increasingly reporting MMR molecular analysis subgroup results, warranting a new review of the literature. In this evidence-based review, we assess clinical trials reporting outcome of treatment in dMMR mCRC patients to different types of chemotherapy with or without monoclonal antibodies. Subsequently, to clarify the clinical findings, we summarize the biological rationale for a differential benefit of chemotherapy with or without monoclonal antibodies in dMMR mCRC patients. Finally, we provide recommendations for clinical investigators and researchers.

## **EFFICACY OF SYSTEMIC THERAPIES IN dMMR mCRC PATIENTS**

We performed a systematic literature search (up to April, 2020) in the PubMed and Embase databases, which included mCRC clinical trials reporting response rates and/or survival outcomes for dMMR mCRC patients receiving a given treatment. Search terms, inclusion and exclusion criteria are presented with the search results in Figure 2. We identified 11 trials which examined standard-of-care chemotherapy with or without monoclonal antibodies effectiveness in dMMR/MSI mCRC patients (ancillary studies<sup>1,23–31</sup> and original studies<sup>32–41</sup>). The study characteristics and treatments examined are described in Table 1. All identified studies included patients receiving first line treatment, often with doublet chemotherapy (5-fluorouracil [5-FU] and oxaliplatin or irinotecan) and targeted treatment (anti-epidermal growth factor (EGFR) or anti-vascular endothelial growth factor (VEGF)). Three studies also examined later treatment lines (first, second and third line)<sup>23,27,32,33,36</sup>. Two studies examined the effect of different first line maintenance treatments<sup>25,31,34,41</sup>, however survival data for dMMR mCRC patients was not reported per type of maintenance treatment and thus could not be analyzed. Immunotherapy trials were excluded as they are not within the scope of this review. The reported treatment efficacy for the different systemic therapy agents are summarized per study (Table 2). We will first discuss the results for studies across multiple treatment lines, subsequently the results for the effectivity of irinotecan-containing treatments, oxaliplatin-containing treatments and lastly, targeted therapies.

Figure 2. PRISMA flow diagram



**Table 1.** Description of included clinical studies in evidence-based review

Author	Study Design	mCRC (n)	dMMR (n)	pMMR (n)	Method MMR/MSI	Intervention
Braun <sup>23</sup> 2008	Phase III RCT (FOCUS <sup>33</sup> )	1313	58	1252	IHC (MLH1, MSH2)	5-FU, FOLFOX, FOLFIRI; sequential versus combination A) 1 <sup>st</sup> line 5-FU; 2 <sup>nd</sup> line irinotecan; B) 1 <sup>st</sup> line 5-FU, 2 <sup>nd</sup> line FOLFIRI or FOLFOX; C) 1 <sup>st</sup> line FOLFIRI or FOLFOX.
Chua <sup>24</sup> 2009	Phase II trials (3 studies <sup>42-44</sup> )	118	2	116	PCR-based (BAT-26)	1 <sup>st</sup> line FOLFOX
Cremolini <sup>32</sup> 2020	Phase III RCT (TRIBE2 <sup>32</sup> )	679	26	528	IHC (MLH1, PMS2, MSH2, MSH6)	A) 1 <sup>st</sup> line FOLFOX-B (8 cycles) then 5-FU-B maintenance, 2 <sup>nd</sup> line FOLFIRI-B B) 1 <sup>st</sup> line FOLFOXIRI-B (8 cycles) then 5-FU-B maintenance, 2 <sup>nd</sup> line FOLFOXIRI-B
Goey <sup>25</sup> 2017	Phase III RCT (CAIRO3 <sup>34</sup> )	558	4	275	IHC (MLH1, PMS2, MSH2, MSH6)	1 <sup>st</sup> line CAPOX-B +/- CAP-B maintenance, reintroduction CAPOX-B upon progression
Innocenti <sup>26</sup> 2019	Phase III RCT (CALGB <sup>35</sup> )	1585	52	755	PCR-based (Bethesda panel)	1 <sup>st</sup> line FOLFOX or FOLFIRI plus bevacizumab, cetuximab or both.
Koopman <sup>27</sup> 2009	Phase III RCT (CAIRO <sup>36</sup> )	515	18	497	1) All: IHC (MLH1, MSH2, PMS2, MSH6); 2) if uninterpretable IHC or negative staining ≥1 protein, PCR-based (BAT-25, BAT-26); 3) if ≥1 instable, extended PCR (BAT-40, D2S123, D5S346, D17S250).	Sequential: 1 <sup>st</sup> line capecitabine (2 <sup>nd</sup> line irinotecan, 3 <sup>rd</sup> line CAPOX) Combination: 1 <sup>st</sup> line CAPRI (2 <sup>nd</sup> line CAPOX)
Nopel <sup>28</sup> 2014	Phase III RCT (AIO <sup>37</sup> )	229	14	190	PCR-based (1 <sup>st</sup> BAT-26, then if unstable full 5 Bethesda panel); IHC (MLH1, MSH2, MSH6)	1 <sup>st</sup> line FOLFOX or CAPOX

Author	Study Design	mCRC (n)	dMMR (n)	pMMR (n)	Method MMR/MSI	Intervention
Smith <sup>29</sup> 2013	Phase III RCT (COIN <sup>38</sup> )	2445	66	1499	PCR-based (BAT-25, BAT-26)	<b>1<sup>st</sup> line</b> CAPOX or FOLFOX A) continuous; B) continuous + cetuximab; C) intermittent (if no progression at 12-wk scan, then chemotherapy-free interval until progression, then re-start) <b>1<sup>st</sup> line</b> Type treatment: see original studies for CAIRO, COIN and FOCUS. CAIRO2: <b>1<sup>st</sup> line</b> CAPOX-B versus <b>1<sup>st</sup> line</b> CAPOX-B with cetuximab
Venderbosch <sup>1</sup> 2014	Pooled analysis of 4 Phase III trials (CAIRO <sup>36</sup> , CAIRO2 <sup>39</sup> , COIN <sup>38</sup> , FOCUS <sup>33</sup> )	3063	153	2910	See CAIRO <sup>27</sup> , CAIRO2 (conform CAIRO), COIN <sup>29</sup> , FOCUS <sup>23</sup> .	
Omrčen <sup>30</sup> 2019	Phase II trial <sup>40</sup>	40	2	38	IHC (MLH1, MSH2, MSH6)	<b>1<sup>st</sup> line</b> Capecitabine + bevacizumab (elderly patients, age ≥ 70 years)
Morano <sup>31</sup> 2019	Phase II, randomized, open-label trial (Valentino <sup>41</sup> )	199	5	194	PCR-based (BAT-25, BAT-26, NR-21, NR-24, MONO-27)	<b>1<sup>st</sup> line</b> Induction FOLFOX + panitumumab with maintenance panitumumab (Arm A) or 5-FU + panitumumab (Arm B)

*Abbreviations:* 5-FU (5-fluorouracil), CAP-B (capecitabine + bevacizumab), CAPIRI (capecitabine + irinotecan), CAPOX-B (capecitabine + oxaliplatin + bevacizumab), dMMR (deficient mismatch repair), FOLFIRI (5-fluorouracil + irinotecan), FOLFIRI-B (5-fluorouracil + irinotecan + bevacizumab), FOLFOX (5-fluorouracil + oxaliplatin), FOLFOX-B (5-fluorouracil + oxaliplatin + bevacizumab), FOLFOXIRI-B (5-fluorouracil + oxaliplatin + irinotecan + bevacizumab), IHC (immunohistochemistry), mCRC (metastatic colorectal cancer), MSI (microsatellite instability), MMR (mismatch repair), PCR (polymerase chain reaction), pMMR (proficient mismatch repair). 1st or 2nd refer to first or second line systemic therapy. Trial and protein abbreviations are not elaborated.

**Table 2.** Results with treatment efficacy for PFS, OS and response rate, for different systemic therapy agents, per study

Article	Treatment	Overall survival (OS)	Progression-free survival (PFS)	Response rate		Hazard Ratio (HR)	
		dMMR	pMMR	dMMR	pMMR	dMMR	pMMR
Braun <sup>23</sup> 2008	Arm A: 1 <sup>st</sup> line 5-FU, 2 <sup>nd</sup> line irinotecan;						
	Arm B: 1 <sup>st</sup> line 5-FU, 2 <sup>nd</sup> line FOLFIRI or FOLFOX;						
	Arm C: 1 <sup>st</sup> line FOLFIRI or FOLFOX.						
				PFS for 1 <sup>st</sup> line (n=58):	PFS for 1 <sup>st</sup> line (n=1252):		
				5-FU HR PFS 1.00	5-FU HR PFS 1.00		
				FOLFIRI HR PFS 0.93 (0.45-1.91)	FOLFIRI HR PFS 0.76 (0.65-0.89)		
				FOLFOX HR PFS 0.68 (0.35-1.31)	FOLFOX HR PFS 0.72 (0.62-0.84)		
				OS (n=58):	OS (n=1262):		
				Sequential HR OS 1.00	Sequential HR OS 1.00		
				Combination HR OS 0.66 (0.37-1.17)	Combination HR OS 0.94 (0.83-1.00)		
Koopman <sup>27</sup> 2009	Sequential: 1 <sup>st</sup> line capecitabine (2 <sup>nd</sup> line irinotecan, 3 <sup>rd</sup> line CAPOX)	12.7 months (7.4-22.2, p=0.47, n=7)	17.2 months (14.7-18.8, n=230)	4.2 months (2.2-10.6, p=0.72, n=7)	5.8 months (4.9-6.3, n=247)	ORR 25% (2/8) (3-65, p=0.57)	ORR 17% (41/239) (13-23)
						DCR 50% (4/8) (16-84, p=0.09)	DCR 76% (182/239) (70-81)
	Combination: 1 <sup>st</sup> line CAPIRI (2 <sup>nd</sup> line CAPOX)	6.2 months (3.6-31.3, n=8, p=0.58)	18.3 months (16.2-20.6, n=210)	4.0 months (2.3-6.5, p=0.02, n=10)	8.3 months (7.6-8.7, n=243)	ORR 25% (2/8) (3-65%, p=0.30)	DCR 63% (5/8) (25-92, p=0.048).

Article	Treatment	Overall survival (OS)	Progression-free survival (PFS)	Response rate	Hazard Ratio (HR)
		dMMR	pMMR	dMMR	pMMR
Cremolini <sup>32</sup> 2020	<i>Doublet:</i> <b>1<sup>st</sup> line</b> FOLFOX-B + 5-FU-B maintenance,				<b>PFS (doublet versus triplet) (n=26):</b> FOLFOXIRI-B HR PFS1 0.97 (0.39-2.38, $p_{interaction}=0.410$ ); HR PFS2 1.11 (0.44-2.81, $p_{interaction}=0.319$ )
	<b>2<sup>nd</sup> line</b> FOLFIRI-B				<b>PFS (doublet versus triplet) (n=528):</b> FOLFOXIRI-B HR PFS1 0.66 (0.55-0.79); HR PFS2 0.68 (0.57-0.83)
	<i>Triplet: 1<sup>st</sup> line</i> FOLFOXIRI-B + 5-FU-B maintenance,				<b>OS doublet versus triplet:</b> FOLFOXIRI-B HR 0.73 (0.60-0.90)
	<b>2<sup>nd</sup> line</b> FOLFOXIRI-B				<b>OS doublet versus triplet:</b> FOLFOXIRI-B HR 1.28 (0.45-3.63, $p_{interaction}=0.330$ )
Chua <sup>24</sup> 2009	<b>1<sup>st</sup> line</b> FOLFOX		<b>ORR</b> 50% (1/2), OR 0.79 (0.05- 12.97, n=2)	<b>ORR</b> OR 1.00 ( $p=0.09$ , n=114)	(n=114) Reference
Goey <sup>25</sup> 2017	<b>1<sup>st</sup> line</b> CAPOX-B (CAIRO3)	13.6 months (9.5-17.7, n=4).	PFS1 2.1 months (0.0- 11.6, n=4). PFS2 8.0 months (0.0- 17.4, n=4).	HR OS 0.40 (0.06-2.92, $p=0.4$ ). HR PFS 0.36 (0.05-2.62, $p=0.3$ )	
		21.4 months (19.1-23.7, $p=0.040$ , n=275).	PFS1 5.7 months (4.9- 6.6, $p=0.748$ , n=275). PFS2 10.7 months (9.4- 12.0, $p=0.268$ , n=275).		

Article	Treatment	Overall survival (OS)	Progression-free survival (PFS)	Response rate	Hazard Ratio (HR)
		dMMR	pMMR	dMMR	pMMR
Nopel <sup>28</sup> 2015	<b>1<sup>st</sup> line</b> CAPOX/ FOLFOX	24 months (95% C.I. 19.1-29.5, n=14).	17.5 months (95% C.I. 15.5-19.7, n=190). p=0.228.	4.5 months (95% C.I. 7.2, n=14).	7.5 months (95% C.I. 6.7-8.4, p=0.431, n=190).
				<b>ORR</b> 50% (7/14) <b>DCR</b> 64.3% (9/14)	<b>ORR</b> 56.3% (107/190), <b>DCR</b> 85.3% (162/190), p=0.782.
Smith <sup>29</sup> 2013	<b>1<sup>st</sup> line</b> COIN CAPOX/ FOLFOX +/- Cetux	For any type chemotherapy: HR 1.27 (95% C.I. 0.65-2.45; p=0.49, n=45).	Cetuximab, any type chemotherapy: HR 1.26 (95% C.I. 0.68-2.34; p=0.47, n=45); <b>BRAF-wt</b> HR 0.98 (0.49-1.95; p=0.95, n=36).	Cetuximab, any type chemotherapy: HR 0.98 (0.86-1.12; p=0.80, n=977).	Adjusted HR (for somatic mutation status, treatment arm, chemotherapy regimen) PFS HR 1.66 (95% C.I. 1.21-2.27; p=1.6 x10-3, n=66). <b>BRAF-wt subgroup</b> HR PFS 1.85 (95% C.I. 1.31-2.61; p=0.0005, n=36), <b>OS</b> HR 1.60 (95% C.I. 1.14-2.24; p=0.0066, n=66); <b>BRAF-wt subgroup</b> OS HR 1.89 (95% C.I. 1.30-2.76; p=0.00085, n=36).

Article	Treatment	Overall survival (OS)	Progression-free survival (PFS)	Response rate				Hazard Ratio (HR)	
		dMMR	pMMR	dMMR	pMMR	dMMR	pMMR	dMMR	pMMR
Innocenti <sup>26</sup> 2019	<b>1* line</b>	<u>Bevacizumab</u>	<u>Bevacizumab</u>	<u>Bevacizumab</u>	<u>Bevacizumab</u>			Adjusted HR (see article).	Adjusted HR (see article).
	CALGB	30.0 months	30.3 months	( <i>n</i> =21) 9.3 months (5.4-29.0);	( <i>n</i> =285) 11.2 months (10.3-12.5)			<u>Cetuximab</u>	<u>Cetuximab</u>
	Bevacizumab	(23.6-NE);	(27.3-34.3)					( <i>n</i> =16) ref.	( <i>n</i> =301) ref.
	Beva+Cetux	<u>Beva + Cetux</u>	<u>Beva+ Cetux</u>	<u>Beva + Cetux</u>	<u>Beva+ Cetux</u>			<u>Bevacizumab</u>	<u>Bevacizumab</u>
	Cetuximab	21.5 months	26.2 months	( <i>n</i> =15) 7.7 months (6.6-17.6);	( <i>n</i> =189) 10.9 months (9.8-12.8)			( <i>n</i> =21) PFS HR 0.16 (0.07-0.37), <i>p</i> <0.001;	( <i>n</i> =285) PFS HR 0.93 (0.77-1.12, <i>p</i> =0.439)
		(16.4-41.1);	(22.6-29.7)					<u>Beva + Cetux</u>	<u>Beva + Cetux</u>
		<u>Cetuximab</u> OS	<u>Cetuximab</u>	<u>Cetuximab</u>	<u>Cetuximab</u>			( <i>n</i> =15) PFS HR 0.44 (0.20-0.98) (0.80-1.21, <i>p</i> =0.881)	( <i>n</i> =189) PFS HR 0.98 (0.80-1.21, <i>p</i> =0.881)
		11.9 months	30.7 months	( <i>n</i> =16) 5.4 months (4.1-8.6).	( <i>n</i> =301) 10.9 months (9.5-12.8)			<u>Bevacizumab</u>	<u>Bevacizumab</u> OS
		(10.3-24.6); <i>p</i> =0.0014.	(27.6-35.0); <i>p</i> =0.2182					OS HR 0.13 (0.06-0.30), <i>p</i> <0.001;	OS HR 1.06 (0.87-1.29, <i>p</i> =0.539)
								<u>Beva + Cetux</u>	<u>Beva + Cetux</u> OS
									HR 1.18 (0.95-1.47, <i>p</i> =0.134), <i>p</i> =0.018.
									<b>MSI-H vs MSS</b>
									PFS HR 1.02 (0.71-1.47, <i>p</i> =0.912).
									OS HR 0.87 (0.60-1.28, <i>p</i> =0.491).

Article	Treatment	Overall survival (OS)	Progression-free survival (PFS)	Response rate	Hazard Ratio (HR)
		dMMR	pMMR	dMMR	pMMR
Venderbosch <sup>1</sup> 2014	<b>1<sup>st</sup> line</b> CAIRO2 (CAPOX-B + - Cetuximab)	15.6 months (12.9-22.3, <i>n</i> =29).	22.0 (20.3- 24.1, <i>n</i> =487).	7.5 months (6.4-10.5, <i>n</i> =29).	PFS HR 1.66 (1.13-2.45); OS HR 1.60 (1.07-2.40).
Omrčen <sup>30</sup> 2019	<b>1<sup>st</sup> line</b> Capecitabine + Bevacizumab	<b>BRAF-m</b> ( <i>n</i> =2) 6.9 months & 13 months (each patient)	<b>Whole cohort</b> 20.5 months (range 3.5- 134, <i>n</i> =40)	<b>Whole cohort</b> Patient 1: 2.3 months & 13 months (each patient)	
Morano <sup>31</sup> [21] 2019	<b>1<sup>st</sup> line</b> FOLFOX + panitumumab (pan) with maintenance pan (Arm B) or 5-FU + pan (Arm B)	<b>BRAF-wt 2</b> year OS-rate 60.0% (29.3- 100, <i>n</i> =5)	<b>BRAF-wt 2</b> year OS-rate 62.9% (56.0- 70.6, <i>n</i> =194)	<b>BRAF-wt 4.1</b> months (95% C.I. 4.0-NA, <i>n</i> =5)	HR PFS 3.03 (95% C.I. 1.24- 7.42, <i>p</i> =0.015) HR 2-year OS rate 1.23 (95% C.I. 0.38-3.92, <i>p</i> =0.732) Adjusted HR for PFS (see article); HR 1.28 (0.47-3.47, <i>p</i> =0.626)

**Abbreviations:** 5-FU (5-fluorouracil), beva (bevacizumab), BRAF-wt (BRAF-wildtype), BRAF-mt (BRAF-mutant), BRAF-wt (BRAF-wildtype), CAP-B (capecitabine + bevacizumab), CAPIRI (capecitabine + irinotecan), CAPOX-B (capecitabine + oxaliplatin + bevacizumab), cetux (cetuximab), DCR (disease control rate), dMMR (deficient mismatch repair), FOLFIRI (5-fluorouracil + irinotecan), FOLFIRI-B (5-fluorouracil + irinotecan + bevacizumab), FOLFOX (5-fluorouracil + oxaliplatin), FOLFOX-B (5-fluorouracil + oxaliplatin + bevacizumab), FOLFOXIRI-B (5-fluorouracil + oxaliplatin + irinotecan + bevacizumab), MSI-H (microsatellite instability high), MSS (microsatellite stable), ORR (objective response rate), PFS1 (first progression-free survival), PFS2 (second progression-free survival), pMMR (proficient mismatch repair). 1<sup>st</sup> refers to first line systemic therapy. Trial abbreviations are not elaborated.

## Studies across multiple treatment lines

Three studies reported outcomes of multiple treatment lines<sup>23,27,32,33,36</sup>. We first assess the results of OS for the different treatment regimens (sequential versus combination regimens in the CAIRO and FOCUS studies and doublet versus triplet in the TRIBE2 study), while later examining the PFS of different treatment types in the corresponding sections for each treatment.

## Combination versus sequential treatment

The CAIRO and FOCUS studies examined sequential versus combination treatment<sup>23,27,36</sup>. The CAIRO trial randomized patients between first line capecitabine, second line irinotecan and third line capecitabine + oxaliplatin (CAPOX; sequential treatment arm) treatment versus first line capecitabine + irinotecan (CAPIRI), and second line CAPOX (combination treatment arm)<sup>36</sup>. The FOCUS trial randomized patients for 3 treatment arms: first line 5-FU/capecitabine monotherapy and second line irinotecan (arm A); first line 5-FU/capecitabine monotherapy and second line combination chemotherapy consisting of 5-FU with irinotecan or oxaliplatin (FOLFIRI or FOLFOX; arm B), or first- and second line combination chemotherapy consisting of FOLFIRI or FOLFOX (arm C)<sup>33</sup>.

In the CAIRO trial, OS was lower in dMMR mCRC patients treated with combination treatment compared to sequential treatment, however with a low number of patients per group and large survival confidence intervals (6.2 months, 3.6-31.3,  $n=8$  versus 12.7 months, 7.4-22.2,  $n=7$ )<sup>27,36</sup>. However, in the FOCUS trial, there was a non-significant trend to longer OS in dMMR mCRC patients receiving combination treatment versus sequential treatment (HR 0.66, 95% C.I. 0.37-1.17,  $n=58$ )<sup>23,33</sup>. In pMMR mCRC patients, OS was similar with combination and sequential treatment (HR 0.94, 95% C.I. 0.83-1.00,  $n=1252$ )<sup>23,33</sup>. With contrasting non-significant results and a low number of dMMR patients per study, no firm conclusions can be drawn regarding the benefit of combination versus sequential treatment. Upfront combination treatment regimens may be beneficial rather than sequential treatments, since the response rates of dMMR mCRC patients to traditional chemotherapy and targeted treatment are low and thus patients may not be able to receive later line treatments.

## Doublet versus triplet treatment

In the TRIBE2 trial, patients were randomized for first line doublet versus triplet treatment with first line FOLFOX-bevacizumab and 5-FU-bevacizumab (5-FU-B) maintenance followed by second line FOLFIRI-bevacizumab in the doublet arm versus first line 5-FU, oxaliplatin, irinotecan + bevacizumab (FOLFOXIRI-B) and 5-FU-B maintenance followed by second line reintroduction FOLFOXIRI-B in the triplet arm<sup>32</sup>. This study showed no benefit of triplet compared to doublet therapy in dMMR mCRC patients in terms of PFS1 (HR 0.97, 95% C.I. 0.39-2.38,  $n=26$ ), PFS2 (HR 1.11, 95% C.I. 0.44-2.81) or OS (HR 1.28, 95% C.I. 0.45-3.63)<sup>32</sup>. Whereas pMMR mCRC patients showed a clear benefit of triplet compared to doublet therapy in PFS1 (HR 0.66, 95% C.I. 0.55-0.79), PFS2 (HR 0.68, 95% C.I. 0.57-0.83), and OS (HR 0.73, 95% C.I. 0.60-0.90,  $n=528$ )<sup>32</sup>. Although no clear benefit for triplet versus doublet treatment in first line setting can be seen for dMMR mCRC patients, the results reflect subgroup analysis with a low number of patients and thus may not have enough power to detect such a difference.

## Irinotecan-containing systemic therapy versus 5-FU monotherapy

The CAIRO and FOCUS randomized studies also report PFS for 5-FU monotherapy versus irinotecan-containing systemic therapy in first line setting in mCRC patients<sup>23,27,33,36</sup>. In the CAIRO trial, dMMR patients receiving capecitabine and irinotecan doublet therapy (CAPIRI) had a similar progression-free survival (PFS) and ORR as dMMR patients receiving mono capecitabine; PFS 4.0 months (95% C.I. 2.3-6.5,  $n=10$ ) versus 4.2 months (95% C.I. 2.2-10.6,  $n=7$ ), respectively<sup>27</sup>. As a reference, the PFS and ORR were higher in pMMR patients receiving CAPIRI compared to mono capecitabine; PFS 8.3 months (95% C.I. 7.6-8.7,  $n=243$ ) versus 5.8 months (95% C.I. 4.9-6.3,  $n=247$ ), respectively<sup>27</sup>. The overall response rate (ORR) in dMMR mCRC patients receiving CAPIRI compared to mono capecitabine were also similar (25% in both treatment lines)<sup>27</sup>. Likewise, dMMR patients in the FOCUS trial receiving irinotecan-containing therapy, versus 5-FU/capecitabine monotherapy, did not have a significantly different PFS with a hazard ratio (HR) of 0.93 (95% C.I. 0.45-1.91)<sup>23</sup>. In the FOCUS trial, the HR for PFS was non-significantly lower for oxaliplatin-containing therapy (HR 0.68, 95% C.I. 0.35-1.31) compared to 5-FU monotherapy in dMMR patients<sup>23</sup>. There was a trend to a better PFS in pMMR patients receiving irinotecan-containing therapy versus 5-FU/capecitabine; HR 0.76 (95% C.I. 0.65-0.89)<sup>23</sup>.

Thus, the FOCUS and CAIRO trials report a similar PFS in dMMR patients receiving 5-FU monotherapy compared to irinotecan doublet therapy, while pMMR patients have a trend to increased PFS upon irinotecan doublet therapy compared to 5-FU monotherapy. However this should be interpreted with caution due to small numbers and results being subgroup analysis. These results may suggest that 5-FU monotherapy may be better suited by offering similar survival benefit but less toxicity compared to irinotecan doublet therapy, however this should be validated prior to implementation, given the weak evidence. This is in contrast to the preclinical study results, summarized later in this review, which showed that dMMR CRC tumours are sensitive to irinotecan therapy.

## **Oxaliplatin-containing systemic therapy**

Seven studies examined the efficacy of first line oxaliplatin-containing systemic treatment (5-FU and oxaliplatin [FOLFOX] or capecitabine and oxaliplatin [CAPOX]) in dMMR mCRC patients (ancillary studies: <sup>1,23-25,28,29</sup> and original studies: <sup>32-34,37-39,41-44</sup>). The FOCUS trial was the only trial to examine treatment arms with and without oxaliplatin<sup>23,33</sup>, while the CAIRO3 trial examined maintenance treatment<sup>25,34</sup>.

In the FOCUS trial, dMMR mCRC patients receiving FOLFOX had a trend to better PFS (HR 0.68, 95% C.I. 0.35-1.31) compared to patients receiving 5-FU or 5-FU and irinotecan doublet therapy (FOLFIRI) (HR 0.93, 95% C.I. 0.45-1.91)<sup>23</sup>. In comparison, FOCUS pMMR patients had similar PFS with FOLFOX (HR 0.72, 95% C.I. 0.62-0.84) and FOLFIRI (HR 0.76, 95% C.I. 0.65-0.89) compared to 5-FU<sup>23,33</sup>. The remaining studies reported the efficacy of first line oxaliplatin-containing systemic therapy in dMMR mCRC patients, without a comparison arm<sup>1,24,25,28</sup>. The AIO trial reported a comparable ORR in 14 dMMR patients compared to 190 pMMR patients (ORR 50% versus 56.3%) receiving oxaliplatin doublet therapy<sup>28</sup>. Three studies<sup>1,25,28</sup> did not find a statistical difference in PFS for dMMR and pMMR mCRC patients receiving oxaliplatin doublet therapy, although the studies may not have been powered to detect a survival difference, considering the small group sizes. Of note, the PFS was shorter in dMMR mCRC versus pMMR mCRC patients across all three studies. In the AIO trial, the median PFS was 4.5 months (95% C.I. 1.8-7.2,  $n=14$ ) versus 7.5 months (95% C.I. 6.7-8.4,  $n=190$ ,  $p=0.431$ ) respectively<sup>28</sup>. In the CAIRO2 trial, dMMR mCRC patients had a higher, but non-significant, risk for shorter PFS (HR 1.66 (95% C.I. 1.13-2.45), with an observed median PFS of 7.5 months (95% C.I. 6.4-10.5,  $n=29$ ) in dMMR mCRC patients compared to 10.5 months (95% C.I. 9.6-11.4,  $n=407$ ) in pMMR mCRC

patients<sup>1</sup>. Lastly, the PFS1 in CAIRO3 patients was similar regardless of MMR status, with an observed PFS1 of 2.1 months (95% C.I. 0.0-11.6,  $n=4$ ) in dMMR mCRC patients versus PFS1 of 5.7 months (95% C.I. 4.9-6.6,  $n=275$ ,  $p=0.748$ ) in pMMR mCRC patients<sup>25</sup>. The median OS ranged from 13.6 to 24.0 months (dMMR) versus 17.5 to 22.0 months (pMMR), which was significantly different between dMMR and pMMR in the CAIRO3 trial, but not in the AIO trial<sup>1,25,28</sup>. Of note is the low percentage of dMMR mCRC patients in the CAIRO3 trial (1%), lower than the expected 5% prevalence, potentially indicating a patient selection due to poor treatment response as all CAIRO3 patients were required to have at least stable disease after 6 cycles of induction CAPOX-B prior to inclusion<sup>25</sup>.

## Targeted therapy (anti-EGFR and anti-VEGF)

In mCRC patients, primary tumour location 'sidedness' and *RAS*-mutational status are predictive biomarkers for treatment response to anti-EGFR treatment, with a worse outcome for patients with a right-sided and/or *RAS*-mutant tumour<sup>35,45,46</sup>. The predictive value of *BRAF*-mutational status in mCRC patients receiving anti-EGFR is inconclusive. The BEACON trial indicates that combination treatment with anti-EGFR, MEK and *BRAF* inhibitors in later line setting provides survival benefit, while other studies indicate that monotherapy anti-EGFR treatment in pre-treated patients should be avoided<sup>47,48</sup>. In analyzing the trial results for dMMR mCRC patients receiving anti-EGFR treatment, we also analyze available results for relevant prognostic subgroups.

In the COIN, Valentino, CALGB, CAIRO3 trials and Omrčen *et al.* study, mCRC patients received first line treatment containing chemotherapy with cetuximab or bevacizumab or combined cetuximab/bevacizumab treatment (ancillary studies: <sup>25,26,29-31</sup> and original trials: <sup>34,35,38,40,41</sup>). The CAIRO3 trial results (where all patients received oxaliplatin doublet and bevacizumab targeted treatment) is discussed in detail above. Given the low number of patients, we shall not discuss the Omrčen *et al.* study in detail<sup>30,40</sup>.

COIN study patients received CAPOX or FOLFOX treatment and were randomized to receive cetuximab<sup>29,38</sup>. In 45 dMMR mCRC patients receiving CAPOX/FOLFOX with cetuximab, there was a trend towards worse PFS (HR 1.26, 95% C.I. 0.68-2.34,  $p=0.47$ ) compared to dMMR mCRC patients without cetuximab<sup>29</sup>. The trend was not present when analyzing the dMMR mCRC *BRAF*-wildtype subgroup (HR 0.98, 95% C.I. 0.49-1.95,  $p=0.95$ ,  $n=36$ )<sup>29</sup>. No subgroup data results are available regarding the effect of primary tumour sidedness or *RAS*-mutational status in dMMR mCRC patients

receiving cetuximab. In comparison, pMMR mCRC patients had similar PFS with or without cetuximab, including the pMMR / *BRAF*-wildtype subgroup<sup>29</sup>. The adjusted HR for PFS and OS, for somatic mutation status, treatment arm and chemotherapy regimen, were significantly worse in dMMR patients compared to pMMR patients; PFS HR 1.66 (95% C.I. 1.21-2.27,  $p=0.0016$ ) and OS HR 1.60 (95% C.I. 1.14-2.24,  $p=0.0066$ )<sup>29</sup>.

In the Valentino trial, *BRAF*/*RAS*-wildtype mCRC patients were randomized between 5-FU and panitumumab maintenance treatment (arm A) and panitumumab (arm B), after induction first line treatment consisting of FOLFOX plus panitumumab<sup>31,41</sup>. PFS was significantly lower (unadjusted HR 3.03, 95% C.I. 1.24-7.42,  $p=0.015$ ) in dMMR versus pMMR mCRC, with a median PFS of 4.1 months (95% C.I. 4.0-not reached,  $n=5$ ) versus 11.1 months (95% C.I. 10.4-13.2,  $n=194$ ) respectively<sup>31</sup>. The 2-year OS rate is lower in dMMR mCRC patients (60.0%, 95% C.I. 9.3-100) versus pMMR mCRC patients (62.9%, 95% C.I. 56.0-70.6) with an unadjusted HR of 1.23 (95% C.I. 0.38-3.92,  $p=0.732$ )<sup>31</sup>. Of note, 40% (2/5) of the dMMR mCRC patients had molecular alterations associated with anti-EGFR primary resistance (*PI3KCA* exon 20 mutation and *PTEN* inactivating mutation)<sup>31</sup>. In multivariable analysis, dMMR status was not associated with PFS when adjusted for other clinical factors (HR 1.28, 95% C.I. 0.47-3.47,  $p=0.626$ ), however this might be limited by the low number (5) of dMMR patients in the trial<sup>31</sup>. Thus, the survival of dMMR mCRC patients receiving first line panitumumab (with FOLFOX and maintenance treatment) is lower than in pMMR mCRC patients, but MMR status was not independently prognostic for survival during anti-EGFR treatment when adjusted for other clinical variables in the Valentino trial.

Patients in the CALGB trial were randomized to receive bevacizumab, cetuximab or both in combination with first line chemotherapy (FOLFOX or FOLFIRI)<sup>26,35</sup>. An adjusted HR for PFS and OS was calculated, adjusting for several factors, including primary tumour location (right versus transverse vs left-sidedness), *BRAF* and *RAS* mutational status<sup>26</sup>. CALGB patients with dMMR tumours receiving cetuximab versus bevacizumab had a significantly worse adjusted HR for PFS and OS, while pMMR patients showed comparable adjusted HR for PFS and OS in all treatment arms<sup>26</sup>. The adjusted HR for PFS in dMMR mCRC patients receiving bevacizumab, combined bevacizumab/cetuximab and cetuximab was HR 0.16 (95% C.I. 0.07-0.37,  $n=21$ ,  $p<0.001$ ) versus HR 0.44 (95% C.I. 0.20-0.99,  $n=15$ ,  $p=0.046$ ) versus HR 1.0 ( $n=16$ ), respectively<sup>26</sup>. The reported PFS in dMMR patients receiving chemotherapy + cetuximab is 5.4 months (4.1-8.6,  $n=16$ ) versus 9.3 months (5.4-29.0,  $n=21$ ) for chemotherapy + bevacizumab<sup>26</sup>.

## BIOLOGICAL RATIONALE FOR DIFFERENTIAL SENSITIVITY OF dMMR AND pMMR CRC TUMOURS TO SYSTEMIC THERAPY

### Assessment of available models used in pre-clinical studies

Preclinical studies have extensively examined the sensitivity, and mechanisms involved, of dMMR CRC to chemotherapy with or without monoclonal antibodies, as outlined in recent reviews<sup>49–52</sup>. An extensive literature review was provided in the recent reviews, as such, we briefly touch on the relevant findings below. These preclinical findings have not resulted in a clear benefit in clinical trials. A potential barrier to interpreting the available pre-clinical data is the choice of model systems used, which traditionally have included cell lines with a dMMR system<sup>53–57</sup> and xenograft models<sup>52,54</sup>.

The use of dMMR CRC cell lines in preclinical studies has key limitations. Cell lines are genetically unstable *in vitro* and acquire a more aggressive subclone through repeated passaging while culturing, thus no longer reflecting the heterogeneity of the original tumours<sup>58</sup>. The dMMR preclinical studies often chose cell lines with only one mutational type (e.g. *hMLH1* deficient cell lines), not accurately reflecting the mutational landscape of dMMR CRC patients. Preclinical studies should comprise *BRAF*-wildtype and *BRAF*-mutant models (to reflect Lynch syndrome and sporadic dMMR patients), and where applicable, confirm results in epigenetically silenced *hMLH1* models. Moreover, dMMR CRC cell lines should display genomic instability reflecting the genomic landscape of dMMR CRC tumours. Due to non-functional mismatch repair machinery, dMMR CRC tumours accumulate somatic mutations in specific oncogenes and tumour suppressor genes, including *BRAF*, *MRE11A*, *KRAS*, *TGF $\beta$ rII* and *IGF2R*<sup>59,60</sup>. The choice of cell line can affect preclinical results, as was seen in irinotecan preclinical studies where cell lines without secondary *MRE11A* mutations had a different irinotecan sensitivity compared to *MRE11A*-mutant cell lines<sup>50,57,61,62</sup>.

Recent studies have shown that immunotherapy can effectively produce anti-tumour immune responses in dMMR CRC tumours, highlighting the crucial role of the tumour microenvironment (TME)<sup>17–19</sup>. In addition to immunotherapy, the TME also affects the sensitivity of other treatments<sup>63</sup>. Thus, when studying dMMR mCRC sensitivity to treatment it is important to use models with TME components where possible. Patient-derived xenograft models, which have stromal components, display intra-tumoural heterogeneity and long-term genomic stability. However, they are limited in clinical

applicability since they require substantial patient tumour substrate, animal models and months to establish<sup>64–66</sup>. Humanized mouse models have been used to model PD-1 blockade using xenografts for non-colorectal cancers<sup>67</sup>. The models would lend well to preclinical dMMR CRC treatment studies, since the mice have a full human immune system as well as the engrafted xenograft. Patient-derived organoids (PDOs) may be a more accurate model for dMMR CRC, since organoids maintain intra-tumoural heterogeneity, have long-term phenotypical and genetical stability and can be easily propagated from patient tissues<sup>68,69</sup>. Organoids allow researchers to perform library drug screens on dMMR mCRC PDOs, revealing the sensitivity of patients to different treatments in one experiment<sup>70</sup>. Co-cultures with TME components can be established using PDOs; moreover, these co-culture models have altered treatment sensitivity compared to PDOs alone<sup>71,72</sup>. Drug screen data derived from CRC PDOs is highly correlated to patient response in the clinic, showing that organoids may be used to predict treatment response<sup>73–75</sup>.

## **KEY PRECLINICAL FINDINGS CONCERNING ALTERED SENSITIVITY OF dMMR CRC TUMOURS TO CHEMOTHERAPY WITH OR WITHOUT MONOCLONAL ANTIBODIES**

To clarify the clinical findings, we summarize the biological rationale for a differential benefit of chemotherapy with or without monoclonal antibodies in dMMR mCRC patients. Preclinical studies have examined the response to the main backbones of chemotherapy (5-FU, irinotecan and oxaliplatin) in dMMR CRC. Chemotoxicity induced by 5-FU treatment is thought to be partially mediated by the MMR pathway<sup>76</sup>. The hMutSQ heterodimer, consisting of the MMR proteins hMSH6 and hMSH2, binds and recognizes 5-FU-modified DNA<sup>77</sup>. Although not all studies agree<sup>78</sup>, preclinical studies have demonstrated that dMMR CRC cell lines are at least 18-fold more resistant to 5-FU than pMMR cell lines, and that the resistance is reversed when the MMR deficiency is corrected<sup>55,56,79</sup>. Regarding topoisomerase inhibitor sensitivity, dMMR CRC tumours are thought to be more sensitive to irinotecan since these tumours often have a secondary frameshift mutation in *MRE11A*, part of the double strand break repair (dsBR) complex<sup>57,80,81</sup>. CRC cell lines which were hMLH1-deficient, especially cell lines with *MRE11A* mutations, were more sensitive to irinotecan than pMMR cell lines, and accordingly, restoring hMLH1 induced resistance to irinotecan<sup>57,61,62</sup>. Other

studies contradict the above results, possibly due to using dMMR cell lines without secondary mutations<sup>50,53</sup>. Thirdly, dMMR CRC cell lines are resistant to cisplatin and carboplatin, but not to oxaliplatin, when compared to pMMR CRC cell lines<sup>54,78,82</sup>. Oxaliplatin forms a 1,2-diaminocyclohexane ligand, which is not recognized by the MMR pathway when incorporated into DNA<sup>83</sup>, thus dMMR CRC tumours do not have a different sensitivity to oxaliplatin compared to pMMR tumours<sup>50</sup>. Altogether, preclinical studies suggest that dMMR mCRC, compared to pMMR mCRC, may be more resistant to 5-FU, more sensitive to irinotecan, but not have an altered sensitivity to oxaliplatin. This in contrast to the trends observed in clinical studies, where 5-FU + irinotecan doublet therapy did not have an increased benefit over 5-FU monotherapy or oxaliplatin doublet treatment.

Regarding the effectivity of targeted treatment, to our knowledge no studies have investigated the efficacy of anti-EGFR or anti-VEGF treatment in preclinical dMMR CRC models. dMMR mCRC patients may be less sensitive to anti-EGFR treatment due to the higher prevalence of *BRAF*-mutations, increased right-sided primary tumour location and decreased EGFR ligand expression<sup>84,85</sup>. For anti-VEGF targeted treatment, contradictory hypotheses exist: mucinous dMMR CRC tumours have decreased *VEGF* expression and decreased microvessel density, which would support lower sensitivity to anti-VEGF targeted treatment<sup>86</sup>, however, the immunostimulatory effect associated with anti-VEGF treatment may potentiate the immunostimulatory environment in dMMR CRC tumours, improving sensitivity<sup>87</sup>. In clinical trials, there is a suggestion of better efficacy of bevacizumab targeted therapy in dMMR patients compared to cetuximab targeted therapy and of reduced benefit to cetuximab targeted therapy in general, which may be explained by the above mechanisms.

## CONCLUSIONS AND FUTURE PERSPECTIVES

This evidence-based review cannot draw a definitive conclusion regarding the relative efficacy of different systemic therapy agents within dMMR mCRC patients, due to the low amount of clinical trials and low number of dMMR patients, despite an increasing number of trials reporting subgroup data for dMMR mCRC patients. We have identified three trends. These trends need to be confirmed by more studies prior to implementation of altered treatment strategies in clinical practice since there are too few trials available which support these results. Firstly, dMMR patients receiving 5-FU + irinotecan doublet therapy did not have an increased benefit over

5-FU monotherapy, suggesting that patients may not benefit from the addition of irinotecan<sup>23,27</sup>. Secondly, one study reported a trend to better efficacy of oxaliplatin-containing doublet treatment versus irinotecan or 5-FU monotherapy in dMMR mCRC patients, suggesting that dMMR mCRC may benefit more from oxaliplatin doublet therapy compared to irinotecan doublet or 5-FU monotherapy<sup>23</sup>. Lastly, there was a suggestion of better efficacy of bevacizumab targeted therapy in dMMR patients compared to cetuximab targeted therapy<sup>26</sup>, and of no significant benefit in PFS of cetuximab (versus no targeted treatment) in *BRAF*-wildtype dMMR mCRC patients<sup>29</sup>. Despite our review, it is difficult to establish the exact predictive value of MMR for each treatment strategy considering the low numbers of patients and heterogeneous studies. Given the sparse clinical evidence supporting each trend, more studies are needed to confirm the findings prior to implementation in clinical practice. Likewise, no conclusion can be drawn about the effect of other relevant factors including sidedness and mutational status.

The strength of this review lies in including all current available evidence, linking the results to preclinical studies and aiming to identify solutions to accelerate research on this topic. Our analysis was restricted by the quality of data available, which is limited by a low number of studies, small number of dMMR mCRC patients per treatment and heterogeneity in study designs. Furthermore, there may be a selection bias in the available evidence since we could only include studies where MMR status was known. As well, a time bias is likely, given that the prognostic value of MMR status in mCRC patients was not known prior to 2009<sup>27</sup>. The majority of included studies were ancillary studies, based on subgroup analysis for which the trials were not powered, which limits the conclusions which can be drawn based on the studies alone. A meta-analysis of the data could provide clarity. However, due to the limited data available, a meta-analysis was impossible and we could not compare treatment results for dMMR mCRC subgroups (e.g. Lynch syndrome versus sporadic dMMR). The limited amount of available data affirms the urgent need for trials examining which systemic treatment is most beneficial in dMMR mCRC patients.

Due to the low prevalence of dMMR in the metastatic CRC population, available trials only identify a small number of dMMR mCRC patients. We advise clinical trial researchers and clinicians to publish molecular analysis results and deposit patient-level data in data repositories. This will enable meta-analyses, hopefully increasing our knowledge of the most effective treatment in dMMR mCRC patients. If relevant

clinical parameters are shared, a comparison of relevant dMMR mCRC patient subgroups (e.g. Lynch syndrome) and adjustment of prognostic clinical parameters is possible. Furthermore, we advise clinicians to test all mCRC patients for MMR status to guide treatment, but also to enable CRC patient registries to register population-based data for research, which is often more reflective of patients in the clinic than trials. Effective preclinical studies can help to confirm differential treatment sensitivity and to elucidate the molecular pathways involved in dMMR CRC. Compared to other preclinical models, patient-derived organoids can accelerate research regarding which systemic therapy is most effective in dMMR mCRC patients, as discussed previously. This technique is especially applicable to low prevalence diseases, such as dMMR mCRC, where patients are under-represented in clinical trials. Cumulatively, these efforts will help future research for dMMR mCRC sensitivity to treatment, as shown in Figure 3.

**Figure 3.** Accelerating research for dMMR mCRC treatment

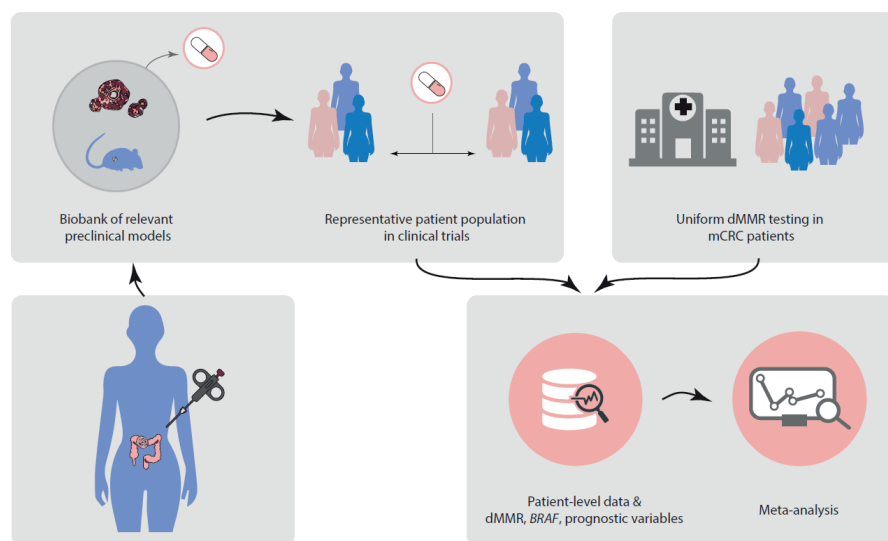


Fig 3. To determine which treatment is most effective in dMMR mCRC patients, representative preclinical models for the dMMR CRC population should be used (e.g. patient-derived organoids). In clinical trials, the patient population should reflect the clinical population (e.g. a mix of germline and sporadic dMMR) and for patient-level data, dMMR status, mutational status and relevant prognostic factor should be made available in depositories so that patient-level meta-analysis is possible.

Despite advances in immune checkpoint inhibitor treatment, it is anticipated that dMMR mCRC patients will continue to receive chemotherapy with or without monoclonal antibodies during their treatment, prior to receiving immunotherapy, upon progression during immunotherapy or in combination with immunotherapy. One strategy to overcome primary resistance to immunotherapy is to combine immune checkpoint inhibitors with chemotherapy and/or targeted treatment<sup>88</sup>. Chemotherapy can increase the efficacy of immunotherapy through promoting tumour antigen release, antigen presentation and stimulating immune effectors, while certain targeted treatments decrease tumour characteristics which have been associated with decreased immunotherapy effectivity<sup>88</sup>. Several phase III trials of combination treatment with immunotherapy in dMMR CRC patients are ongoing, including the COMMIT trial (NCT02997228; first line FOLFOX, bevacizumab with atezolizumab versus atezolizumab single agent treatment for mCRC patients), the A021502 trial (NCT02912559; chemotherapy combined with atezolizumab versus chemotherapy as adjuvant therapy) and the CheckMate-142 trial (NCT02060188; which includes an intervention arm with combination nivolumab + ipilimumab and cobimetinib treatment). These and other ongoing studies will help clarify the role of chemotherapy and targeted treatment in increasing the efficacy of immunotherapy.

The available evidence highlights differences in sensitivity to systemic therapy in patients with dMMR versus pMMR mCRC. However, no definitive conclusion can be drawn regarding which chemotherapy with or without monoclonal antibodies is more effective in dMMR mCRC patients. A better understanding of which chemotherapy with or without monoclonal antibodies is most effective in dMMR mCRC patients is urgently needed. Several strategies can help accelerate research into this field: using preclinical models which more accurately reflect dMMR CRC patients (e.g. PDOs comprising multiple molecular subgroups), publishing patient-level data to allow for meta-analyses and increased testing of MMR status in patients in daily clinical care to increase known dMMR mCRC patients in population registries. With these solutions, clinicians and researchers will hopefully be able to reveal how best to treat dMMR mCRC patients, improving their prognosis.

## REFERENCES

1. Venderbosch S., Nagtegaal I.D., Maughan T.S., Smith C.G., Cheadle J.P., Fisher D., et al. Mismatch repair status and BRAF mutation status in metastatic colorectal cancer patients: A pooled analysis of the CAIRO, CAIRO2, COIN, and FOCUS studies. *Clin Cancer Res.* (2014);**20**(20):5322–30.
2. Gelsomino F., Barbolini M., Spallanzani A., Pugliese G., Cascinu S. The evolving role of microsatellite instability in colorectal cancer: A review. *Cancer Treat Rev.* (2016);**51**:19–26.
3. Weisenberger D.J., Siegmund K.D., Campan M., Young J., Long T.I., Faasse M.A., et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet.* (2006);**38**(7):787–93.
4. Mensenkamp A.R., Vogelaar I.P., Van Zelst-Stams W.A.G., Goossens M., Ouchene H., Hendriks-Cornelissen S.J.B., et al. Somatic mutations in MLH1 and MSH2 are a frequent cause of mismatch-repair deficiency in lynch syndrome-like tumors. *Gastroenterology.* (2014);**146**(3):643–646.e8.
5. Børresen A.-L., Lothe R.A., Meling G.I., Lystad S., Morrison P., Lipford J., et al. Somatic mutations in the hMSH2 gene in microsatellite unstable colorectal carcinomas. *Hum Mol Genet.* (1995);**4**(11):2065–72.
6. Tran B., Kopetz S., Tie J., Gibbs P., Jiang Z. Impact of BRAF Mutation and Microsatellite Instability on the Pattern of Metastatic Spread and Prognosis in Metastatic Colorectal Cancer. *Cancer.* (2011);**117**:4623–32.
7. Aasebø K., Dragomir A., Sundström M., Mezheyski A., Edqvist P.H., Eide G.E., et al. Consequences of a high incidence of microsatellite instability and BRAF-mutated tumors: A population-based cohort of metastatic colorectal cancer patients. *Cancer Med.* (2019);**8**(7):3623–35.
8. Wensink G.E., Elferink M.A.G., May A.M., Mol L., Hamers P.A.H., Bakker S.D., et al. Survival of patients with deficient mismatch repair metastatic colorectal cancer in the pre-immunotherapy era. *Br J Cancer.* (2021);**124**(2):399–406.
9. Kim C.G., Ahn J.B., Jung M., Beom S.H., Kim C., Kim J.H., et al. Effects of microsatellite instability on recurrence patterns and outcomes in colorectal cancers. *Br J Cancer.* (2016);**115**(1):25–33.
10. André T., Shiu K.-K., Kim T.W., Jensen B.V., Jensen L.H., Punt C., et al. Pembrolizumab in Microsatellite-Instability–High Advanced Colorectal Cancer. *NEJM.* (2020);**383**(23):2207–18.
11. Taieb J., Shi Q., Pederson L., Alberts S., Wolmark N., Cutsem E. Van, et al. Prognosis of microsatellite instability and / or mismatch repair deficiency stage III colon cancer patients after disease recurrence following adjuvant treatment: results of an ACCENT pooled analysis of seven studies. *Ann Oncol.* (2019);**30**:1466–71.
12. Chong L.C., Townsend A.R., Young J., Roy A., Piantadosi C., Hardingham J.E., et al. Outcomes for Metastatic Colorectal Cancer Based on Microsatellite Instability: Results from the South Australian Metastatic Colorectal Cancer Registry. *Target Oncol.* (2019);**14**(1):85–91.
13. Goldstein J., Tran B., Ensor J., Gibbs P., Wong H.L., Wong S.F., et al. Multicenter retrospective analysis of metastatic colorectal cancer (CRC) with high-level microsatellite instability (MSI-H). *Ann Oncol.* (2014);**25**:1032–8.

14. Cohen R., Buhard O., Cervera P., Hain E., Dumont S., Andre T. Clinical and molecular characterisation of hereditary and sporadic metastatic colorectal cancers harbouring microsatellite instability/DNA mismatch repair deficiency. *Eur J Cancer*. (2017);**86**:266–74.
15. Boeckx N., Janssens K., Van Camp G., Rasschaert M., Papadimitriou K., Peeters M., et al. The predictive value of primary tumor location in patients with metastatic colorectal cancer: A systematic review. *Crit Rev Oncol Hematol*. (2018);**121**:1–10.
16. Marisa L., Svrcek M., Collura A., Becht E., Cervera P., Wanherdrick K., et al. The Balance Between Cytotoxic T-cell Lymphocytes and Immune Checkpoint Expression in the Prognosis of Colon Tumors. *JNCI*. (2018);**110**(1):68–77.
17. Le D.T., Uram J.N., Wang H., Bartlett B.R., Kemberling H., Eyring A.D., et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *NEJM*. (2015);**372**(26):2509–20.
18. Overman M.J., Mcdermott R., Leach J.L., Lonardi S., Lenz H., Morse M.A., et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *Lancet*. (2017);**18**(9):1182–91.
19. Overman M.J., Lonardi S., Yeung K., Wong M., Lenz H., Gelsomino F., et al. Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair–Deficient/ Microsatellite Instability–High Metastatic Colorectal Cancer. *JCO*. (2018);**36**(8):773–9.
20. Le D.T., Kim T.W., Van Cutsem E., Geva R., Jäger D., Hara H., et al. Phase II Open-Label Study of Pembrolizumab in Treatment-Refractory, Microsatellite Instability–High/ Mismatch Repair–Deficient Metastatic Colorectal Cancer: KEYNOTE-164. *J Clin Oncol*. (2020);**38**(1):11–9.
21. Guetz G.D.E.S., Uzzan B., Nicolas P., Schischmanoff O., Perret G., Morere J. Microsatellite Instability does not Predict the Efficacy of Chemotherapy in Metastatic Colorectal Cancer. A Systematic Review and Meta-analysis. *Anticancer Res*. (2009);**29**(5):1615–20.
22. Van Cutsem E., Cervantes A., Adam R., Sobrero A., Van Krieken J.H., Aderka D., et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol*. (2016);**27**(8):1386–422.
23. Braun M.S., Richman S.D., Quirke P., Daly C., Adlard J.W., Elliott F., et al. Predictive Biomarkers of Chemotherapy Efficacy in Colorectal Cancer: Results From the UK MRC FOCUS Trial. *J Clin Oncol*. (2008);**26**(16):2690–8.
24. Chua W., Goldstein D., Lee C.K., Dhillon H., Michael M., Mitchell P., et al. Molecular markers of response and toxicity to FOLFOX chemotherapy in metastatic colorectal cancer. *Br J Cancer*. (2009);**101**(6):998–1004.
25. Goey K.K.H., Elias S.G., van Tinteren H., Laclé M.M., Willems S.M., Offerhaus G.J.A., et al. Maintenance treatment with capecitabine and bevacizumab versus observation in metastatic colorectal cancer: Updated results and molecular subgroup analyses of the phase 3 CAIRO3 study. *Ann Oncol*. (2017);**28**(9):2128–34.
26. Innocenti F., Ou F.-S., Qu X., Zemla T.J., Niedzwiecki D., Tam R., et al. Mutational Analysis of Patients With Colorectal Cancer in CALGB/SWOG 80405 Identifies New Roles of Microsatellite Instability and Tumor Mutational Burden for Patient Outcome. *J Clin Oncol*. (2019);**37**(14):1217–27.
27. Koopman M., Kortman G.A.M., Mekenkamp L., Ligtenberg M.J.L., Hoogerbrugge N., Antonini N.F., et al. Deficient mismatch repair system in patients with sporadic advanced colorectal cancer. *Br J Cancer*. (2009);**100**(2):266–73.

28. Nöpel-Dünnebacke S., Schulmann K., Reinacher-Schick A., Porschen R., Schmiegel W., Tannapfel A., et al. Prognostic value of microsatellite instability and p53 expression in metastatic colorectal cancer treated with oxaliplatin and fluoropyrimidine-based chemotherapy. *Z Gastroenterol.* (2014);**52**(12):1394–401.
29. Smith C.G., Fisher D., Claes B., Maughan T.S., Idziaszczyk S., Peuteman G., et al. Somatic profiling of the epidermal growth factor receptor pathway in tumors from patients with advanced colorectal cancer treated with chemotherapy  $\pm$  cetuximab. *Clin Cancer Res.* (2013);**19**(15):4104–13.
30. Omrčen T., Katić A., Tomić S., Eterović D., Vrdoljak E. Predictors of outcome in elderly patients with metastatic colorectal cancer: the final results of a prospective phase II study of bevacizumab in combination with capecitabine as first-line treatment. *Anticancer Drugs.* (2020);**31**(5):518–22.
31. Morano F., Corallo S., Lonardi S., Raimondi A., Cremolini C., Rimassa L., et al. Negative hyperselection of patients with RAS and BRAF wild-type metastatic colorectal cancer who received panitumumab-based maintenance therapy. *J Clin Oncol.* (2019);**37**(33):3099–110.
32. Cremolini C., Antoniotti C., Rossini D., Lonardi S., Loupakis F., Pietrantonio F., et al. Upfront FOLFOXIRI plus bevacizumab and reintroduction after progression versus mFOLFOX6 plus bevacizumab followed by FOLFIRI plus bevacizumab in the treatment of patients with metastatic colorectal cancer (TRIBE2): a multicentre, open-label, phase 3, randomised, controlled trial. *Lancet Oncol.* (2020);**21**(4):497–507.
33. Seymour M.T., Maughan T.S., Ledermann J.A., Topham C., James R., Gwyther S.J., et al. Different strategies of sequential and combination chemotherapy for patients with poor prognosis advanced colorectal cancer (MRC FOCUS ): a randomised controlled trial. *Lancet.* (2007);**370**:143–52.
34. Simkens L.H.J., Van Tinteren H., May A., Ten Tije A.J., Creemers G.J.M., Loosveld O.J.L., et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): A phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet.* (2015);**385**(9980):1843–52.
35. Venook A.P., Niedzwiecki D., Lenz H.J., Innocenti F., Fruth B., Meyerhardt J.A., et al. Effect of first-line chemotherapy combined with cetuximab or bevacizumab on overall survival in patients with KRAS wild-type advanced or metastatic colorectal cancer a randomized clinical trial. *JAMA.* (2017);**317**(23):2392–401.
36. Koopman M., Antonini N.F., Douma J., Wals J., Honkoop A.H., Erdkamp F.L.G., et al. Sequential versus combination chemotherapy with capecitabine, irinotecan, and oxaliplatin in advanced colorectal cancer (CAIRO): a phase III randomised controlled trial. *Lancet.* (2007);**370**:135–42.
37. Porschen R., Arkenau H.T., Kubicka S., Greil R., Seufferlein T., Freier W., et al. Phase III study of capecitabine plus oxaliplatin compared with fluorouracil and leucovorin plus oxaliplatin in metastatic colorectal cancer: A final report of the AIO colorectal study group. *J Clin Oncol.* (2007);**25**(27):4217–23.
38. Maughan T.S., Adams R.A., Smith C.G., Meade A.M., Seymour M.T., Wilson R.H., et al. Addition of cetuximab to oxaliplatin-based first-line combination chemotherapy for treatment of advanced colorectal cancer : results of the randomised phase 3 MRC COIN trial. *Lancet.* (2011);**377**(9783):2103–14.

39. Tol J., Koopman M., Rodenburg C.J., Cats A., Creemers G.J., Schrama J.G., et al. A randomised phase III study on capecitabine, oxaliplatin and bevacizumab with or without cetuximab in first-line advanced colorectal cancer, the CAIRO2 study of the Dutch Colorectal Cancer Group (DCCG). An interim analysis of toxicity. *Ann Oncol.* (2008);**19**(4):734–8.
40. Vrdoljak E., Omrčen T., Boban M., Hrabar A. Phase II study of bevacizumab in combination with capecitabine as first-line treatment in elderly patients with metastatic colorectal cancer. *Anticancer Drugs.* (2011);**22**(2):191–7.
41. Pietrantonio F., Morano F., Corallo S., Miceli R., Lonardi S., Raimondi A., et al. Maintenance Therapy with Panitumumab Alone vs Panitumumab Plus Fluorouracil-Leucovorin in Patients with RAS Wild-Type Metastatic Colorectal Cancer: A Phase 2 Randomized Clinical Trial. *JAMA Oncol.* (2019);**5**(9):1268–75.
42. Goldstein D., Mitchell P., Michael M., Beale P., Friedlander M., Zalberg J., et al. Australian experience of a modified schedule of FOLFOX with high activity and tolerability and improved convenience in untreated metastatic colorectal cancer patients. *Br J Cancer.* (2005);**92**(5):832–7.
43. Mitchell P.L., Goldstein D., Michael M., Beale P., Friedlander M., Zalberg J., et al. Addition of gabapentin to a modified FOLFOX regimen does not reduce oxaliplatin-induced neurotoxicity. *Clin Colorectal Cancer.* (2006);**6**(2):146–51.
44. Clarke S., Goldstein D., Mitchell P., Michael M., Beale P., Friedlander M., et al. Modification of leucovorin dose within a simplified FOLFOX regimen improves tolerability without compromising efficacy. *Clin Colorectal Cancer.* (2007);**6**(8):578–82.
45. Tejpar S., Stintzing S., Ciardiello F., Tabernero J., van Cutsem E., Beier F., et al. Prognostic and Predictive Relevance of Primary Tumor Location in Patients With RAS Wild-Type Metastatic Colorectal Cancer Retrospective Analyses of the CRYSTAL and FIRE-3 Trials. *JAMA Oncol.* (2017);**3**(2):194–201.
46. Van Cutsem E., Köhne C.H., Hitre E., Zaluski J., Chien C.R.C., Makhson A., et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *NEJM.* (2009);**360**(14):1408–17.
47. Taieb J., Lapeyre-Prost A., Laurent Puig P., Zaanan A. Exploring the best treatment options for BRAF-mutant metastatic colon cancer. *Br J Cancer.* (2019);**121**(6):434–42.
48. Kopetz S., Grothey A., Yaeger R., Van Cutsem E., Desai J., Yoshino T., et al. Encorafenib, binimetinib, and cetuximab in BRAF V600E-mutated colorectal cancer. *NEJM.* (2019);**381**(17):1632–43.
49. Zhang C.M., Lv J.F., Gong L., Yu L.Y., Chen X.P., Zhou H.H., et al. Role of deficient mismatch repair in the personalized management of colorectal cancer. *Int J Environ Res Public Health.* (2016);**13**(9):892.
50. Hewish M., Lord C.J., Martin S.A., Cunningham D., Ashworth A. Mismatch repair deficient colorectal cancer in the era of personalized treatment. *Nat Rev Clin Oncol.* (2010);**7**(4):197–208.
51. Valentini A.M., Armentano R., Pirrelli M., Caruso M.L. Chemotherapeutic agents for colorectal cancer with a defective mismatch repair system: The state of the art. *Cancer Treat Rev.* (2006);**32**(8):607–18.
52. Lee K., Tosti E., Edelmann W. Mouse models of DNA mismatch repair in cancer research. *DNA Repair (Amst).* (2016);**38**:140–6.

53. Fedier A., Schwarz V.A., Walt H., Carpini R.D., Haller U., Fink D. Resistance to topoisomerase poisons due to loss of DNA mismatch repair. *Int J Cancer*. (2001);**93**(4):571–6.
54. Fink D., Zheng H., Nebel S., Norris P.S., Aebi S., Lin T.P., et al. In vitro and in vivo resistance to cisplatin in cells that have lost DNA mismatch repair. *Cancer Res*. (1997);**57**(10):1841–5.
55. Meyers M., Wagner M.W., Hwang H.S., Kinsella T.J., Boothman D.A. Role of the hMLH1 DNA mismatch repair protein in fluoropyrimidine-mediated cell death and cell cycle responses. *Cancer Res*. (2001);**61**(13):5193–201.
56. Carethers J.M., Chauhan D.P., Fink D., Nebel S., Bresalier R.S., Howell S.B., et al. Mismatch repair proficiency and in vitro response to 5-fluorouracil. *Gastroenterology*. (1999);**117**:123–31.
57. Vilar E., Scaltriti M., Balmã J., Saura C., Guzman M., Arribas J., et al. Microsatellite instability due to hMLH1 deficiency is associated with increased cytotoxicity to irinotecan in human colorectal cancer cell lines. *Br J Cancer*. (2008);**99**(10):1607–12.
58. Pampaloni F., Reynaud E.G., Stelzer E.H.K. The third dimension bridges the gap between cell culture and live tissue. *Nat Rev Mol Cell Biol*. (2007);**8**(10):839–45.
59. Vilar E., Gruber S.B. Microsatellite instability in colorectal cancer-the stable evidence. *Nat Rev Clin Oncol*. (2010);**7**:152–162.
60. Duval A., Hamelin R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: Toward a new concept of target genes for instability. *Cancer Res*. (2002);**62**(9):2447–54.
61. Jacob S., Aguado M., Fallik D., Praz F. The role of the DNA mismatch repair system in the cytotoxicity of the topoisomerase inhibitors camptothecin and etoposide to human colorectal cancer cells. *Cancer Res*. (2001);**61**(17):6555–62.
62. Magrini R., Bhone M.R., Hanski M.L., Notter M., Scherbl H., Richard Boland C., et al. Cellular effects of CPT-11 on colon carcinoma cells: Dependence on p53 and hMLH1 status. *Int J Cancer*. (2002);**101**(1):23–31.
63. Son B., Lee S., Youn H., Kim E., Kim W., Youn B. The role of tumor microenvironment in therapeutic resistance. *Oncotarget*. (2017);**8**(3):3933–45.
64. Kopetz S., Lemos R., Powis G. The promise of patient-derived xenografts: The best laid plans of mice and men. *Clin Cancer Res*. (2012);**18**(19):5160–2.
65. Bertotti A., Migliardi G., Galimi F., Sassi F., Torti D., Isella C., et al. A molecularly annotated platform of patient- derived xenografts (“xenopatients”) identifies HER2 as an effective therapeutic target in cetuximab-resistant colorectal cancer. *Cancer Discov*. (2011);**1**(6):508–23.
66. Hidalgo M., Bruckheimer E., Rajeshkumar N. V., Garrido-Laguna I., De Oliveira E., Rubio-Viqueira B., et al. A pilot clinical study of treatment guided by personalized tumorgrafts in patients with advanced cancer. *Mol Cancer Ther*. (2011);**10**(8):1311–6.
67. Wang M., Yao L.C., Cheng M., Cai D., Martinek J., Pan C.X., et al. Humanized mice in studying efficacy and mechanisms of PD-1-targeted cancer immunotherapy. *FASEB J*. (2018);**32**(3):1537–49.
68. Sato T., Vries R.G., Snippert H.J., Van De Wetering M., Barker N., Stange D.E., et al. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*. (2009);**459**(7244):262–5.

69. Huch M., Gehart H., Van Boxtel R., Hamer K., Blokzijl F., Verstegen M.M.A., et al. Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell*. (2015);**160**(1–2):299–312.
70. Van De Wetering M., Francies H.E., Francis J.M., Bounova G., Iorio F., Pronk A., et al. Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell*. (2015);**161**(4):933–45.
71. Dijkstra K.K., Cattaneo C.M., Weeber F., Chalabi M., van de Haar J., Fanchi L.F., et al. Generation of Tumor-Reactive T Cells by Co-culture of Peripheral Blood Lymphocytes and Tumor Organoids. *Cell*. (2018);**174**(6):1586–1598.e12.
72. Tsai S., McOlash L., Palen K., Johnson B., Duris C., Yang Q., et al. Development of primary human pancreatic cancer organoids, matched stromal and immune cells and 3D tumor microenvironment models. *BMC Cancer*. (2018);**18**(1):1–13.
73. Ooft S.N., Weeber F., Dijkstra K.K., McLean C.M., Kaing S., van Werkhoven E., et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Sci Transl Med*. (2019);**11**(513):eaay2574.
74. Ganesh K., Wu C., O'Rourke K.P., Szeglin B.C., Zheng Y., Sauvé C.-E.G., et al. A rectal cancer organoid platform to study individual responses to chemoradiation. *Nat Med*. (2019);**25**(10):1607–14.
75. Vlachogiannis G., Hedayat S., Vatsiou A., Jamin Y., Fernández-mateos J., Khan K., et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science*. (2018);**359**(6378):920–6.
76. Tokunaga E., Oda S., Fukushima M., Maehara Y., Sugimachi K. Differential growth inhibition by 5-fluorouracil in human colorectal carcinoma cell lines. *Eur J Cancer*. (2000);**36**(15):1998–2006.
77. Tajima A., Hess M.T., Cabrera B.L., Kolodner R.D., Carethers J.M. The mismatch repair complex hMutSa recognizes 5-fluorouracil-modified DNA: Implications for chemosensitivity and resistance. *Gastroenterology*. (2004);**127**(6):1678–84.
78. Aebi S., Fink D., Gordon R., Kim H.K., Zheng H., Fink J.L., et al. Resistance to cytotoxic drugs in DNA mismatch repair-deficient cells. *Clin Cancer Res*. (1997);**3**(10):1763–7.
79. Arnold C.N., Goel A., Boland C.R. Role of hMLH1 promoter hypermethylation in drug resistance to 5-fluorouracil in colorectal cancer cell lines. *Int J Cancer*. (2003);**106**(1):66–73.
80. Pommier Y. Topoisomerase I inhibitors: Camptothecins and beyond. *Nat Rev Cancer*. (2006);**6**(10):789–802.
81. Rodriguez R., Hansen L.T., Phear G., Scorch J., Spang-Thomsen M., Cox A., et al. Thymidine selectively enhances growth suppressive effects of camptothecin/irinotecan in MSI+ cells and tumors containing a mutation of MRE11. *Clin Cancer Res*. (2008);**14**(17):5476–83.
82. Fink D., Nebel S., Aebi S., Zheng H., Cenm B., Nehmā A., et al. The Role of DNA Mismatch Repair in Platinum Drug Resistance. *Cancer Res*. (1996);**56**(619):4881–6.
83. Scheeff E.D., Briggs J.M., Howell S.B. Molecular modeling of the intrastrand guanine-guanine DNA adducts produced by cisplatin and oxaliplatin. *Mol Pharmacol*. (1999);**56**(3):633–43.
84. Lee M.S., McGuffey E.J., Morris J.S., Manyam G., Baladandayuthapani V., Wei W., et al. Association of CpG island methylator phenotype and EREG/AREG methylation and expression in colorectal cancer. *Br J Cancer*. (2016);**114**(12):1352–61.

85. Pentheroudakis G., Kotoula V., De Roock W., Kouvatsas G., Papakostas P., Makatsoris T., et al. Biomarkers of benefit from cetuximab-based therapy in metastatic colorectal cancer: Interaction of EGFR ligand expression with RAS/RAF, PIK3CA genotypes. *BMC Cancer*. (2013);**13**(1):1.
86. Wendum D., Boëlle P.Y., Rigau V., Sebbagh N., Olschwang S., Mourra N., et al. Mucinous colon carcinomas with microsatellite instability have a lower microvessel density and lower vascular endothelial growth factor expression. *Virchows Arch*. (2003);**442**(2):111–7.
87. Tian L., Goldstein A., Wang H., Lo H.C., Kim I.S., Welte T., et al. Mutual regulation of tumour vessel normalization and immunostimulatory reprogramming. *Nature*. (2017);**544**(7649):250–4.
88. Meric-Bernstam F., Larkin J., Tabernero J., Bonini C. Enhancing anti-tumour efficacy with immunotherapy combinations. *Lancet*. (2020);**6736**(20).



# MISMATCH REPAIR STATUS IN PATIENT-DERIVED COLORECTAL CANCER ORGANOIDS DOES NOT AFFECT INTRINSIC TUMOUR CELL SENSITIVITY TO SYSTEMIC THERAPY

*Cancers*, 2021, 13, 5434.

Emerens Wensink<sup>†</sup>  
Emre Küçükköse<sup>†</sup>  
Celine Roelse  
Susanne van Schelven  
Daniëlle Raats  
Sylvia Boj  
Miriam Koopman  
Jamila Laoukili  
Jeanine Roodhart<sup>‡</sup>  
Onno Kranenburg<sup>‡</sup>

<sup>†</sup> Both authors share equal contribution.

<sup>‡</sup> Both authors share last authorship.

## ABSTRACT

DNA mismatch repair deficiency (dMMR) in metastatic colorectal cancer (mCRC) is associated with poor survival and a poor response to systemic treatment. However, it is unclear whether dMMR results in a tumour cell-intrinsic state of treatment resistance, or whether alternative mechanisms play a role. To address this, we generated a cohort of MMR-proficient and -deficient patient-derived organoids (PDOs) and tested their response to commonly used drugs in the treatment of mCRC, including 5-fluorouracil (5-FU), oxaliplatin, SN-38, binimetinib, encorafenib, and cetuximab. MMR status did not correlate with the response of PDOs to any of the drugs tested. In contrast, the presence of activating mutations in the *KRAS* and *BRAF* oncogenes was significantly associated with resistance to chemotherapy and sensitivity to drugs targeting oncogene-activated pathways. We conclude that mutant *KRAS* and *BRAF* impact the intrinsic sensitivity of tumour cells to chemotherapy and targeted therapy. By contrast, tumour cell-extrinsic mechanisms—for instance signals derived from the microenvironment—must underlie the association of MMR status with therapy response. Future drug screens on rationally chosen cohorts of PDOs have great potential in developing tailored therapies for specific CRC subtypes including, but not restricted to, those defined by *BRAF/KRAS* and MMR status.

## INTRODUCTION

A deficient DNA mismatch repair (dMMR) system underlies the formation of a specific subtype of colorectal cancer (CRC), accounting for approximately 15% of all cases of primary CRC. In general, dMMR is associated with a lower risk of distant metastasis formation and a relatively good prognosis<sup>1</sup>. However, a minority of patients with dMMR CRC do develop metastatic disease. This group, making up 3–5% of all cases of metastatic CRC (mCRC), is associated with an unfavorable prognosis<sup>2–4</sup> which has been linked to a reduced sensitivity to systemic treatment<sup>5</sup>. Indeed, patients with metastatic dMMR tumours receiving chemotherapy and targeted treatment have a lower response rate (5% versus 44%) and a shorter progression-free survival, when compared to patients with proficient-MMR tumours (pMMR)<sup>6,7</sup>. Moreover, patients with non-metastatic dMMR CRC benefit less from adjuvant treatment with 5-FU monotherapy than patients with pMMR CRC<sup>8–10</sup>.

Immune checkpoint inhibitors (ICI) have revolutionized the treatment of patients with dMMR mCRC, producing durable responses that result in a significantly improved survival<sup>11</sup>. Despite the success of ICI therapy, approximately one third of the patients receiving such treatment display resistance<sup>11</sup>. For those patients, chemotherapy and targeted treatment remain a relevant (often the only) treatment option. However, all mCRC patients receive the same (non-immunotherapy) systemic treatment, regardless of MMR status. Since dMMR status is associated with a worse prognosis and response rate, it is imperative to understand the basis of this differential outcome. Eventually, such knowledge may allow the design of more effective treatment strategies that are based on patient stratification according to MMR status<sup>5</sup>.

Organoid technology can be used to generate patient-derived tumour organoid (PDO) cultures, in which the genetic and functional characteristics of the original tumour are maintained<sup>12–14</sup>. Importantly, clinical responses of colorectal tumours to systemic therapy are faithfully reproduced in PDO-based drug screens *in vitro*<sup>13–15</sup>. In the present study, we generated a series of genetically characterized PDOs from dMMR and pMMR CRC to assess whether tumour cell-intrinsic parameters could explain the clinically observed differences in response to systemic therapy between these CRC subgroups. The generated drug response data were also compared to the presence of driver mutations in *KRAS* and *BRAF*.

## MATERIALS AND METHODS

### Human Specimens

Tissue samples from CRC patients were collected during surgery, or by fine-needle biopsies, within the Biobanking protocol HUB-Cancer TCBIO #12-093, which was approved by the medical ethical committee of the University Medical Center Utrecht (UMCU). Written informed consent from the donors was obtained prior to acquisition of the specimen for research use in the present study.

### In Vitro Organoid Culture

Culturing organoids was performed by embedding in ice-cold Matrigel<sup>®</sup> (Corning), mixed with a CRC culture medium (Table S1) in a 3:1 ratio. For passaging, the tumour organoids were dissociated with TrypLE Express (Gibco, Breda, Netherlands, #12604021) for 5–10 min at 37 °C and re-plated in a pre-warmed 6-well plate. Rho-associated kinase (ROCK) inhibitor Y-27632 (Tocris, Abingdon, United Kingdom, #1254, 10  $\mu$ M) was added to culture medium upon plating for two days.

### Immunohistochemistry (IHC)

PDOs were cultured for 10 days and harvested using 2 mg/mL Dispase type II (Sigma-Aldrich, #D4693) by 15 min incubation at 37 °C. After washout of Dispase type II with PBS (Sigma-Aldrich, Zwijndrecht, Netherlands), the organoids were fixed in 4% formaldehyde solution at room temperature for 20 min. The formaldehyde solution was aspirated and 200  $\mu$ L 2% Agar (Merck) solution was added and hardened on a pre-cooled dish. Subsequently, the agar-droplets containing PDOs were embedded in paraffin blocks. For assessment of mismatch-repair status, immunohistochemistry staining was performed on a BenchMark Ultra Autostainer (Ventana Medical Systems). Briefly, paraffin sections were cut at 4  $\mu$ m and deparaffinized in the instrument with EZ prep solution (Ventana Medical Systems) at 75 °C for 8 min. Heat-induced antigen retrieval was carried out using Cell Conditioning 1 (CC1, Ventana Medical Systems) for 32 min at 100 °C. The following antibodies were used: anti-hMLH1 (BD Pharmingen Amsterdam, Netherlands, G168-15, 1:20), anti-hPMS2 (Roche, Woerden, Netherlands, EPR3947, ready-to-use), anti-hMSH2 (Roche, G219-1129, ready-to-use), and anti-hMSH6 (Abcam, Cambridge, United Kingdom, EPR3945, 1:200). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems).

## DNA Isolation and Whole-Genome-Sequencing (WGS)

PDOs were harvested as described previously in 2.3. Total DNA was isolated using the DNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. Extracted genomic material concentration and quality was measured using NanoDrop 2000 (Thermo Scientific) and Bioanalyzer 2100 (Agilent). Whole genome sequencing of the PDOs was performed on an Illumina NovaSeq 6000 (2 × 150 bp). Mapping and assessment of germline single-nucleotide-variation (SNV) and insertion and/or deletion variants was performed using the "Illumina Analysis Pipeline" of Utrecht Bioinformatics Expertise Core within the Center for Molecular Medicine at the UMC.

## Drug Screens

All drug screens were performed twice on different days. PDOs were mechanically and enzymatically dissociated into single cells by incubating in TrypLE for 5–10 min and replated to allow for the formation of organoids over the course of 72 h. After 72 h, PDOs were collected, incubated with Dispase II (Sigma-Aldrich, #D4693, 2 mg/mL) for 15 min at 37 °C to remove Matrigel, and counted using a hemocytometer and trypan blue. PDOs (50 organoids/μl) were resuspended in 13 mL CRC culture medium without NAC and with low EGF concentration (1 ng/mL) supplemented with 5% Matrigel. PDO suspension (40 μl/well) was dispensed in clear-bottomed, black-walled 384-well plates (Corning, #3904) using an automated Multidrop™ Combi Reagent Dispenser. We generated a twelve-step, threefold drug matrices of 5-FU, oxaliplatin, SN-38, cetuximab, encorafenib, and binimetinib, as described in Table S2 using a Tecan D300e digital dispenser. Readouts were obtained at day six in empty wells, positive control (20 μg/mL puromycin), negative control (1% Dimethyl sulfoxide or 0.3% Tween-20, depending on solvent of agent used), and the drug-treated wells<sup>16</sup>. Quantification was done on a SpectraMax plate reader (Molecular Devices) by measuring cell viability using CellTiter-Glo 3D (Promega, #G9681, 40 μL/well) according to the manufacturer's instructions. Screens with strictly standardized mean difference (SSMD) values <3.0 were classified as poor quality and excluded from the analysis<sup>17</sup>.

## Dose–Response Curve (DRC) Fitting and Statistical Analysis

Drug responses were calculated on the basis of normalized viability data as previously described<sup>16</sup>, and were fitted by log10-transformed drug concentrations

using a 4-parameter logistic regression model using the *nplr* package (v.0.1-7) in R. Nonmonotonic DRCs of 5-FU and encorafenib were fitted using Dr-fit software<sup>18</sup>. Encorafenib's maximum concentration was adjusted after initial analysis at 1  $\mu$ M due to biphasic stimulatory effect on curve fitting at higher concentrations, presumably due to complex formation.

One DRC parameter was analyzed: area under the curve (AUC), inferred by integrating fitted curves using Simpson's rule. To analyze the reproducibility between drug screens, we calculated Pearson's R and corresponding *p*-value using the response values in the drug matrix. Pre-specified CRC subgroups (MMR status and *KRAS/BRAF* mutational-status) were compared using a Mann–Whitney U-test or Kruskal–Wallis test. *p*-values of <0.05 were considered significant.

## RESULTS

### A Collection of dMMR and pMMR CRC Patient-Derived Organoids

To study the tumour cell-intrinsic differences in response to commonly used therapies between dMMR and pMMR CRC, we generated a PDO cohort of six dMMR and six pMMR CRC-derived PDOs with annotated primary tumour location (sidedness) and *BRAF*- and *KRAS*-mutational status (Table 1). The MMR status of all 12 PDOs was assessed by immunohistochemistry analysis of the expression of the MMR proteins MLH1, PMS2, MSH2, and MSH6 (Figure 1). All six pMMR PDOs had normal expression of all MMR proteins. By contrast, five (out of six) dMMR PDOs displayed loss of MLH1 and PMS2 expression, and one PDO (dMMR3) displayed loss of MSH2/MSH6 expression. Whole-genome (DNA) sequencing revealed that dMMR3 and dMMR5 had a loss of function mutation in *MSH2* (p.Glu483\*), and inactivating mutations in *MLH1* (p.I219V) and *MSH3* (p.W1111R), indicating that these PDOs were derived from patients with Lynch syndrome (Figure 2).

**Table 1.** Characteristics of PDOs.

Organoid	MMR Status (IHC)	KRAS Status	BRAF Status	Source of Tissue	Primary Tumour Location (Sidedness)
dMMR1	MLH1/ PMS2	WT	V600E	Metastatic (liver)	Left
dMMR2	MLH1/ PMS2	WT	WT	Metastatic (inguinal lymph node)	Left
dMMR3	MSH2/ MSH6	A146T	WT	Metastatic (peritoneum)	Right
dMMR4	MLH1/ PMS2	A146T	V600E	Primary (synchronous metastatic disease)	Right
dMMR5	MLH1/ PMS2	WT	WT	Primary (unknown if synchronous)	Right
dMMR6	MLH1/ PMS2	WT	V600E	Primary (unknown if synchronous)	Right
pMMR1	Normal	WT	V600E	Primary (synchronous metastatic disease)	Right
pMMR2	Normal	G12A	WT	Metastatic (liver)	Right
pMMR3	Normal	WT	WT	Metastatic (liver)	Right
pMMR4	Normal	G12V	WT	Primary (unknown if synchronous)	Unknown
pMMR5	Normal	G12A	WT	Primary (synchronous metastatic disease)	Right
pMMR6	Normal	WT	WT	Primary (unknown if synchronous)	Unknown

Sidedness of the primary tumour was defined as right-sided (coecum-transverse colon), left-sided (splenic flexure-sigmoid), and rectosigmoid/rectal. Synchronous metastatic disease: if developed metastasis within six months of primary CRC diagnosis.

*Abbreviations:* IHC (immunohistochemistry), dMMR (deficient mismatch repair), MMR (mismatch repair), PDOs (patient-derived organoids), pMMR (proficient mismatch repair), WT (wildtype).

Figure 1.

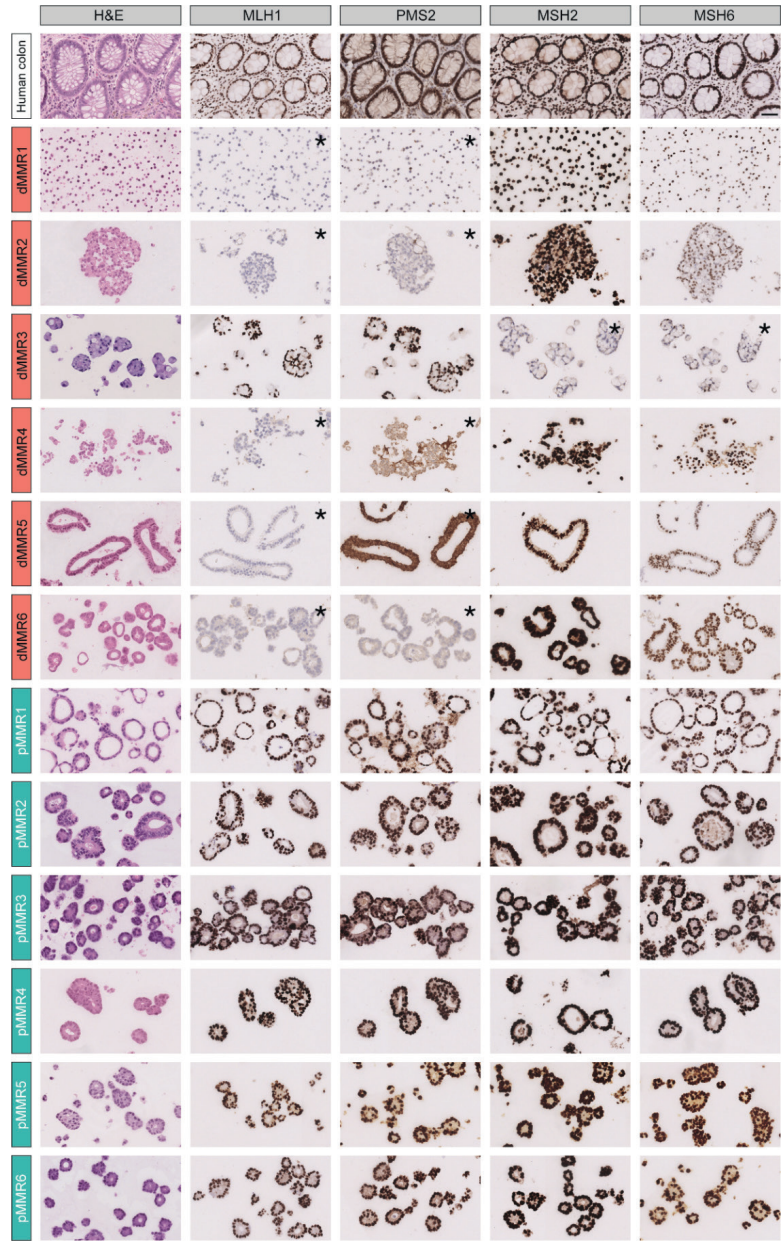


Fig 1. Expression of mismatch repairs proteins (MLH1, PMS2, MSH2 and MSH6) in PDOs analyzed by immunohistochemistry. \* demonstrates loss of expression. Scale bar shown in human colon MSH6 staining is 50  $\mu$ m and applies for all images.

Figure 2.

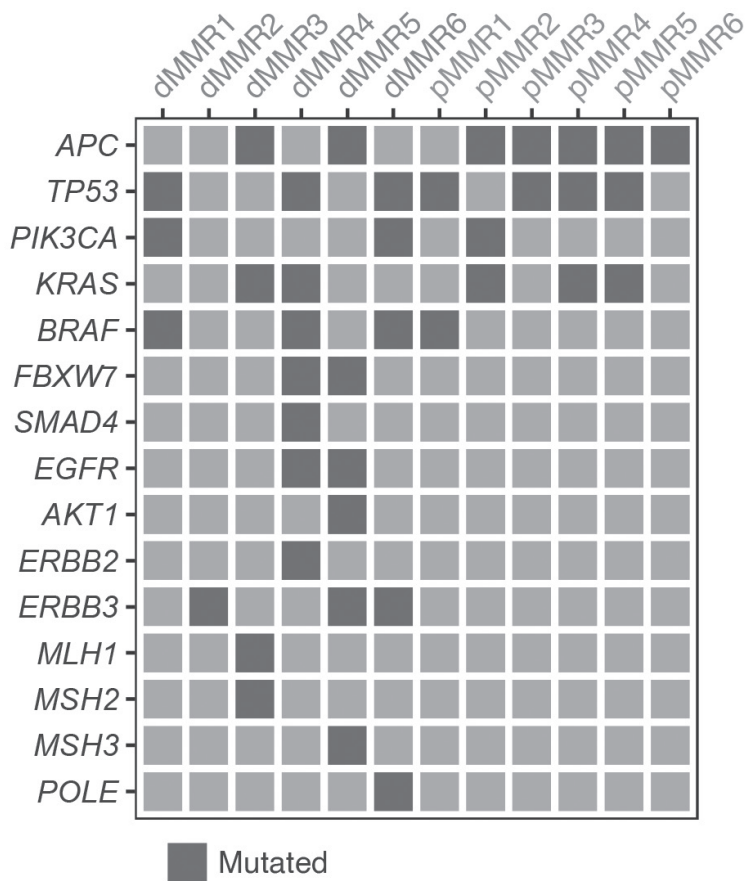


Fig 2. Mutational status of genes frequently associated with CRC. Overview of all PDOs with mutations in genes commonly mutated in CRC. “Mutated” was defined as a given variant being predicted pathogenic by COSMIC.

### Comparable Sensitivity of dMMR and pMMR PDOs to Chemotherapy and Targeted Therapy

All PDOs were screened for their sensitivity to commonly used drugs in the treatment of metastatic CRC, including the chemotherapeutic drugs 5-FU, oxaliplatin and irinotecan (SN-38), the EGFR-targeting antibody cetuximab, the BRAF inhibitor encorafenib, and the MEK inhibitor binimetinib (Figure 3A). All 12 PDOs were

exposed to concentration series of all drugs for a period of three days. Cell Titer Glo was then used to measure ATP levels as a proxy for the amount of viable (metabolically active) cells. Drug effects were calculated by normalizing the recorded values against those from drug solvent-exposed control wells (Table S3). The quality of all drug screens was analyzed using the strictly standardized mean difference (SSMD) and the observed correlation between replicates. Screens were excluded from subsequent analyses if they had an SSMD < 3.0. The SSMD of all assays for all PDOs are shown in Figure S1. Based on these quality assessments, six drug screens were excluded and 24 screens were included in the subsequent analyses. All PDOs were screened against all drugs twice on different days. The normalized cell viabilities of the biological triplicates of the analyzed assays were significantly correlated (*Pearson's*  $R = 0.73 - 0.76$ ;  $p < 0.05$ , Figure S2). In addition, the variation was minimal (1.09%) between the duplicate assays of each PDO (Figure S3).

To investigate whether MMR status correlated with the sensitivity of CRC PDOs to any of the tested drugs, we generated dose response curves (DRC) and calculated the area under the curve (AUC) values for each of the fitted curves. We found no significant differences in sensitivity to any of the drugs between the groups of dMMR and pMMR PDOs (Figures 3B,C and S4). Thus, as a group, dMMR PDOs are not intrinsically more or less sensitive to chemotherapy and targeted treatment when compared to pMMR PDOs *in vitro*.

Figure 3.

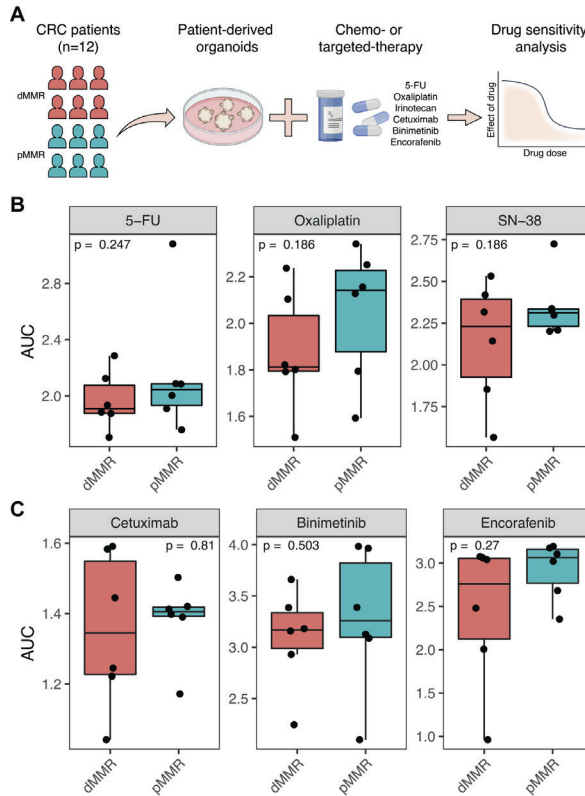


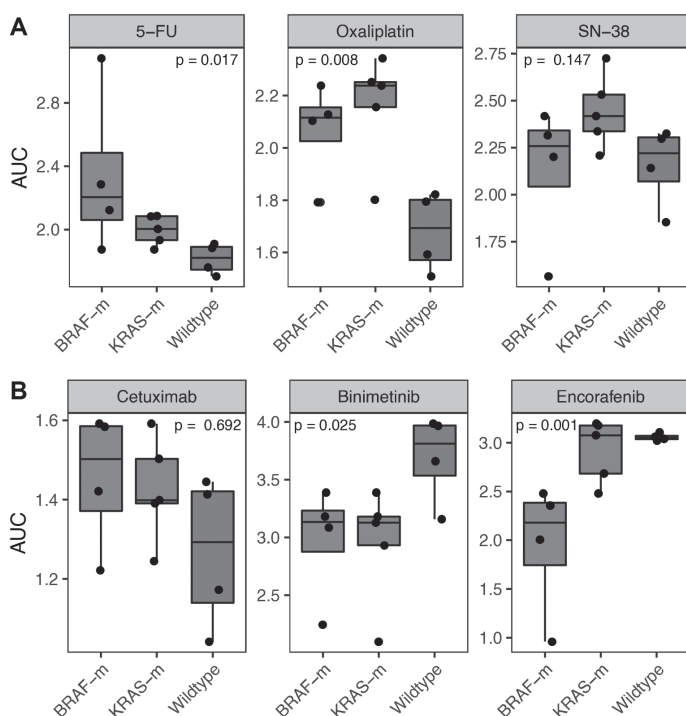
Fig 3. Sensitivity of dMMR versus pMMR PDOs to systemic treatment. (A) Schematic overview of the study. Per treatment type, boxplots display the average of each drug response curve (DRC) parameter per organoid line, grouped per mismatch repair (MMR) status, with significance tested using the Mann-Whitney U-test. (B) The AUC for chemotherapy (5-FU, oxaliplatin and SN-38 (irinotecan)) are displayed. (C) Similarly, the AUCs for targeted treatment (cetuximab, binimetinib and encorafenib) are displayed.

## Association of BRAF and KRAS Status with Response to Chemotherapy and Targeted Therapy

While the study was designed to compare PDOs from dMMR and pMMR CRC, additional information on the PDO cohort allowed us to also compare drug responses between subgroups differing in driver gene mutation status (*BRAF*, *KRAS*). These analyses revealed that *BRAF*- and *KRAS*-mutant PDOs were significantly ( $p = 0.017$  and  $p = 0.008$ ) more resistant to 5-FU and oxaliplatin when compared to *BRAF/KRAS* wildtype PDOs (Figures 4A and S5). By contrast, *BRAF/KRAS* status had no significant effect on the sensitivity of PDOs to irinotecan/SN-38 (Figure 4A).

*BRAF*-mutant PDOs showed a significantly higher sensitivity to the *BRAF* inhibitor encorafenib than *KRAS*-mutant or *BRAF/KRAS* wildtype PDOs (Figure 4B). Both *BRAF*- and *KRAS*-mutated PDOs displayed a significantly increased sensitivity to the MEK inhibitor binimetinib when compared to wildtype PDOs. There was no significant difference in cetuximab sensitivity among PDOs based on *BRAF/KRAS* mutational status (Figure 4B). This is perhaps surprising, given the clinical correlation between the presence of mutations in *KRAS* and *BRAF* with resistance to cetuximab<sup>19–21</sup>. However, one of the cetuximab-resistant PDOs (dMMR2) with wildtype *KRAS* and *BRAF* did contain an activating mutation in *ERBB3*, which is also known to confer resistance to EGFR inhibition<sup>22</sup> (Figure 2).

**Figure 4.**



**Fig 4.** Sensitivity of *BRAF*-mutant versus *KRAS*-mutant versus wildtype PDOs to systemic treatment. Per treatment type (5-FU, oxaliplatin, SN-38 (irinotecan), cetuximab, binimetinib, encorafenib), boxplots display the average of each drug response curve (DRC) parameter per organoid line, grouped per mutational status (*BRAF*-mutant (*BRAF*-m), *KRAS*-mutant (*KRAS*-m) or wildtype for *BRAF* and *KRAS* (Wildtype)). (A) The AUC for chemotherapy (5-FU, oxaliplatin and SN-38 (irinotecan) are displayed, significance tested using the Kruskal-Wallis test. (B) Similarly, the AUCs for targeted treatment (cetuximab, binimetinib and encorafenib) are displayed. Organoid line dMMR4 has both a *KRAS* and *BRAF* mutation, so was included in both categories.

## DISCUSSION

The data presented in this study show that MMR status does not affect the intrinsic sensitivity of PDOs to chemotherapy or targeted therapy. The poor prognosis of dMMR mCRC patients can therefore not simply be explained by a reduced sensitivity of dMMR tumour cells to systemic therapy. An alternative explanation could be that the distinct microenvironments of pMMR and dMMR CRC could influence treatment response<sup>23</sup>. Addressing this experimentally would require the generation of PDO co-culture models in which these distinct microenvironments are modeled, for instance by co-culturing PDOs with specific immune cell types, or by transplanting PDOs into mice with a human immune system<sup>24,25</sup>. Insight into the mechanisms underlying the poor clinical response of dMMR mCRC to systemic therapy is essential in order to design more effective treatment strategies. Until such mechanistic insight is provided, the only available treatment options for immunotherapy-refractory dMMR mCRC remain standard chemotherapy schedules with or without targeted therapy. Some studies indicate that MMR status may be associated with an increased or reduced benefit from specific treatment schedules, but solid evidence for such relationships is currently lacking<sup>26–29</sup>.

This study was designed to study the effect of MMR status on the tumour cell-intrinsic responsiveness to a series of clinically relevant drugs. A potentially relevant variable that we were unable to address in this study was the origin of the dMMR phenotype: sporadic versus inherited (Lynch syndrome). Likewise, tumour sidedness is a potentially important variable in determining treatment responses. Studies with specifically selected (or generated) PDO cohorts (i.e., sporadic dMMR versus Lynch syndrome; left-sided versus right sided, etc.) are required to address these questions in the future.

Many studies have found correlations between the presence of specific genetic alterations (including but not limited to mutations in *KRAS* and *BRAF*) with resistance to chemotherapy. However, none of these are currently used to select patients for, or exclude them from, treatment with chemotherapy<sup>30</sup>. By contrast, specific genetic alterations are used to select patients for targeted therapy. Examples are the selection of patients with mutant *BRAF* tumours for treatment with BRAF inhibitor combination therapies, the selection of patients with dMMR tumours for immunotherapy, and the exclusion of patients with *KRAS*-mutant tumours from treatment with EGFR-targeting therapy<sup>30</sup>. Although this study was not designed to study a potential effect of mutant *BRAF* and *KRAS* on treatment response, we found that within the PDO

cohort used, the presence of these oncogenes was significantly associated with reduced sensitivity to 5-FU and oxaliplatin. An intrinsic (relative) resistance of tumour cells with mutant *BRAF* or *KRAS* to chemotherapy may contribute to the poor prognosis that is associated with the presence of these oncogenes<sup>31,32</sup>. The data also show that mutant *BRAF* PDOs are more sensitive to inhibitors of BRAF and that *KRAS*- and *BRAF*-mutant PDOs are more sensitive to inhibition of the BRAF target MEK. The results therefore support the concept of targeted inhibition of oncogene-activated signaling pathways as a basis for future treatment strategies, involving the concomitant suppression of feedback pathways.

## CONCLUSIONS

By using PDOs as a model system for evaluating clinically relevant drug responses, we conclude that *KRAS* and *BRAF*, but not MMR status, are dominant factors in determining the intrinsic sensitivity of tumour cells to chemotherapy and targeted therapy. Future drug screens on rationally chosen cohorts of PDOs have great potential in developing tailored therapies for specific CRC subtypes including, but not restricted to, those defined by *BRAF/KRAS* and MMR status.

## ACKNOWLEDGMENTS

The authors thank all patients involved in the current study.

## REFERENCES

1. Popat S., Hubner R., Houlston R.S. Systematic Review of Microsatellite Instability and Colorectal Cancer Prognosis. *J Clin Oncol.* (2005);**23**(3):609–18.
2. Biller L.H., Schrag D. Diagnosis and treatment of metastatic colorectal cancer: A review. *JAMA.* (2021);**325**(7):669–85.
3. Wang B., Li F., Zhou X., Ma Y., Fu W. Is microsatellite instability-high really a favorable prognostic factor for advanced colorectal cancer? A meta-analysis. *World J Surg Oncol.* (2019);**17**(1):1–13.
4. Venderbosch S., Nagtegaal I.D., Maughan T.S., Smith C.G., Cheadle J.P., Fisher D., et al. Mismatch repair status and BRAF mutation status in metastatic colorectal cancer patients: A pooled analysis of the CAIRO, CAIRO2, COIN, and FOCUS studies. *Clin Cancer Res.* (2014);**20**(20):5322–30.
5. Wensink E., Bond M., Kucukkose E., May A., Vink G., Koopman M., et al. A review of the sensitivity of metastatic colorectal cancer patients with deficient mismatch repair to standard-of-care chemotherapy and monoclonal antibodies, with recommendations for future research. *Cancer Treat Rev.* (2021);**95**:102174.
6. Aasebø K., Dragomir A., Sundström M., Mezheyski A., Edqvist P.H., Eide G.E., et al. Consequences of a high incidence of microsatellite instability and BRAF-mutated tumors: A population-based cohort of metastatic colorectal cancer patients. *Cancer Med.* (2019);**8**(7):3623–35.
7. Tougeron D., Sueur B., Zaanen A., de la Fouchardière C., Sefrioui D., Lecomte T., et al. Prognosis and chemosensitivity of deficient MMR phenotype in patients with metastatic colorectal cancer: An AGEO retrospective multicenter study. *Int J Cancer.* (2020);**147**:285–96.
8. Cohen R., Taieb J., Fiskum J., Yothers G., Goldberg R., Yoshino T., et al. Microsatellite Instability in Patients With Stage III Colon Cancer Receiving Fluoropyrimidine With or Without Oxaliplatin: An ACCENT Pooled Analysis of 12 Adjuvant Trials. *J Clin Oncol.* (2020);**39**:642–51.
9. Sargent D.J., Marsoni S., Monges G., Thibodeau S.N., Labianca R., Hamilton S.R., et al. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J Clin Oncol.* (2010);**28**(20):3219–26.
10. Ribic C.M., Sargent D.J., Moore M.J., Thibodeau S.N., French A.J., Goldberg R.M., et al. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N Engl J Med.* (2003);**349**(3):247–57.
11. André T., Shiu K.-K., Kim T.W., Jensen B.V., Jensen L.H., Punt C., et al. Pembrolizumab in Microsatellite-Instability–High Advanced Colorectal Cancer. *NEJM.* (2020);**383**(23):2207–18.
12. Sato T., Stange D.E., Ferrante M., Vries R.G.J., Van Es J.H., Van Den Brink S., et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology.* (2011);**141**(5):1762–72.
13. Vlachogiannis G., Hedayat S., Vatsiou A., Jamin Y., Fernández-mateos J., Khan K., et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science.* (2018);**359**(6378):920–6.

14. Ooft S.N., Weeber F., Dijkstra K.K., McLean C.M., Kaing S., van Werkhoven E., et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Sci Transl Med.* (2019);**11**(513):eaay2574.
15. Wensink G.E., Elias S.G., Mullenders J., Koopman M., Boj S.F., Kranenburg O.W., et al. Patient-derived organoids as a predictive biomarker for treatment response in cancer patients. *npj Precis Oncol.* (2021);**5**(30).
16. Driehuis E., Kretzschmar K., Clevers H. Establishment of patient-derived cancer organoids for drug-screening applications. *Nat Protoc.* (2020);**15**(10):3380–409.
17. Mpindi J.P., Swapnil P., Dmitrii B., Jani S., Saeed K., Wennerberg K., et al. Impact of normalization methods on high-throughput screening data with high hit rates and drug testing with dose-response data. *Bioinformatics.* (2015);**31**(23):3815–21.
18. Di Veroli G.Y., Fornari C., Goldlust I., Mills G., Koh S.B., Bramhall J.L., et al. An automated fitting procedure and software for dose-response curves with multiphasic features. *Sci Rep.* (2015);**5**:14701.
19. Roock W. De, Jonker D.J., Nicolantonio F. Di, Sartore-bianchi A., Simes J., Karapetis C.S., et al. Association of KRAS p . G13D Mutation With Outcome in Patients With Chemotherapy-Refractory Metastatic Colorectal Cancer Treated with Cetuximab. *JAMA.* (2010);**304**(16):1812–20.
20. Karapetis C. s., Khambata-Ford S., Jonker D.J., O'Callaghan C.J., Tu D., Tebbutt N.C., et al. K-ras Mutations and Benefit from Cetuximab in Advanced Colorectal Cancer. *NEJM.* (2008);**359**(17):1757–65.
21. Di Nicolantonio F., Martini M., Molinari F., Sartore-Bianchi A., Arena S., Saletti P., et al. Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J Clin Oncol.* (2008);**26**(35):5705–12.
22. Bosch-Vilaró A., Jacobs B., Pomella V., Asbagh L.A., Kirkland R., Michel J., et al. Feedback activation of HER3 attenuates response to EGFR inhibitors in colon cancer cells. *Oncotarget.* (2017);**8**(3):4277–88.
23. McMillin D.W., Negri J.M., Mitsiades C.S. The role of tumour–stromal interactions in modifying drug response: challenges and opportunities. *Nat Rev Drug Discov.* (2013);**12**(March):217–28.
24. Dijkstra K.K., Cattaneo C.M., Weeber F., Chalabi M., van de Haar J., Fanchi L.F., et al. Generation of Tumor-Reactive T Cells by Co-culture of Peripheral Blood Lymphocytes and Tumor Organoids. *Cell.* (2018);**174**(6):1586–1598.e12.
25. Neal J.T., Li X., Zhu J., Zheng G.X.Y., Davis M.M., Kuo C.J., et al. Organoid Modeling of the Tumor Immune Resource Organoid Modeling of the Tumor Immune Microenvironment. *Cell.* (2018);**175**(7):1972–1988.e16.
26. Braun M.S., Richman S.D., Quirke P., Daly C., Adlard J.W., Elliott F., et al. Predictive Biomarkers of Chemotherapy Efficacy in Colorectal Cancer: Results From the UK MRC FOCUS Trial. *J Clin Oncol.* (2008);**26**(16):2690–8.
27. Innocenti F., Ou F.-S., Qu X., Zemla T.J., Niedzwiecki D., Tam R., et al. Mutational Analysis of Patients With Colorectal Cancer in CALGB/SWOG 80405 Identifies New Roles of Microsatellite Instability and Tumor Mutational Burden for Patient Outcome. *J Clin Oncol.* (2019);**37**(14):1217–27.

28. Koopman M., Kortman G.A.M., Mekenkamp L., Ligtenberg M.J.L., Hoogerbrugge N., Antonini N.F., et al. Deficient mismatch repair system in patients with sporadic advanced colorectal cancer. *Br J Cancer*. (2009);**100**(2):266–73.
29. Smith C.G., Fisher D., Claes B., Maughan T.S., Idziaszczyk S., Peuteman G., et al. Somatic profiling of the epidermal growth factor receptor pathway in tumors from patients with advanced colorectal cancer treated with chemotherapy  $\pm$  cetuximab. *Clin Cancer Res*. (2013);**19**(15):4104–13.
30. Punt C.J.A., Koopman M., Vermeulen L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat Rev Clin Oncol*. (2017);**14**(4):235–46.
31. Ogino S., Shima K., Meyerhardt J.A., McCleary N.J., Ng K., Hollis D., et al. Predictive and Prognostic Roles of BRAF Mutation in Stage III Colon Cancer: Results from Intergroup Trial CALGB 89803. *Clin Cancer Res*. (2012);**18**(3):890–900.
32. Hutchins G., Southward K., Handley K., Magill L., Beaumont C., Stahlschmidt J., et al. Value of mismatch repair, KRAS, and BRAF mutations in predicting recurrence and benefits from chemotherapy in colorectal cancer. *J Clin Oncol*. (2011);**29**(10):1261–70.

## SUPPLEMENTARY MATERIAL

**Supplementary Table 1.** Composition of organoid culture medium

Component	Source (catalogue number)	Concentration
Advanced (DMEM/F12) medium	Gibco (12634-010)	1x (500 mL total)
HEPES Buffer	Lonza (17737E)	10 mM
Penicillin/Streptomycin	Gibco (15070-063)	50 U/ml
GlutaMAX	Gibco (35050-038)	2 mM
R-Spondin conditioned medium	293T-HA-RspoI-F cell line	20%
Noggin conditioned medium	293T-mNoggin-Fc cell line	100 ng/ml
B27	Invitrogen (17504-044)	1x
Nicotinamide	Sigma-Aldrich (N0636)	10 mM
Prostaglandin E2	Tocris (2296-10)	10 nM
Gastrin	Sigma-Aldrich (G9145)	10 nM
N-acetylcysteine (NAC)	Sigma-Aldrich (A9165)	1 mM
A83-01	SignalChem (A09-900-05)	500 nM
Human recombinant EGF	Sigma Aldrich (A9165)	50 ng/mL
SB202190	Gentaur (A1632)	10 $\mu$ M

The table describes the composition of the organoid culture medium.

*Abbreviations:* DMEM/F12 (Dulbecco's Modified Eagle Medium/Ham's F-12), EGF (epidermal growth factor).

**Supplementary Table 2.** Overview of chemotherapies and targeted treatments used in drug screens

Drug	Target	Patient C <sub>MAX</sub>	Chosen concentration range			
5-FU	DNA synthesis	0.88 $\mu$ M <sup>1</sup>	1.5	-	300	$\mu$ M
Oxaliplatin	DNA-alkalyting agent	4.96 $\mu$ M <sup>1</sup>	0.5	-	200	$\mu$ M
SN-38 (irinotecan)	Topoisomerase I	0.014-0.143 $\mu$ M <sup>1</sup>	0.0002	-	0.1	$\mu$ M
Binimetinib	MEK	1.0-1.2 $\mu$ M <sup>2</sup>	0.001	-	15	$\mu$ M
Cetuximab	EGFR	205 $\mu$ g/ml <sup>1</sup>	10	-	150	$\mu$ g/ml
Encorafenib	<i>BRAF</i>	1.12-4.17 $\mu$ M <sup>2</sup>	0.01	-	50	$\mu$ M
Puromycin	Protein synthesis	-			20	$\mu$ g/ml

The table indicates the agents and concentration range used in drug screen assays. As well, the published patient C<sub>max</sub> values are also indicated with the appropriate references.

## References

1. Liston, D. R. & Davis, M. Clinically relevant concentrations of anticancer drugs: A guide for nonclinical studies. *Clin. Cancer Res.* 23, 3489–3498 (2017);
2. Assessment report Mektovi (binimetinib). European Medicine Agency (2018). Available at: <https://www.ema.europa.eu/en/medicines/human/EPAR/mektovi>;
3. Van Geel, R. M. J. M. et al. A phase Ib dose-escalation study of encorafenib and cetuximab with or without alpelisib in metastatic *BRAF*-mutant colorectal cancer. *Cancer Discov.* 7, 610–619 (2017).

**Supplementary Table 3.** Overview drug response AUCs per organoid and agent.

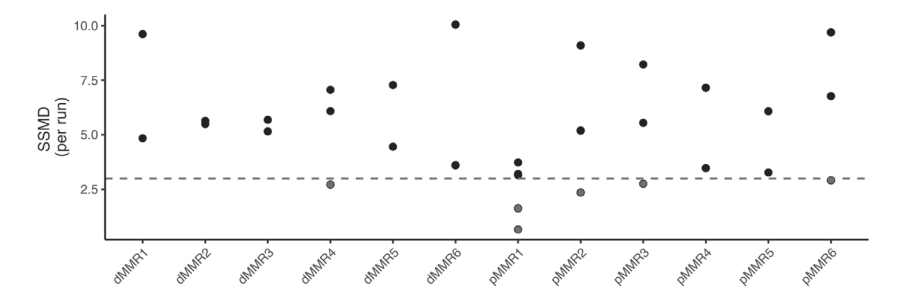
Organoid ID	<i>BRAF</i> status	<i>KRAS</i> status	Agent	AUC
dMMR1	<i>BRAF</i> -m (V600E)	WT	5-FU	2,12
dMMR2	WT	WT	5-FU	1,71
dMMR3	WT	<i>KRAS</i> -m (A146T)	5-FU	1,93
dMMR4	<i>BRAF</i> -m (V600E)	<i>KRAS</i> -m (A146T)	5-FU	1,87
dMMR5	WT	WT	5-FU	1,88
dMMR6	<i>BRAF</i> -m (V600E)	WT	5-FU	2,29
pMMR1	<i>BRAF</i> -m	WT	5-FU	3,08
pMMR2	WT	<i>KRAS</i> -m	5-FU	2,09
pMMR3	WT	WT	5-FU	1,91
pMMR4	WT	<i>KRAS</i> -m	5-FU	2,09
pMMR5	WT	<i>KRAS</i> -m	5-FU	2,00
pMMR6	WT	WT	5-FU	1,76
dMMR1	<i>BRAF</i> -m (V600E)	WT	Binimetinib	2,24
dMMR2	WT	WT	Binimetinib	3,16
dMMR3	WT	<i>KRAS</i> -m (A146T)	Binimetinib	2,93
dMMR4	<i>BRAF</i> -m (V600E)	<i>KRAS</i> -m (A146T)	Binimetinib	3,18
dMMR5	WT	WT	Binimetinib	3,66
dMMR6	<i>BRAF</i> -m (V600E)	WT	Binimetinib	3,39
pMMR1	<i>BRAF</i> -m	WT	Binimetinib	3,09
pMMR2	WT	<i>KRAS</i> -m	Binimetinib	3,13
pMMR3	WT	WT	Binimetinib	3,97
pMMR4	WT	<i>KRAS</i> -m	Binimetinib	3,39
pMMR5	WT	<i>KRAS</i> -m	Binimetinib	2,10
pMMR6	WT	WT	Binimetinib	3,98
dMMR1	<i>BRAF</i> -m (V600E)	WT	Cetuximab	1,22
dMMR2	WT	WT	Cetuximab	1,04

<b>Organoid ID</b>	<b><i>BRAF</i> status</b>	<b><i>KRAS</i> status</b>	<b>Agent</b>	<b>AUC</b>
dMMR3	WT	<i>KRAS</i> -m (A146T)	Cetuximab	1,25
dMMR4	<i>BRAF</i> -m (V600E)	<i>KRAS</i> -m (A146T)	Cetuximab	1,59
dMMR5	WT	WT	Cetuximab	1,44
dMMR6	<i>BRAF</i> -m (V600E)	WT	Cetuximab	1,58
pMMR1	<i>BRAF</i> -m	WT	Cetuximab	1,42
pMMR2	WT	<i>KRAS</i> -m	Cetuximab	1,40
pMMR3	WT	WT	Cetuximab	1,41
pMMR4	WT	<i>KRAS</i> -m	Cetuximab	1,39
pMMR5	WT	<i>KRAS</i> -m	Cetuximab	1,50
pMMR6	WT	WT	Cetuximab	1,17
dMMR1	<i>BRAF</i> -m (V600E)	WT	Encorafenib	0,96
dMMR2	WT	WT	Encorafenib	3,04
dMMR3	WT	<i>KRAS</i> -m (A146T)	Encorafenib	3,08
dMMR4	<i>BRAF</i> -m (V600E)	<i>KRAS</i> -m (A146T)	Encorafenib	2,48
dMMR5	WT	WT	Encorafenib	3,06
dMMR6	<i>BRAF</i> -m (V600E)	WT	Encorafenib	2,01
pMMR1	<i>BRAF</i> -m	WT	Encorafenib	2,35
pMMR2	WT	<i>KRAS</i> -m	Encorafenib	2,68
pMMR3	WT	WT	Encorafenib	3,02
pMMR4	WT	<i>KRAS</i> -m	Encorafenib	3,19
pMMR5	WT	<i>KRAS</i> -m	Encorafenib	3,18
pMMR6	WT	WT	Encorafenib	3,11
dMMR1	<i>BRAF</i> -m (V600E)	WT	Oxaliplatin	1,79
dMMR2	WT	WT	Oxaliplatin	1,82
dMMR3	WT	<i>KRAS</i> -m (A146T)	Oxaliplatin	1,80
dMMR4	<i>BRAF</i> -m (V600E)	<i>KRAS</i> -m (A146T)	Oxaliplatin	2,24
dMMR5	WT	WT	Oxaliplatin	1,51
dMMR6	<i>BRAF</i> -m (V600E)	WT	Oxaliplatin	2,10
pMMR1	<i>BRAF</i> -m	WT	Oxaliplatin	2,13
pMMR2	WT	<i>KRAS</i> -m	Oxaliplatin	2,16
pMMR3	WT	WT	Oxaliplatin	1,79
pMMR4	WT	<i>KRAS</i> -m	Oxaliplatin	2,34
pMMR5	WT	<i>KRAS</i> -m	Oxaliplatin	2,25
pMMR6	WT	WT	Oxaliplatin	1,59
dMMR1	<i>BRAF</i> -m (V600E)	WT	SN-38	2,32

Organoid ID	<i>BRAF</i> status	<i>KRAS</i> status	Agent	AUC
dMMR2	WT	WT	SN-38	2,14
dMMR3	WT	<i>KRAS</i> -m (A146T)	SN-38	2,53
dMMR4	<i>BRAF</i> -m (V600E)	<i>KRAS</i> -m (A146T)	SN-38	2,42
dMMR5	WT	WT	SN-38	1,85
dMMR6	<i>BRAF</i> -m (V600E)	WT	SN-38	1,57
pMMR1	<i>BRAF</i> -m	WT	SN-38	2,20
pMMR2	WT	<i>KRAS</i> -m	SN-38	2,72
pMMR3	WT	WT	SN-38	2,30
pMMR4	WT	<i>KRAS</i> -m	SN-38	2,34
pMMR5	WT	<i>KRAS</i> -m	SN-38	2,21
pMMR6	WT	WT	SN-38	2,33

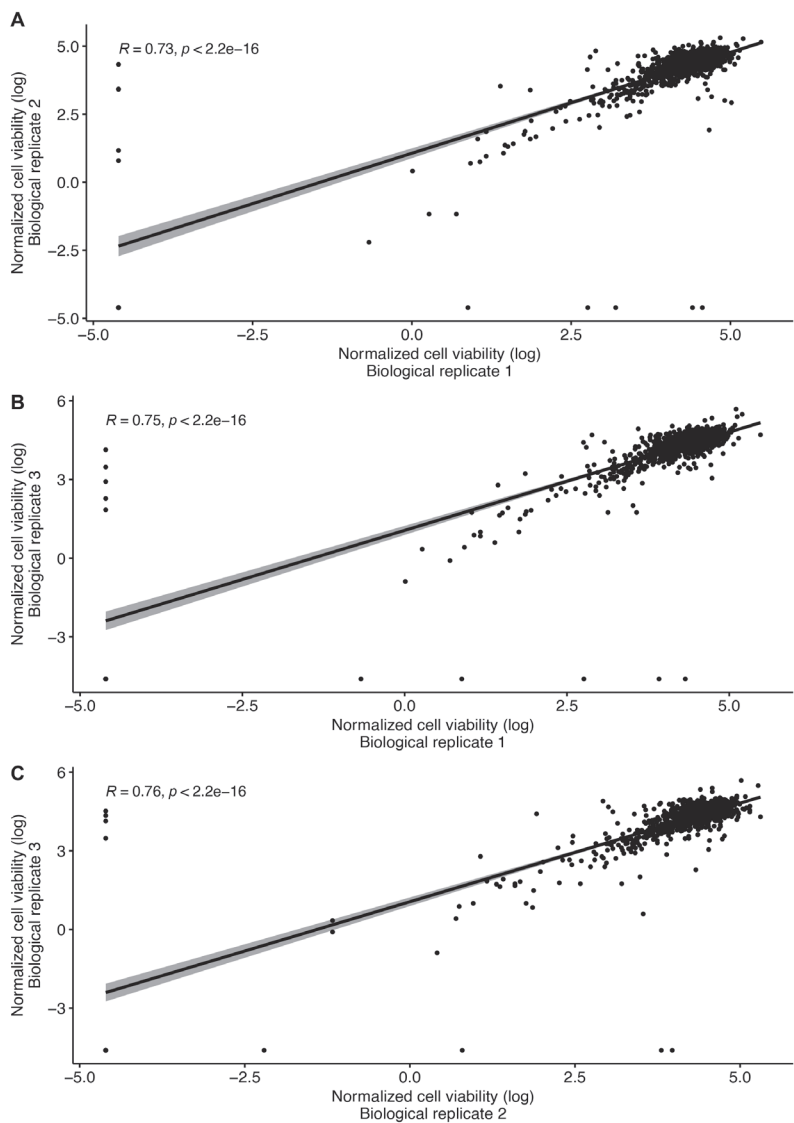
Abbreviations: 5-FU (5-fluorouracil), AUC (area under the curve, obtained using Simpson's rules), *BRAF*-m (*BRAF*-mutant), *KRAS*-m (*KRAS*-mutant), WT (*BRAF* & *KRAS* wildtype).

Supplementary Figure 1. Quality control parameters of drug screens.



Suppl. Fig. 1. The Strictly Standardized Mean Difference (SSMD) values are plotted for each analyzed run for a given organoid line. Organoid drug screen runs which were excluded from the analysis (due to SSMD  $\leq 3.0$ ) are indicated with a gray point.

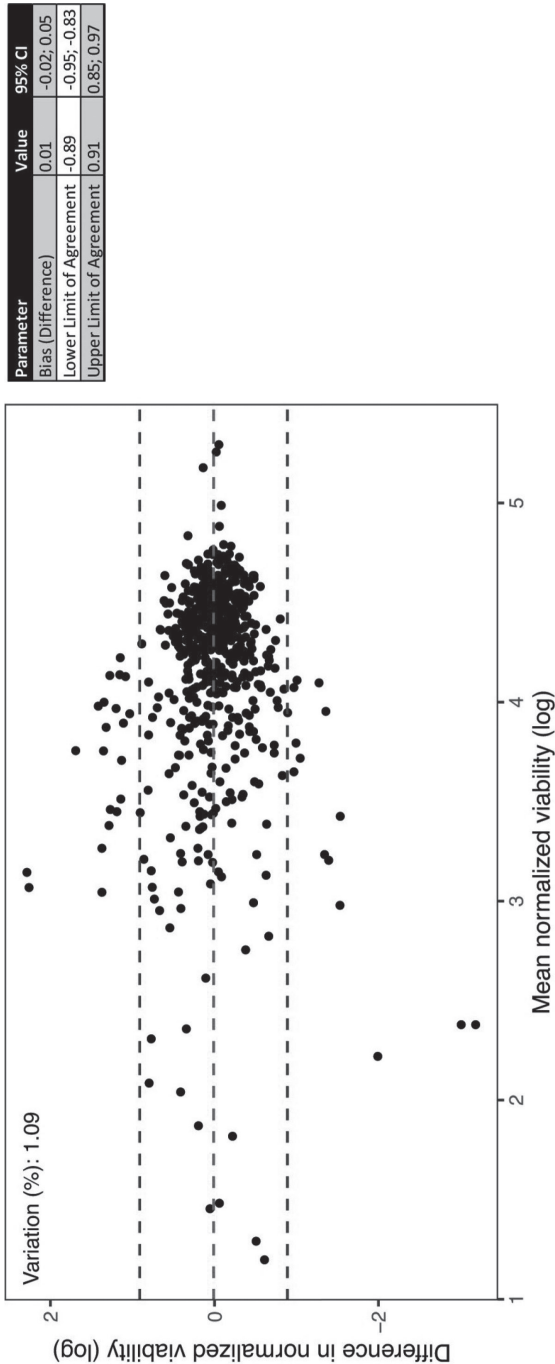
**Supplementary Figure 2.** Correlation of biological replicates of normalized cell viability.



Supplementary Figure S2

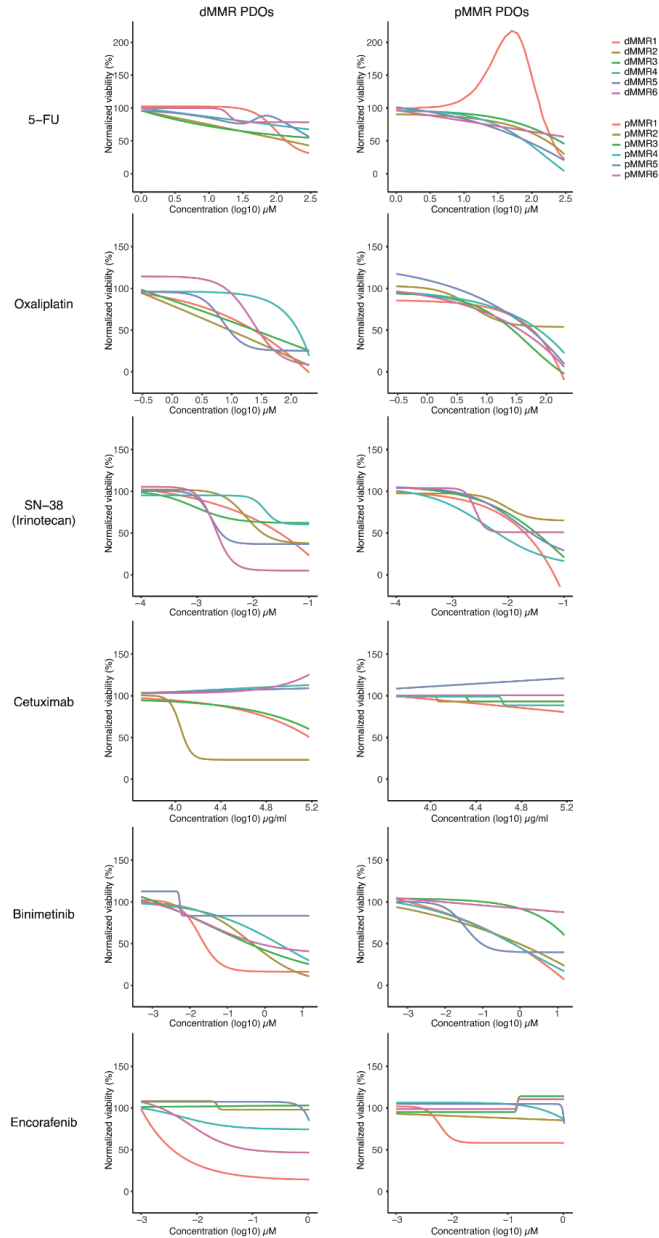
Suppl. Fig. 2. Normalized cell viability values (log) for all analyzed drug screens are plotted for each biological and technical replicate and the Pearson's R correlation values are displayed. A) Biological replicate 1 versus 2; B) Biological replicate 1 versus 3; C) Biological replicate 2 versus 3.

Supplementary Figure 3. Difference between duplicate assays.



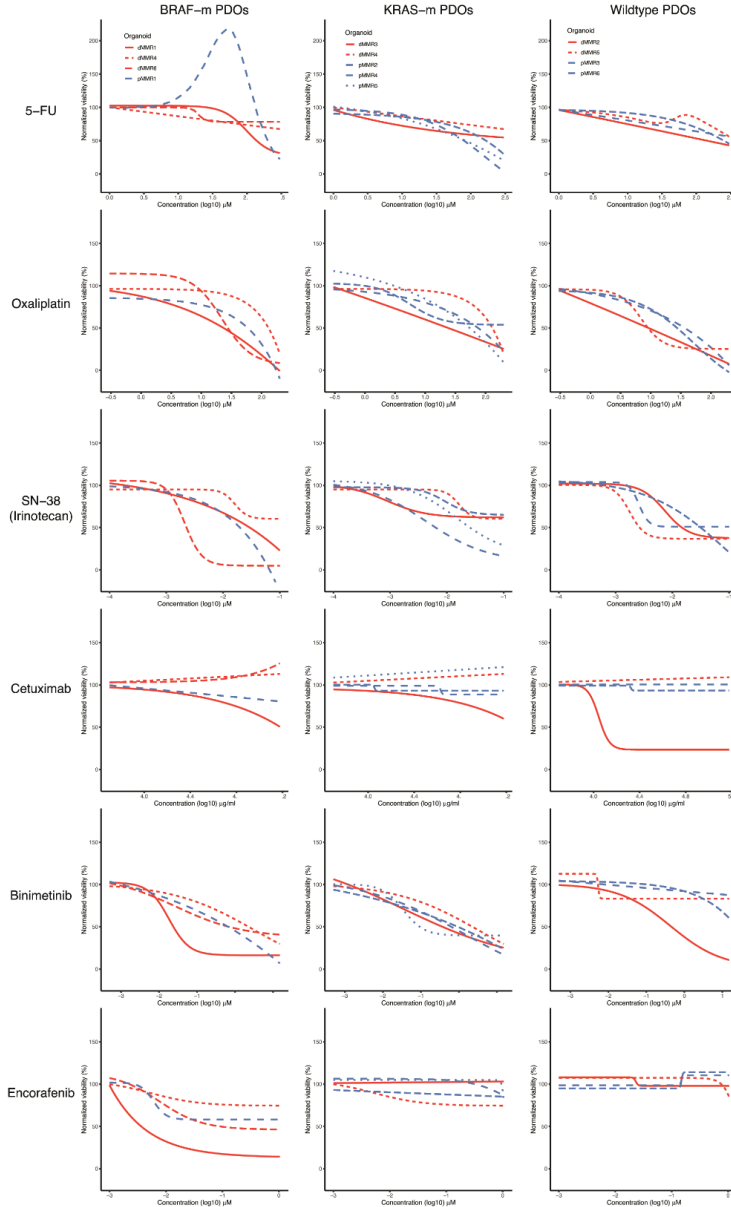
Suppl. Fig. 3. A Bland-Altman plot of the mean normalized viability of the replicates (log) versus the difference in normalized viability between replicates (log) is displayed. The middle dashed line indicates the average difference observed between the replicates and the outer dashed lines are the standard deviation of the differences.

**Supplementary Figure 4.** Drug response curves for each systemic therapy agent (dMMR vs pMMR PDOs).



Suppl. Fig. 4. An overlay of the individual drug response curves (DRC) for each organoid are displayed. Two side-by-side plots were created per treatment type: dMMR PDOs versus pMMR PDOs.

**Supplementary Figure 5.** Drug response curves for each systemic therapy agent (*BRAF*-mutant versus *KRAS*-mutant versus wildtype PDOs).



Suppl. Fig. 5. An overlay of the individual drug response curves (DRC) for each organoid are displayed, for run 2. Two side-by-side plots were created per treatment type: *BRAF*-mutant (*BRAF*-m) in dark red color, *KRAS*-mutant (*KRAS*-m) in orange versus *BRAF* & *KRAS* wildtype (Wildtype) PDOs in blue.





# **PATIENT-DERIVED ORGANOIDS AS A PREDICTIVE BIOMARKER TO GUIDE CANCER TREATMENT**

**PART II**





# **PATIENT-DERIVED ORGANOIDS AS A PREDICTIVE BIOMARKER FOR TREATMENT RESPONSE IN CANCER PATIENTS**

*npj Precision Oncology*, 2021, 5, 30.

Emerens Wensink  
Sjoerd Elias  
Jasper Mullenders  
Miriam Koopman  
Sylvia Boj  
Onno Kranenburg  
Jeanine Roodhart

## ABSTRACT

Effective predictive biomarkers are needed to enable personalized medicine and increase treatment efficacy and survival for cancer patients, thereby reducing toxic side effects and treatment costs. Patient-derived organoids (PDOs) enable individualized tumour response testing. Since 2018, 17 publications have examined PDOs as a potential predictive biomarker in the treatment of cancer patients. We review and provide a pooled analysis of the results regarding the use of PDOs in individualized tumour response testing, focusing on evidence for analytical validity, clinical validity and clinical utility. We identify future perspectives to accelerate the implementation of PDOs as a predictive biomarker in the treatment of cancer patients.

## INTRODUCTION

Despite advances in the treatment of cancer patients, the burden of cancer deaths remains high, with 9,5 million cancer deaths reported worldwide in 2018<sup>1</sup>. A key limitation in cancer treatment is the lack of valid predictive biomarkers, which reduces the efficacy of treatments<sup>2</sup>. Oncologists are largely unable to predict treatment response for individual patients, resulting in patients receiving ineffective treatment with unnecessary exposure to toxic side effects and high treatment costs. Effective predictive biomarkers are needed to enable personalized medicine and increase survival for cancer patients. Personalized medicine strategies include protein-, RNA- and genome-based stratification, though in oncology, precision medicine has been largely based on genomic biomarkers<sup>3</sup>. However, less than half of patients are eligible for genetically-matched treatment<sup>4,5</sup> and for the majority of anticancer agents no genetic markers are available.

A promising predictive biomarker is individualized tumour response testing using patient-derived organoids (PDOs), in which anticancer agents are screened *ex vivo* on PDOs to predict clinical response (Figure 1). PDOs have been developed for a variety of tumours and are stem-cell derived, three-dimensional self-organizing structures comprised of epithelial cells, mimicking its corresponding tumour<sup>6,7</sup>. PDOs represent a superior preclinical model system compared to previous models through their inherent heterogeneity, long-term stability, applicability for high-throughput screens and enhanced capacity to capture tumour characteristics<sup>8-10</sup>. In 2018 it was first demonstrated that PDOs may predict treatment response in cancer patients<sup>11</sup>.

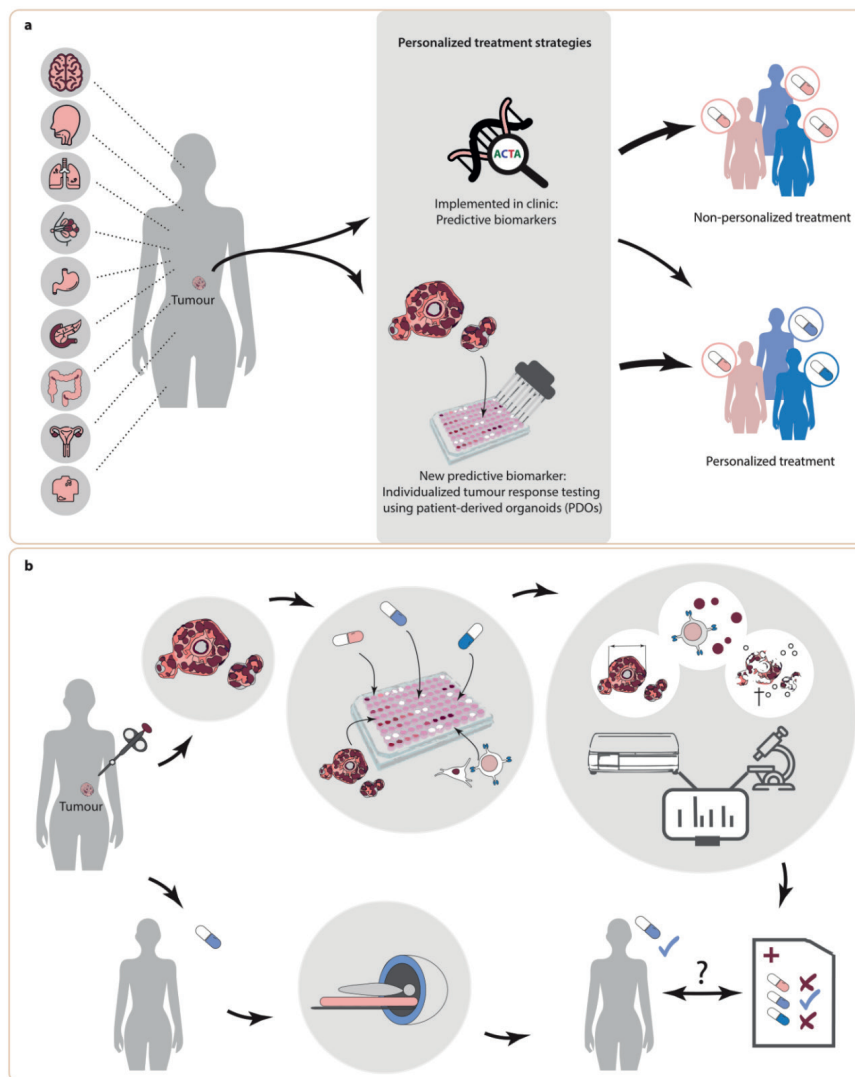
**Figure 1.** PDO-based individualized tumour response testing as a predictive biomarker

Fig 1. a) Illustrates the tumour types for which patient-derived organoids (PDOs) have been tested for clinical validity (listed in full in Table 1). Personalized treatment strategies currently implemented in oncology treatment largely comprise of genomic biomarkers. However, this only results in a personalized treatment strategy for a minority of patients. Individualized tumour response testing using PDOs is a new biomarker which may be used in personalized treatment and increases access to personalized treatment. b) For individualized tumour response testing, tissue from a patient's tumour is obtained to culture organoids, perform drug screens and various read-outs can be obtained to define PDO drug screen response (including organoid size, viability and co-culture cytokine measurements). A predictive biomarker test is developed using the PDO drug screen results and clinical response seen in patients.

**Table 1.** Study characteristics

Publication (author, year, study name)	Study design	Tumour type & stage	Drug screen cohort		Clinical response association cohort		Treatment
			Patients	PDOs	Patients	PDOs	
Ooft, 2019 <sup>21</sup> TUMOROID	Prospective cohort, observational.	mCRC	29	35	29	35	FOLFIRI (>1 <sup>st</sup> line); Irinotecan (>1 <sup>st</sup> line); FOLFOX (mixed lines).
Chalabi, 2020 <sup>22</sup> NICHE	Prospective cohort (within phase II trial), observational.	CRC (Stage III)	11	12	11	12	Nivolumab + ipilimumab (neoadjuvant).
Ganesh, 2019 <sup>29</sup>	Observational cohort.	RC (non-metastatic & metastatic)	14	23	9	17	5-FU & FOLFOX; Radiation.
Yao, 2020 <sup>26</sup> CinClare	Prospective, observational (within phase III trial).	LARC	80	80	80	80	Chemoradiation (capecitabine versus CAPIRI; neoadjuvant).
Narasimhan, 2020 <sup>28</sup> APOLLO	Prospective, offers assay- guided treatment to treatment refractory patients.	mCRC (peritoneal)	15	17	9	9	FOLFOX, FOLFIRI, regorafenib, vandetanib, gemcitabine.
Vlachogiannis, 2018 <sup>11</sup>	Prospective, observational, using PDOs from 4 prospective phase I/II trials.	mCRC, mGC, mGOC	15	19	15	19*	TAS-102, Cetuximab, Regorafenib (CRC); Paclitaxel (GC); 5-FU + cisplatin (GOC).
Steele, 2019 <sup>31</sup>	Observational cohort.	GC (non-metastatic & metastatic)	6	6	2	2	EOX
Tiriac, 2018 <sup>23</sup>	Observational cohort.	Pancreatic cancer (Stage II-IV)	57	66	9	12	5-FU; Gemcitabine + nab-paclitaxel; 5-FU + SN-38 + gemcitabine; 5-FU + SN-38 + oxaliplatin; 5-FU + oxaliplatin; 5-FU + gemcitabine.
Sharick, 2020 <sup>17</sup>	Observational cohort.	Pancreatic cancer (non- metastatic); Breast cancer (not specified)	24	24	10	10	Gemcitabine + 5-FU, oxaliplatin + 5-FU, 5-FU or FOLFIRINOX (pancreatic cancer). AC-T (breast cancer).

Publication (author, year, study name)	Study design	Tumour type & stage	Drug screen cohort		Clinical response association cohort		Treatment
			Patients	PDOs	Patients	PDOs	
Li, 2018 <sup>27</sup>	Observational cohort.	Esophageal cancer (non-metastatic)	8	8	5	5	ECX, ECF, CF.
Driehuis, 2019 <sup>35</sup>	Observational cohort.	HNSCC (non- metastatic)	14	14	7	7	Radiation (postoperative with curative intent, primary and adjuvant).
Sachs, 2018 <sup>14</sup>	Observational cohort.	Breast cancer (metastatic)	NR	12	2	2	Tamoxifen.
Phan, 2019 <sup>18</sup>	Observational cohort.	Ovarian carcinoma (Stage IV)	4	4	2	2	Carboplatin.
De Witte, 2020 <sup>32</sup>	Observational cohort.	Ovarian carcinoma (non-metastatic & metastatic)**	23	36	5	7	Carboplatin + paclitaxel.
Votanopoulos, 2019 <sup>19</sup>	Observational cohort.	Melanoma (Stage III-IV)	7	9	5	7	Pembrolizumab, nivolumab, ipilimumab, dabrafenib/ trametinib.
Mazzocchi, 2018 <sup>20</sup>	Observational cohort.	Mesothelioma (metastatic)	2	2	2	2	Cisplatin + pemetrexed.
Jacob, 2020 <sup>30</sup>	Observational cohort.	Glioblastoma (WHO grade IV)	7	8	5	6	Radiation + temozolomide.

A description of the types of tumours (and stages), cohort size and examined treatments (lines) used for the comparison between *ex vivo* PDO drug screen results and clinical treatment response are reported. The number of patients and PDOs in the study in which drug screens were performed are reported, as well as the number of patients and PDOs for which an association was made with the clinical treatment response. \*The authors report diagnostic results for 21 organoids (2 organoids were tested for >1 treatment line), so that the clinical cohort consists of 19 unique organoids. \*\*The clinical comparison cohort ( $n=5$ ) comprised of high-grade serous ovarian cancer patients who underwent interval debulking surgery.

**Abbreviations:** 5-FU (5-fluorouracil), AC-T (doxorubicin + cyclophosphamide + paclitaxel), CAPIRI (capecitabine + irinotecan), CF (cisplatin + 5-FU), ECF (epirubicin + cisplatin + 5-FU), ECX (epirubicin + capecitabine), EOX (epirubicin + oxaliplatin + 5-FU), FOLFIRI (5-FU + irinotecan), FOLFIRINOX (5-FU + oxaliplatin + irinotecan), FOLFOX (5-FU + oxaliplatin), HNSCC (head and neck squamous cell carcinoma), LARC (locally advanced rectal cancer), mCRC (metastatic colorectal cancer), mGC (metastatic gastric cancer), mGOC (metastatic gastroesophageal cancer), mRC (metastatic rectal cancer), NR (not reported), PDO (patient-derived organoid), RC (rectal cancer), SN-38 (irinotecan), TAS-102 (trifluridine/tipiracil), WHO (World Health Organization).

In order to perform individualized tumour response testing with PDOs, tissue is obtained from a patient's tumour to culture PDOs and perform drug screens (Figure 1). Treatment efficacy is measured by analyzing potential end points which are correlated with treatment sensitivity, including organoid size, viability and co-culture cytokine measurements. Finally, drug screen results and clinical response data are combined to create predictive biomarker tests which are capable of predicting treatment response in patients for a given treatment. Three qualities must be fulfilled for PDOs to function as an effective biomarker: analytic validity, clinical validity and clinical utility<sup>12</sup>. Tests to derive the predictive biomarker should be accurate, reproducible and robust (analytically valid) and results must correlate with clinical end points (clinically valid)<sup>12</sup>. The use of the predictive biomarker should result in improved patient outcome (clinical utility) compared to standard-of-care treatment, in a cost-effective manner.

In this systematic review, we identified 17 oncological studies which report data regarding PDO-based drug screen results and their predictive value or association with the patient's response to treatment in the clinic (Table 1, search strategy is described in Supplementary Table 1). We evaluate the analytical validity by reviewing different drug screen methods used. Next, we investigate the clinical validity by evaluating if clinical studies demonstrate a correlation between PDO-based drug screen results and clinical treatment response in patients and assess if this is impacted by intra-patient heterogeneity. We explore aspects related to the clinical utility and the feasibility of using PDOs in the clinic, including establishment rate and time needed to obtain PDO drug screen results. Lastly, we offer perspectives for future research.

## **ANALYTIC VALIDITY: PDO-BASED DRUG SCREEN EXPERIMENTAL SET-UP**

Prior to performing drug screens, it is essential to perform quality control to verify that the PDOs have been cultured in adequate growth medium to avoid selection bias during establishment and represent the patient's tumour without overgrowth of normal tissue<sup>13</sup>. The 17 studies in this review used varying medium compositions. Depending on the tumour type, growth factor requirements may vary (e.g. neuregulin in breast cancer organoids<sup>14</sup> and  $\beta$ -estradiol in ovarium cancer organoids<sup>15</sup>), however, even within the same tumour type medium compositions and culturing techniques differ. Serum-free growth factors are increasingly becoming available, which allows establishment of organoids in serum-free medium, without undefined differentiation-inducing components<sup>16</sup>. Of note, 3 studies used serum-based media<sup>17–19</sup>, while one study did not specify which medium was used to establish organoids<sup>20</sup>. All studies used at least one of the following quality control assays to verify that PDOs reflect the original tumour: histopathology morphological assessment, DNA and/or RNA sequencing, niche-dependency assays and/or engraftment of organoids in mice (summarized in Supplementary Table 2). In 4 studies, genomic analysis included criteria for which PDOs could be excluded from analysis due to poor quality<sup>14,21–23</sup>. Of note, in several studies quality control was performed on a subset of PDOs, not the whole cohort. A recent protocol for establishing PDOs for drug screening suggests to perform genetic analysis of PDOs and the original tumour tissue to assess that the PDOs are representative and match with the tumour<sup>13</sup>.

The experimental set-up used for PDO- based individualized tumour response testing differed for each study (Supplementary Table 2). PDO drug screens were performed using PDOs embedded in a matrix, in suspension and in a co-culture model. The duration of exposure to drugs varied from 2-24 days. Different end point read-outs were chosen: often cell viability in a luminescence assay (11/17 studies), but also including immunofluorescence with a dead/alive staining, and quantification of interferon-gamma (IFN- $\gamma$ ) in CD8+ T-cells. The TUMOROID trial included a baseline viability measurement in its drug screen set-up, which allowed the determination of growth rate inhibition metrics (GR)<sup>21</sup>, an approach which takes into account the proliferation rate, a known source of variance in drug screens<sup>24</sup>. Sharick *et al.* used optical metabolic imaging (OMI) to measure the metabolic state of single cells within PDOs relative to the average of control cells, which is unique in capturing metabolic heterogeneity during treatment in addition to treatment effect size<sup>17</sup>.

When specified how *in vitro* response was defined, the most frequently used index test was area under the drug response curve (AUC; in 7 studies) rather than other drug response curve (DRC) parameters (listed in Table 2 as the index test). Not all studies provided a definition for *in vitro* response. The parameters which are most informative in predicting patient response may be drug or disease specific<sup>24</sup>. The AUC of a DRC, which combines the potency and efficacy of a drug, is a robust parameter when aiming to compare one agent across multiple tissue lines exposed to the same concentration range and may be more accurate than IC<sub>50</sub> (50% inhibitory concentration)<sup>24,25</sup>. Lastly, for combination treatment two approaches were used to define *in vitro* response: analyzing each agent separately for a combined response classification<sup>19,21,23,26,27</sup> or analyzing the response to combination treatment directly *in vitro*<sup>17,20,21,26–32</sup>. The CinClare trial and de Witte *et al.* report evidence of synergism for combination treatment<sup>32,33</sup>. The TUMOROID trial reported a significant difference in the drug screen results for irinotecan double treatment between PDOs with progressive disease (PD) versus partial response/ stable disease (PR/SD;  $p=0.0260$ ), whereas there was no significant difference in the individual drug parameter drug screen results<sup>21</sup>. Analyzing combination drug screen results, rather than each drug separately, may more accurately discriminate the clinical response in patients.

Table 2. Clinical association with drug screen results

Tumour type [Study]	Treatment	n	PDos	Events	Index test (PDO)	Reference test (clinic)	AUROC [95% C.I.]	Sens/ Spec/ PPV/ NPV	Group values or correlation results
mCRC [Ooft <sup>21</sup> ]	FOLFIRI (>1 <sup>st</sup> line)	12	12	6	GR50 (single treatments)	RECIST response of target lesion <sup>1</sup>	SN-38: 0.83 [0.59-1.08]; 5-FU: 0.80 [0.53 - 1.08]		
					AUC (single treatments)	“	SN-38: 0.75 [0.44-1.06]; 5-FU: 0.80 [0.53 - 1.08]		
					Pan-matrix GR score	“	0.83 [0.53 - 1.13]	<b>Sens: 1.00; Spec: 0.83; PPV: 0.86; NPV: 1.00</b>	
					GR score- based FOLFIRI classifier	“	0.89 [0.67- 1.11 ]	<b>Sens: 1.00; Spec: 0.83; PPV: 0.86; NPV: 1.00</b>	GR score PD 0.73 [0.13 - 1.43]; GR score PR/SD -1.88 [-1.61 - -2.13; <i>p</i> = 0.0260.
					GR score- based FOLFIRI classifier	PFS			Median PFS of 50% most sensitive PDos 169 days; Median PFS of 50% most resistant PDos 58 days; log rank <i>p</i> = 0.0278.
					GR score- based FOLFIRI classifier	RECIST response of target lesion <sup>1</sup> (on irinotecan monotherapy cohort)	0.84 [0.54- 1.14]	<b>Sens: 1.00; Spec: 0.83; PPV: 0.80; NPV: 1.00</b>	Correct classification 83.3% patients ( <i>p</i> = 0.0017).

Tumour type [Study]	Treatment	n	PDOs	Events	Index test (PDO)	Reference test (clinic)	AUROC [95% C.I.]	Sens/ Spec/ PPV/ NPV	Group values or correlation results
mCRC [Ooft <sup>21</sup> ]	Irinotecan (>1 <sup>st</sup> line)	10	10	5	GR50	RECIST response of target lesion <sup>1</sup>	0.96 [0.84 – 1.11]		Median GR50 PD 0.0053 [0.0047 – 0.0058]; Median GR50 SD 0.0023 [0.0019 – 0.0027]; <i>p</i> = 0.0159. Median AUC PD 0.70 [0.67 - 0.72]; Median AUC SD 0.54 [0.52-0.57]; <i>p</i> = 0.0079. Correct classification 80% patients ( <i>p</i> = 0.0061).
					AUC	“	(reference in figure)		
					GR score- based irinotecan classifier	“			
					GR score at 3.2 nM SN-38	“	0.96 [0.84 – 1.11]	<b>Sens:</b> 1.00; <b>Spec:</b> 0.80; <b>PPV:</b> 0.83; <b>NPV:</b> 1.00	Median GR score PD 0.70 [0.63-0.77]; Median GR score SD 0.33 [0.26 - 0.40]; Mann Whitney test <i>p</i> = 0.0159.

Tumour type [Study]	Treatment	n	PDOs	Events	Index test (PDO)	Reference test (clinic)	AUROC [95% C.I.]	Sens/ Spec/ PPV/ NPV	Group values or correlation results
mCRC [Ooft <sup>21</sup> ]	FOLFOX (mixed lines)	10	16	4	GR50 (single treatments)	RECIST response of target lesion <sup>1</sup>	Oxal.: 0.69 [0.40-0.99]; 5-FU: 0.55 [0.17-0.94]	Sens/ Spec/ PPV/ NPV	GR50 5-FU PD 5.17 [3.22-8.31]; GR50 5-FU PR/SD 3.91 [3.00 - 5.36]; <i>p</i> = 0.7105. GR50 oxaliplatin PD 104.2 [89.2-123.2]; GR50 oxaliplatin PR/ SD 71.3 [51.6 - 97.1], <i>p</i> = 0.3301. AUC 5-FU PD 0.81 [0.78-0.83]; AUC 5-FU PR/SD 0.77 [0.74-0.81]; <i>p</i> = 0.9399. AUC oxaliplatin PD 0.96 [0.95-0.96]; AUC oxaliplatin PR/ SD 0.89 [0.84-0.94]; <i>p</i> = 0.3301. GR score PD 1.52 [0.84-2.24]; GR score PR/SD 1.71 [1.38 - 2.04]; not significant.
					AUC (single treatments)	"	Oxal.: 0.69 [0.40-0.99]; 5-FU: 0.56 [0.22-0.89]		
					GR score- based (diagonal DRC combination)	"	0.56 [0.11- 1.01]		
					Pan-matrix GR score	"	0.59 [0.19 - 0.98]	<b>Sens:</b> 0.89 <sup>^</sup> ; <b>Spec:</b> 0.50 <sup>^</sup> ; <b>PPV:</b> 0.80 <sup>^</sup> ; <b>NPV:</b> 0.67 <sup>^</sup>	

Tumour type [Study]	Treatment	n	PDOs	Events	Index test (PDO)	Reference test (clinic)	AUROC [95% C.I.]	Sens/ Spec/ PPV/ NPV	Group values or correlation results
CRC [Chalabi <sup>33</sup> ]	Immune checkpoint inhibitors (neoadjuvant)	11	12	6 <sup>a</sup>	IFN- $\gamma$ (CD-8+ T-cells)	Pathological: MPR( $\leq$ 10% RVT). PR ( $\leq$ 50% RVT) or MPR/PR are considered response		Sens: 0.50 <sup>^</sup> ; Spec: 1.00 <sup>^</sup> ; PPV: 1.00 <sup>^</sup> ; NPV: 0.67 <sup>^</sup>	
RC [Ganesh <sup>29</sup> ]	5-FU & FOLFOX	7	14	7	AUC	PFS			Spearman $r = -0.86$ , $p = 0.024$ for AUC & PFS.
	Radiation	7	11	3	AUC	Endoscopic response to radiation (% tumour decrease)			% AUC matched clinical endoscopic clinical response to radiation.
LARC [Yao <sup>36</sup> ]	Chemoradiation (cap. or CAPIRI; neoadjuvant)	80	80	43	Organoid size decrease cut- off (sensitive if 0-1 (good); $\geq 1/3$ sensitive) TRG	Pathological: TRG	0.88 [0.76 – 0.99]	Sens: 0.92; <sup>2</sup> Spec: 0.79; <sup>2</sup> PPV: 0.79; NPV: 0.92	
	(Radiation only)				"	2-3 (poor)			
	(5-FU only)				"	"		Sens: 0.41; Spec: 0.98; PPV: 0.94; NPV: 0.66	
	(Irinotecan only)				"	"		Sens: 0.59; Spec: 0.88; PPV: 0.81; NPV: 0.72	
mCRC [Narasimhan <sup>28</sup> ]	FOLFOX	5	5	3	AUC	PD vs. PR/SD		Sens: 0.78; Spec: 0.79; PPV: 0.78; NPV: 0.79	AUC in PR/SD vs. PD, no statistically significant difference.

Tumour type [Study]	Treatment	n	PDOs	Events	Index test (PDO)	Reference test (clinic)	AUROC [95% C.I.]	Sens/ Spec/ PPV/ NPV	Group values or correlation results
mGIC [Vlachogiannis 11]	Combined	15	15'	14"	Not specified	Not specified		Sens: 1.00^; Spec: 0.93^; PPV: 0.88^; NPV: 1.00^	
Gastric cancer [Steele <sup>11</sup> ]	EOX	2	2	1	% cell viability decrease (flow cytometry)	Pathological: grade (% residual tumour)		Sens: 1.00; Spec: 1.00; PPV: 1.00; NPV: 1.00	
Pancreatic cancer [Tiriak <sup>23</sup> ]	Combined	9	12	3	AUC (tertiles of sensitive/ intermediate/ resistant; separately/ drug)	PFS (relative to expected median)		Sens: 1.00; Spec: 0.67; PPV: 0.86; NPV: 1.00	Patients with PFS > median (n=6): 5/6 PDOs had >1 sensitive drug and no resistant drugs <i>in vitro</i> .
Pancreatic cancer [Sharick <sup>17</sup> ]	Combined (adjuvant)	7	7	3	OMI index (metabolic state)	RFS (>12 months)		Sens: 1.00; Spec: 1.00; PPV: 1.00; NPV: 1.00	Median OMI index in RFS >12 months 2.2 [1.2 - 3.2]; Median OMI index in RFS <12 months 0.4 [0.2 - 0.5]. Decreased heterogeneity of treatments/controls in RFS >12 months vs RFS <12 months.

Tumour type [Study]	Treatment	n	PDOs	Events	Index test (PDO)	Reference test (clinic)	AUROC [95% C.I.]	Sens/ Spec/ PPV/ NPV	Group values or correlation results
Esophageal cancer [Li <sup>27</sup> ]	Combined	5	5	5	AUC (monotherapy and combination)	Pathological: (resistant: TRG 4-5)		Sens: -; Spec: 0.60; PPV: 0.00; NPV: 1.00	
	ECX	3	3	3	"	"		Sens: -; Spec: 0.67; PPV: 0.00; NPV: 1.00	
	ECF	1	1	1	"	"		Sens: -; Spec: 1.00; PPV: -; NPV: 1.00	
	CF	1	1	1	"	"		Sens: -; Spec: 0.00; PPV: 0.00; NPV: -	
HNSCC [Driehuis <sup>35</sup> ]	Radiotherapy (curative intent)	7	7	3	AUC	Relapse after radiation		Sens: 0.75; Spec: 1.00; PPV: 1.00; NPV: 0.75	
Ovarian cancer [Phan <sup>18</sup> ]	Carboplatin	2	2	2	Decrease in viability	Not specified		Sens: -; Spec: 1.00; PPV: -; NPV: 1.00	Descriptive results (n=2).

Tumour type [Study]	Treatment	n	PDOs	Events	Index test (PDO)	Reference test (clinic)	AUROC [95% C.I.]	Sens/ Spec/ PPV/ NPV	Group values or correlation results
Ovarian cancer [De Witte, 2020 <sup>32]</sup>	Carboplatin + paclitaxel	5	7	4 <sup>#</sup>	AUC	1) Pathological [CRS=1 vs. CRS=2]			AUC higher in CRS = 1 vs. CRS =2; Wilcoxon signed- rank test $p<0.001$ .
					“	2) CA-125 normalization ( $<35$ kU/L during 1 <sup>o</sup> treatment)			AUC higher in non-CA-125 normalization vs. normalization; $p =$ 0.0004.
					“	3) RECIST response (SD vs. PR)			AUC higher in RECIST SD versus PR; $p = 0.0092$ .
					“	4) 6-month PFS			AUC did not correlate with 6-month PFS; $p=0.9663$ .
					“	5) OS			AUC did correlate with OS ( $<17$ -months); Wilcoxon signed- rank test corrected for multiple testing $p=0.0016$ .
Melanoma [Votanopoulos <sup>19]</sup>	Immune checkpoint inhibitors	5	7	5	Decrease in viability	Not specified			Sens: -; Spec: 0.60; PPV: 0.00; NPV: 1.00
	Dabrefinib/ trametinib	1	2	1 <sup>*</sup>	“	“			Sens: 1.00 <sup>^</sup> ; Spec: 1.00 <sup>^</sup> ; PPV: 1.00 <sup>^</sup> ; NPV: 1.00 <sup>^</sup>

Tumour type [Study]	Treatment	n	PDOs	Events	Index test (PDO)	Reference test (clinic)	AUROC [95% C.I.]	Sens/ Spec/ PPV/ NPV	Group values or correlation results
Mesothelioma [Mazzocchi <sup>20</sup> ]	Cisplatin + pemetrexed	2	2	1	Decrease in viability	Not specified		Sens: 1.00; Spec: 1.00; PPV: 1.00; NPV: 1.00	
Glioblastoma [Jacob <sup>30</sup> ]	Radiation + temozolomide	5**	6	3	Decrease in viability	Not specified		Sens: 1.00; Spec: 1.00; PPV: 1.00; NPV: 1.00	

Results concerning clinical validity results for PDO drug screens and clinical response are displayed, along with the publication, tumour type & stage, parameters compared, treatment, number of patients, PDOs and events per treatment. Where applicable, results were digitized from graphs using WebPlotDigitizer software. <sup>^</sup>Sensitivity/specificity/PPV/NPV are calculated using patient-level values, except for: <sup>21</sup> outcomes are reported for 2 PDOs per patient for 3 patients with PDOs before & after treatment; <sup>33</sup> includes 2 PDOs for 1 patient with synchronous dMMR and pMMR tumours; <sup>11</sup> results for PDO-level (not available for patient-level). \*The authors report diagnostic results for 21 organoids (2 organoids were tested for >1 treatment line), so that the clinical cohort consists of 19 unique organoids, of which regorafenib *in vitro* tests were performed using xenograft models (not a PDO drug screen). \*\* Drug screen performed in 7 patients, but clinical response reported for 5 patients. <sup>1</sup>indicates response was measured after 3 cycles of treatment; <sup>2</sup>the reported diagnostic values differ from the article due to alternative designation of true positive and true negative values, <sup>3</sup>Only reported for PDOs, not patient-level<sup>11</sup>; patient had 2 PDOs with mixed clinical response<sup>9,32</sup> or patient had synchronous tumour (responders & non-responder) with 2 PDOs<sup>33</sup>.

**Abbreviations:** 1° (primary treatment), 5-FU (5-fluorouracil), AUC (area under the curve), AUROC (area under the receiver operator curve), C.I. (confidence interval), capec. (capecitabine), CAPIRI (capecitabine + irinotecan), CAPOX (capecitabine + oxaliplatin), CF (cisplatin + 5-FU), CR (complete response), CRS (chemotherapy response score), dMMR (deficient mismatch repair), DRC (drug response curve), ECF (epirubicin + cisplatin + 5-FU), ECX (epirubicin + cisplatin + capecitabine), EOX (epirubicin + oxaliplatin + 5-FU), FOLFIRI (5-FU + irinotecan), FOLFOX (5-FU + oxaliplatin), GR (growth rate inhibition metrics), GR<sub>50</sub> (concentration at which GR is 50%), HNSCC (head and neck squamous cell carcinoma), IFN- $\gamma$  (interferon gamma), LARC (locally advanced rectal cancer), mCRC (metastatic colorectal cancer), mGIC (metastatic gastrointestinal cancer), MPR (major pathological response), NPV (negative predictive value), OMI (optical metabolic imaging), OS (overall survival), Oxal. (oxaliplatin), PD (progressive disease), PDO (patient-derived organoid), PFS (progression-free survival), pMMR (proficient mismatch repair), PPV (positive predictive value), PR (partial response), RC (rectal cancer), RECIST (Response evaluation criteria in solid tumours), RFS (recurrence-free survival), RVT (residual viable tumour), SD (stable disease), sens. (sensitivity), SN-38 (irinotecan), spec. (specificity), TRG (tumour regression grade score).

## CLINICAL VALIDITY: CORRELATION OF PDO DRUG SCREEN SENSITIVITY WITH CLINICAL RESPONSE

The 17 studies in this review assessed the clinical validity of PDOs as a predictive biomarker for treatment response in the clinic. The studies were heterogeneous, varying in study design, patient population and treatments (Table 1). All studies were observational, with the exception of the APOLLO trial which was the first study to offer patients assay-guided treatment<sup>28</sup>. The results encompassed a variety of tumour types and stages of disease. Colorectal cancer (CRC) studies were the most frequent among the publications (5/17) and also the largest in patient cohort size<sup>21,22,28,29,33</sup>. Many studies (7/17) derived PDOs from patients with metastatic disease<sup>11,14,18,20,21,28,30</sup>. Lastly, the treatments examined included systemic chemotherapy, targeted therapy, (chemo) radiation and immunotherapy.

In general, the patient cohorts for which *ex vivo* drug response results and clinical response are available were small, varying from 2-80 patients per study, with a median of 7 patients per study and a median of 3 patients per type of treatment per study (Table 1 and Table 2). An exception is the Phase 3 CinClare trial, which examined PDO drug response in 80 locally advanced rectal cancer (LARC) patients receiving neoadjuvant chemoradiation, randomized for capecitabine versus capecitabine with irinotecan (CAPIRI)<sup>33</sup>.

The results regarding the correlation of PDO-based drug screen results and clinical response per study are described per tumour type and treatment type below (Table 2 and Supplementary Table 3). We summarized the clinical validity results for all studies into an evidence landscape figure (Figure 2)<sup>34</sup>. Five of the 17 studies reported a statistically significant correlation and/or predictive value for PDO-based drug screen results and clinical response for a given treatment<sup>11,21,29,32,33</sup>. A trend for a correlation or predictive value was seen in 11 studies for a given treatment<sup>14,17-20,23,27,29-31,35</sup>, whereas 3 studies reported no correlation<sup>21,22,28</sup> and 1 study was unable to test for an association<sup>28</sup>. To compare PDO-based drug screen results and clinical response, certain studies chose a clinical parameter which reflects the lesion from which the PDO was obtained rather than the patient's clinical response, while the latter is most clinically relevant (Table 2 listed as the reference test). In the following sections, we analyze the results in more detail and report pooled results of the clinical validity results.

**Figure 2.** Evidence landscape of PDO drug screen parameters and clinical response

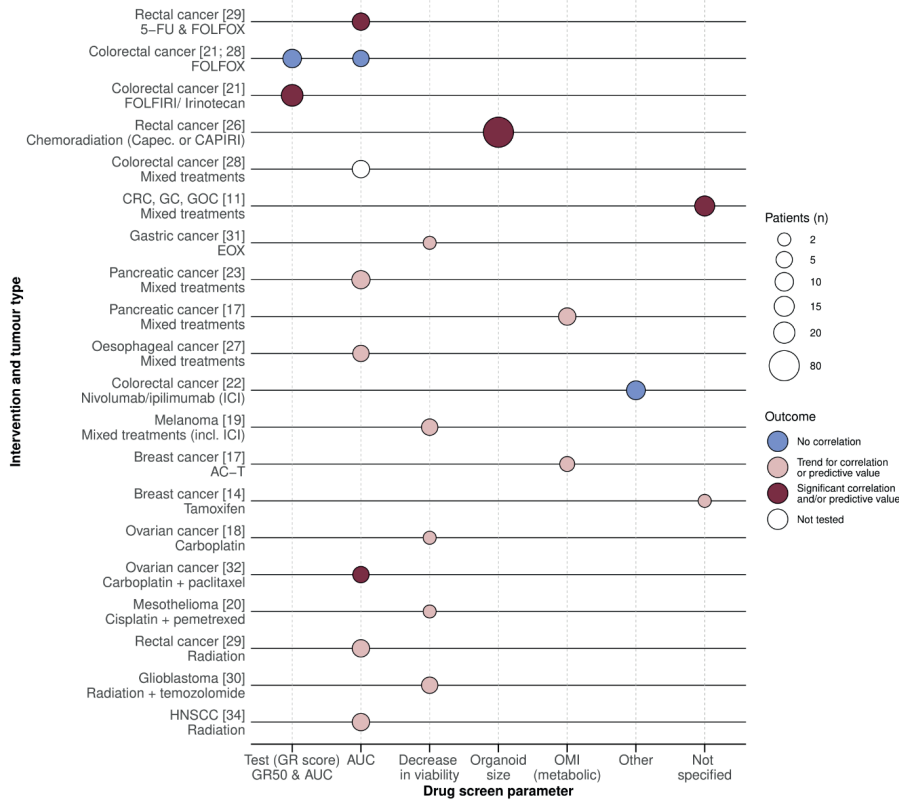


Fig 2. Illustrates the clinical validity results for PDOs as a predictive biomarker for treatment response (dark red: significant correlation and/or predictive value found, pink: trend for correlation or predictive value found, blue: no correlation and white: not tested), with the size of the circle representing the patient cohort size, specified per treatment and tumour type (y-axis) and *ex vivo* drug response parameter (x-axis).

**Abbreviations:** 5-FU (5-fluorouracil), AC-T (doxorubicin + cyclophosphamide + paclitaxel), AUC (area under the curve), Capec. (capecitabine), CAPIRI (capecitabine + irinotecan), CRC (colorectal cancer), EOX (epirubicin + oxaliplatin + 5-FU), FOLFIRI (5-FU + irinotecan), FOLFOX (5-FU + oxaliplatin), GC (gastric cancer), GOC (gastroesophageal cancer), GR (growth rate inhibition metrics), GR50 (value with 50% viable GR), HNSCC (head and neck squamous cell carcinoma), ICI (immune checkpoint inhibitors), OMI (optical metabolic imaging), PDO (patient-derived organoid).

## Systemic chemotherapy and targeted therapy

### *CRC patients*

Four studies reported results regarding the predictive value of PDO drug screen results for treatment response in CRC patients of various disease stages receiving systemic chemotherapy<sup>21,28,29,33</sup>. The TUMOROID and CinClare trials showed that PDO drug screen results were associated with the observed clinical response in patients treated with irinotecan-based regimens<sup>21,33</sup>. In the TUMOROID study, the examined *ex vivo* drug screen parameters derived from growth rate inhibition metrics were predictive for the best RECIST response to irinotecan-based treatment observed in the lesion from which the PDO was obtained in metastatic CRC (mCRC) patients ( $n=10$  irinotecan and  $n=12$  irinotecan-doublet)<sup>21</sup>. A drug response-based cut-off correctly discriminated between the best RECIST response observed in the lesion from which the PDO was derived in 92% (95% C.I. 65-99%; calculated using Wilson's method, 11/12) of patients receiving irinotecan-doublet treatment<sup>21</sup>. Moreover, 50% of patients with the most sensitive *in vitro* results had significantly longer progression-free survival (PFS, median 169 versus 58 days (digitized from figure),  $p=0.0278$ )<sup>21</sup>. In the CinClare trial, *ex vivo* PDO drug screen results (organoid size) were predictive for clinical response (tumour regression grade upon resection) in 80 locally advanced rectal cancer (LARC) patients receiving neoadjuvant chemoradiation, randomized for capecitabine or capecitabine and irinotecan (CAPIRI), and correctly classified 95% (95% C.I. 76-91%; 68/80) of the patients<sup>33</sup>.

These results are promising and suggest that PDO drug screen results are predictive for clinical response in CRC patients receiving irinotecan-based treatment. However, there are conflicting results regarding if PDO drug screens are associated with clinical response for oxaliplatin-based treatment. Ganesh *et al.*<sup>29</sup> reported an association in a RC cohort while the TUMOROID trial<sup>21</sup> and APOLLO trial<sup>28</sup> did not find an association in mCRC patients. The drug response results (AUC) was significantly associated with the observed PFS in 7 RC patients (Ganesh *et al.*<sup>29</sup>). This is in contrast to the TUMOROID results in mCRC patients, where none of the *ex vivo* drug screen parameters were predictive of the RECIST response in the lesion from which the PDO was derived<sup>21</sup>. The APOLLO results support the findings from the TUMOROID study, since the drug response results (AUC) were not different for 9 mCRC patients with peritoneal metastases who clinically had response versus no response to 5-fluorouracil and oxaliplatin (FOLFOX) treatment<sup>28</sup>. Interestingly, both Ganesh *et al.* and the TUMOROID trial reported results for a cohort of patients in which the majority had

received oxaliplatin-based treatment prior to deriving the organoids (71% and 60%, respectively)<sup>21,29</sup>. With the reported positive result in the Ganesh *et al.* study, despite including patients with prior oxaliplatin treatment, it seems that prior treatment does not affect the predictive value of *ex vivo* PDO drug screens. One aspect which may have influenced the conflicting results between the Ganesh *et al.* study and the TUMOROID and APOLLO trials is the choice of solvent used to dissolve oxaliplatin for the PDO drug screens, since this is known to affect the activity of oxaliplatin<sup>36</sup>.

### ***Upper-gastrointestinal cancer***

In a cohort of 15 metastatic gastrointestinal cancer patients treated with different systemic therapies, a binary *ex vivo* sensitive versus resistant classification of the drug screen results had predictive value, with a 100% sensitivity, 93% specificity, 88% positive predictive value and 100% negative predictive value<sup>11</sup>. Four studies provided descriptive results to compare PDO drug screen results with clinical response in patients with pancreatic cancer<sup>17,23</sup>, non-metastatic gastric cancer (GC)<sup>31</sup> and non-metastatic esophageal cancer<sup>27</sup> patients. The results from the Tiriack *et al.* and Sharick *et al.* studies in pancreatic cancer patients receiving chemotherapy regimens suggest that PDOs can function as a predictive biomarker for pancreatic cancer patients. However, the results in gastric and esophageal cancer patients are preliminary, based on small cohort sizes, and do not yet convincingly show that PDO drug screen results are associated with clinical response in patients. The PDO drug screen results (AUC) of 8/9 (89%, 95% C.I. 57-98%) pancreatic cancer patients within the Tiriack *et al.* study, were consistent with clinical outcome (Table 2 with 1-3 patients received a given treatment)<sup>23</sup>. Similarly, in the Sharick *et al.* study, a cut-off of the optical metabolic imaging index (OMI) correctly discriminated 7/7 patients based on the clinical recurrence-free survival (RFS)<sup>17</sup>. Patients with a RFS >12 months had a higher OMI index compared to patients with <12 months RFS<sup>17</sup>. The drug screen of several pancreatic PDOs also contained cancer associated fibroblasts, for which the *in vitro* response generally matched the clinical response<sup>17</sup>. Steele *et al.* reported that in two GC patients receiving epirubicin, oxaliplatin and 5-FU, the clinical response matched the drug response (% cell viability decrease)<sup>31</sup>. Lastly, Li *et al.* found mixed results in 5 non-metastatic esophageal cancer patients receiving one of 3 combination treatment regimens, with 3/5 having a matching PDO drug screen result with the observed clinical response, while 2/5 patients with a tumour regression grade of 4-5 had limited sensitivity in the PDO drug screens<sup>27</sup>.

### **Other tumour types**

Two publications examined the clinical validity of PDO drug screens in patients with ovarian cancer<sup>18,32</sup>. PDO drug response (AUC) was correlated with clinical response in 5 patients with high-grade serous ovarian cancer who underwent interval debulking surgery for three clinical parameters: histopathological chemotherapy response score, normalization of CA-125 and RECIST response ( $p<0.01$ )<sup>32</sup>. No correlation was found with 6-months PFS ( $p=0.97$ ), however, PDO drug response did correlate with overall survival ( $<17$  months,  $p=0.0016$ )<sup>32</sup>. The remaining studies reported descriptive results for 2-3 patients each, shown in Table 2 and Supplementary Table 3, with metastatic ovarian cancer patients receiving carboplatin treatment<sup>18</sup>, metastatic mesothelioma patients receiving cisplatin and pemetrexed<sup>20</sup>, metastatic breast cancer patients receiving tamoxifen<sup>14</sup> and breast cancer patients receiving neoadjuvant chemotherapy<sup>17</sup>. Considering the small sample size and descriptive nature of the results, we will not elaborate on the results.

### **(Chemo)radiation**

In all four studies examining (chemo)radiation, a possible association is reported between PDO drug screens results and clinical response for RC<sup>29,33</sup>, HNSCC<sup>35</sup> and glioblastoma<sup>30</sup> patients. The results for the CinClare study were described above, which combined neoadjuvant radiation with capecitabine or CAPIRI<sup>33</sup>, showing a clear association between PDO drug screen results (organoid size) and tumour regression grade upon resection in 80 LARC patients. Ganesh *et al.* reported that the PDO drug screen AUC was associated with endoscopic clinical response in 7 RC patients, based on a descriptive comparison (Table 2)(22). Similarly, in non-metastatic HNSCC patients, the PDO drug screen AUC descriptively matched clinical response in 86% (95% C.I. 49-97%; 6/7) of patients receiving postoperative radiotherapy and was inconsistent in 1 patient<sup>35</sup>. A descriptive comparison of the PDO drug screens for radiation and temozolomide treatment (decrease in cell viability) matched the clinical response in 5 glioblastoma patients, while for 2 patients no clinical response was reported<sup>30</sup>. Although the results should be validated in larger trials, the results for the predictive value of PDO drug screens in predicting radiation treatment response are promising.

### **Immune checkpoint inhibition**

Two studies used co-cultures of PDOs with immune cells to examine the effectivity of immune checkpoint inhibitors (ICI), which require the immune system to orchestrate cytotoxicity<sup>19,22</sup>. The studies showcase the potential of using PDOs in more complex

tumour microenvironment co-culture models to predict a variety of treatments. Votanopoulos *et al.* reported that in 86% (95% C.I. 49-97%; 6/7) of melanoma patients, immune-enhanced PDO drug screen results (decrease in cell viability) recapitulated the clinical response<sup>19</sup>. In the NICHE trial, the PDO-immune cell co-culture drug screen results (based on the IFN- $\gamma$  production by CD-8+ T-cells) matched the clinical response to neoadjuvant nivolumab and ipilimumab in 6 CRC patients, all non-responders with proficient mismatch repair (pMMR) tumours<sup>22</sup>. However, the drug screen results were inconsistent in 3/6 patients with clinical response, comprising patients with a pMMR or deficient mismatch repair (dMMR) tumour<sup>22</sup>. Thus, in melanoma patients, immune-enhanced PDO cell viability was associated with treatment response to ICI, indicating that PDO-immune co-cultures drug screen results correlated with treatment response, while in CRC patients receiving neoadjuvant immunotherapy, the IFN- $\gamma$  production by CD-8+ T-cells in co-culture with PDOs did not predict clinical response.

## Pooled clinical validity results

To summarize the clinical validity results, we pooled the sensitivity and specificity of PDO-based drug screen results for predicting treatment response. The pooled sensitivity and specificity values for clinical response through PDO-based screening were 0.81 (95% C.I. 0.69-0.89) and 0.74 (95% C.I. 0.64-0.82), respectively (Figure 3 demonstrates the paired forest plots), with a  $\chi^2$  test for heterogeneity of 11.6 for sensitivity ( $p=0.56$ ) and 6.4 for specificity ( $p=0.93$ ). Considering the small sample sizes, we repeated the meta-analysis for studies with  $\geq 5$  responder and non-responder patients<sup>21,22,26</sup> and obtained similar results (pooled sensitivity 0.84 (95% C.I. 0.56-0.95) and specificity 0.81 (95% C.I. 0.68-0.89)), with a  $\chi^2$  test of 8.8 for sensitivity ( $p<0.05$ ) and 0.9 for specificity ( $p=0.83$ ). The pooled sensitivity and specificity values are likely an overestimation, since studies that did not report quantitative results necessary for the meta-analysis could not be included and since not all studies used a pre-defined index test. As such, we cannot exclude publication or outcome reporting bias in the results. The area under the receiver operator curve (AUROC) for discriminating clinical response using various index tests clinical validity results is summarized in a forest plot in Supplementary Figure 1. The pooled results are an indication of the overall performance of PDOs in predicting response across different tumour types and treatments for the available evidence. Despite heterogeneity in tumour type, treatment and end points used in the studies, we do not see heterogeneity in our data. However the pooled results should be interpreted cautiously considering that PDOs may predict differently between tumour types and treatments. Future studies will enable meta-analysis per tumour and treatment type; due to limited evidence this is currently not possible.

Figure 3. Forest plots of sensitivity and specificity (clinical validity pooled results)

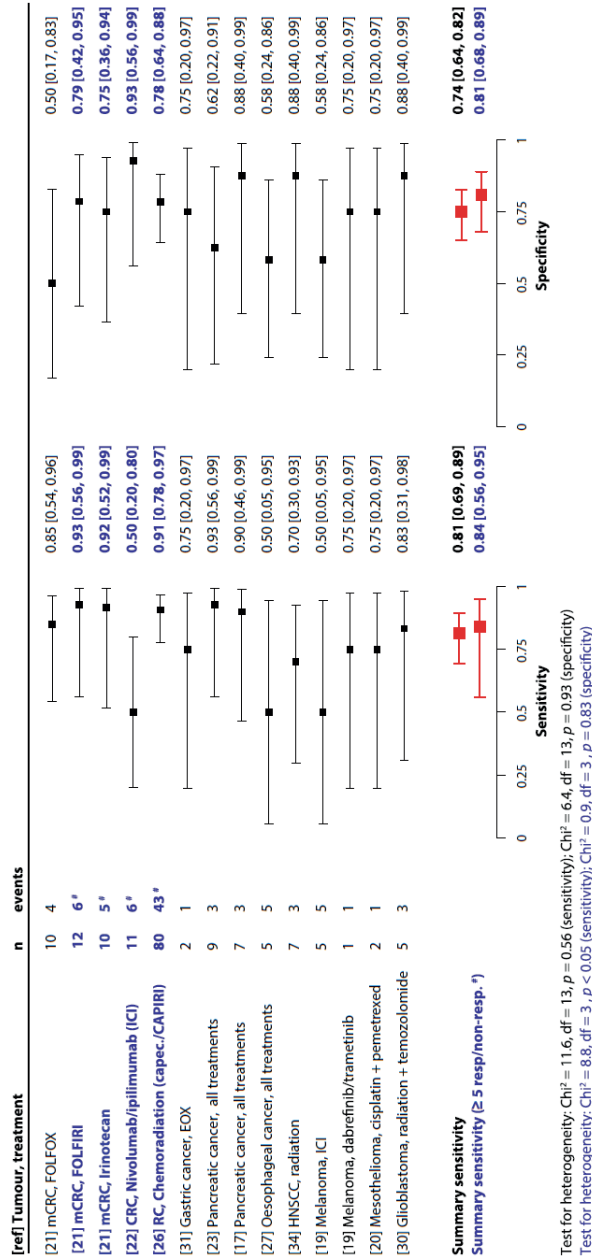


Fig 3. A paired forest plot of the sensitivity and specificity of each study and treatment type is shown with 95% confidence intervals. A bivariate meta-analysis was performed to obtain a pooled summary estimate for sensitivity and specificity indicated in the forest plots (1: for all studies that reported results that could be included in this pooled analysis and 2: for studies with  $\geq 5$  responders/non-responders). The analysis was performed in R (Version 3.6.1) using the “mada” package<sup>52</sup>. In blue and bold font (\*) the studies were indicated which were included in the analysis for  $\geq 5$  responders/non-responders. Patients who contributed to multiple accuracy estimates: 2 patients received ICI and dabrafenib/trametinib<sup>19</sup>; 3 patients received FOLFOX and irinotecan-based treatment<sup>21</sup>; 3 patients had 2 PDOs each (before and after FOLFOX treatment)<sup>21</sup> and 1 patient had a synchronous tumour (responder & non-responder)<sup>22</sup>.  
*Abbreviations:* capec. (capecitabine), CAPIRI (capecitabine + irinotecan), EOX (epirubicin + oxaliplatin + 5-FU), FOLFIRI (5-fluorouracil + irinotecan), FOLFOX (5-fluorouracil + oxaliplatin), HNSCC (head & neck squamous cell carcinoma), ICI (immune checkpoint inhibitors), mCRC (metastatic colorectal cancer), RC (rectal cancer), ref (reference), resp. (responder clinically), non-resp. (non-responder clinically).

## Effect of spatial intrapatient heterogeneity on clinical validity

For PDO-based drug screen results to be clinically valid in patients with advanced cancer, PDOs should be able to act as a predictive biomarker for treatment response in the patient as a whole, without being limited by intratumoural heterogeneity<sup>37</sup>. Although PDOs are heterogenous<sup>9,17</sup>, a PDO is derived from a single biopsy or surgical resection, representing a snapshot of one spatial lesion. Intra-patient heterogeneous PDO drug screen responses were relatively uncommon in a cohort of CRC patients, with pharmacological profiles of PDOs obtained from multiple CRC liver metastases in 10 patients largely clustering together patient-wise and inter-metastatic heterogeneity being observed in  $<1/10^{\text{th}}$  of all drug-patient comparisons<sup>38</sup>. Interestingly, Sharick *et al.* were able to assess the heterogeneity on a cellular level of PDOs during treatment using OMI<sup>17</sup>. The PDOs of pancreatic cancer patients with a RFS  $>12$  months had a lower degree of metabolic heterogeneity during treatment (versus control PDOs), while PDOs from patients with a RFS  $<12$  months had increased heterogeneity in treatment PDOs compared to control PDOs<sup>17</sup>.

Five studies reported spatial intrapatient heterogeneity in drug responses in PDOs derived from distinct cancer lesions per patient, although the clinical implications are unclear<sup>11,21,23,28,32</sup>. Vlachogiannis *et al.* demonstrated that the mixed clinical response seen in a mCRC patient to trifluridine/tipiracil (TAS-102) treatment, with progressive disease (PD) in one liver metastasis and stable disease (SD) in a second liver metastasis, was reflected in the PDO drug screen results, with an eightfold difference in  $GI_{50}$  between the PDO derived from the sensitive metastasis and the PDOs derived from the PD metastasis<sup>11</sup>. In 7 ovarian cancer patients (2-4 PDOs/patient), De Witte *et al.* demonstrated that all related PDOs exhibited a differential drug response to at least one drug, and that the differential drug response could only be partially linked to genetic heterogeneity<sup>32</sup>. Based on a small subset of patients with paired PDOs from different tumour regions, the remaining studies demonstrated heterogeneity in drug response, without correlating this to the response seen in the patient. These studies show that the effect is treatment<sup>21,23</sup> and patient specific<sup>28</sup>.

The studies which reported a correlation or predictive value for PDO individualized tumour response testing and treatment response are based on results using the response in the lesion from which the PDO was obtained<sup>21,33</sup> and analyzing the patient's response as a whole, e.g. RECIST response<sup>11,29,32</sup>. The TUMOROID study, which primarily examined the clinical response in the lesion from which the PDO was obtained, also showed that

patients with the 50% most sensitive drug screen results had a significantly longer PFS for 5-FU + irinotecan combination treatment<sup>21</sup>. Although intrapatient heterogeneity in PDO drug screen results has been shown, the current available results indicate that PDOs are able to act as a predictive biomarker for the patient's treatment response as a whole, and thus are not significantly impeded by intrapatient heterogeneity. However, these results should be confirmed in future studies.

## **CLINICAL UTILITY: FEASIBILITY OF USING PDO DRUG SCREENS TO PREDICT TREATMENT RESPONSE**

If PDOs are to be effectively translated to the clinic for precision medicine, their clinical utility must be proven. Three feasibility aspects are important for the clinical utility of PDO-based individualized tumour response testing: 1) having a sufficiently high PDO establishment rate to balance the burden incurred through diagnostic interventions to obtain tissue to culture PDOs; 2) avoiding unnecessary treatment delay by minimizing the time from obtaining tissue for culturing PDOs to analysis of PDO drug screen results and 3) the use of PDO-guided treatment should be beneficial for patients through increased survival and/or quality of life by either using PDOs to select patients for the most optimal standard-of-care treatment option or through identifying novel treatment candidates. The clinical benefit of PDO-guided treatment compared to (standard-of-care) physician-guided treatment has not been assessed. However, the APOLLO trial is the first to provide PDO-guided treatment to patients. We discuss all three feasibility aspects below.

The organoid establishment rate was reported in 12 studies, ranging from 31-90%. We performed a random effects pooled analysis of the reported organoid establishment rates per study using a generalized linear mixed model. Two analyses were performed: sample-level (the proportion of established organoids per total number of samples obtained) and patient-level (the proportion of patients with established organoids per total number of patients sampled), including all reported establishment rates (except when only an approximation was reported<sup>35</sup>). The pooled organoid establishment rate was 68.5% (95% C.I. 56.5-78.5%;  $I^2 = 89\%$ ) in 7 studies reporting sample-level organoid establishment rates<sup>11,14,21,23,26,27,29</sup> and 68.0% (95% C.I. 54.9-78.8%;  $I^2 = 83\%$ ) in 8 studies reporting patient-level organoid establishment rates<sup>17,19,21,23,27-30</sup> (Supplementary Figure 2 demonstrates the forest plots). An establishment rate of approximately 70% may be high enough to balance potential burdens for patient in obtaining tissue for culturing PDOs and may improve

through developments in establishment techniques. The highest establishment rates were observed in melanoma (resections: 90%<sup>19</sup>) and rectal cancer patients (77%-86%<sup>29,33</sup>). Only tumour biopsy cellularity was found to be associated to PDO establishment rate<sup>11</sup>, while site of tissue sampling (primary tumour versus metastasis) and prior treatment were not found to be different in patients with or without established PDOs<sup>17,21</sup>.

The acceptable time from tissue sampling to obtaining drug screen results, without delaying treatment will vary depending on the clinical situation. Unfortunately, only two studies reported the time needed from tissue sampling to obtaining drug screen results (<8 weeks for all patients<sup>28</sup> and <20 days in a pilot for 1 patient<sup>32</sup>). The time required to establish PDOs after obtaining tissue can vary greatly, as exemplified in a study for pancreatic cancer resections (median 10 days, inter-quartile range, IQR 6-12) versus breast cancer biopsies (median 34 days, IQR 27-51)<sup>17</sup>. The period from obtaining tissue to drug screen results may be reduced by minimizing the number of PDOs needed and duration of drug exposure while maintaining analytical validity.

The APOLLO trial shows promise that PDO-guided treatment is feasible and may offer additional treatment options for treatment refractory patients<sup>28</sup>. Patients with peritoneal mCRC and disease progression despite standard systemic treatment were screened using an adapted CRC-focused panel of clinically available treatments and two patients received PDO-guided off-label treatment (described in Supplementary Table 3). Although this study illustrates the feasibility of performing organoid-based treatment stratification, considering the small number of patients no firm conclusion can be drawn concerning the clinical utility of PDO-guided treatment. However, the study does highlight the potential for PDO drug screens to identify novel treatment candidates for patients which otherwise would not have been available. PDO drug screens can be performed in a high throughput manner, enabling rapid screening of large libraries of therapeutic agents to identify new agents or new combinations of agents for a patient or subgroup of patients<sup>39</sup>. Effective anti-cancer treatments are often combination regimens and thus libraries of single targeted agents may not result in identification of clinically effective agents. In conclusion, demonstrating the clinical utility of PDOs requires demonstrating that patients benefit through increased survival or quality of life, which can potentially be achieved by either using PDOs to identify which standard-of-care drugs are most effective (avoiding exposure to ineffective drugs and their associated toxicity) or through identifying new therapy candidates through library screening of non-standard-of-care drugs.

## RECOMMENDATIONS FOR FUTURE STUDIES

Several recommendations for future research can help accelerate the implementation of PDO drug screens as a predictive biomarker in the clinic. We will first address recommendations for methodology and reporting of studies. Secondly, new innovations in PDO drug screens are examined which can improve the reproducibility and automation of drug screens. And lastly, we will describe aims of future studies to accelerate the transition of PDOs to the clinic.

### Standardized methodology and reporting

Researchers should aim to adhere to methodological standards when reporting results, to facilitate study quality assessment (including potential biases) and study result interpretation<sup>40</sup>. The REporting recommendations for tumour MARKer prognostic studies (REMARK) guidelines can be used to standardize reporting for studies examining PDOs as a predictive biomarker in patients with cancer<sup>40</sup>. We wish to highlight several related aspects applicable to the methodology and reporting for future studies, which are specific for PDO predictive biomarker studies.

Given the heterogeneity in PDO drug screen set-up used, a validated, standardized experimental design may offer benefits. It allows researchers to avoid unnecessary time in validating a new experimental design, to use previously tested organoid lines to validate results and to prospectively validate published PDO-based diagnostic tests. This may be achievable, since the APOLLO trial demonstrated that using a similar experimental design in two laboratories resulted in significantly correlated results (Pearson's  $r=0.96$ ,  $p<0.05$ )<sup>28</sup>. Furthermore, physicians could use a database of published drug screen results to assess if a patient's drug screen is relatively resistant or sensitive.

To start with, PDO culturing and screens should aim to use materials which are not animal-derived or serum-based. The majority of organoid studies used animal-derived extracellular matrices (e.g. Matrigel®), which are biologically variable and contain animal-derived growth factors<sup>41</sup>. Animal-derived matrices can theoretically reduce the reproducibility of drug screens and influence PDOs in culture, while also reducing the extent to which the model reflects the physiological setting. New synthetic hydrogels, which are fully defined and growth factor-free, have proven to support the establishment of human PDOs from single cell suspensions and are amenable to drug screens<sup>42–45</sup>. Future studies should explore if synthetic hydrogels can be used to establish

PDOs from human tissue and whether the use of synthetic hydrogels improves the reproducibility of drug screens. Similarly, serum-free Wnt growth factor supplements are increasingly available, enabling organoid culture medium to become serum-free<sup>16</sup>.

Subsequently, one important aspect is transparent reporting of the chosen definition for clinical response and *in vitro* response, to ensure reproducible study results and interpretation for clinical applicability. In the methods, it should be clear if the chosen end points were part of a pre-defined statistical analysis. As well, drug screens of combination treatments should analyze response of the combination treatment by adding both agents directly *in vitro*, rather than analyzing the response separately, to best model the treatment given to the patient.

Finally, as mentioned previously, detailed reporting of the establishment rate, including how successful establishment of PDOs was defined and quality control to verify that PDOs represent the original tumour, and time needed to obtain PDO drug screen results from obtaining tissue will help in validating the feasibility of using PDOs as a biomarker. Reporting results of feasibility aspects for PDO establishment, including the location and type of tissue obtained, establishment rate per patient and per sample obtained, and features (e.g. patient demographics, molecular status etc.) found to be associated with successful establishment of PDOs, will aid researchers in improving PDO establishment techniques. If PDOs derived from primary disease (or earlier treatment lines) have predictive value for patients with metastatic disease or later treatment lines, which is currently unknown, the time to obtaining results can be minimized by culturing PDOs early in the course of the disease.

## Innovations in PDO drug screens

Organoid-based drug screens are developing rapidly, offering new techniques and materials which can improve the reproducibility, high throughput design and automation of PDO screening. A newly developed automated microfluidic platform for PDO drug screening enables the addition of drugs at different time points, allowing drug screens to more closely resemble combination treatment regimens given to patients<sup>46</sup>. Such automated platforms may be compatible with image-based analysis<sup>47</sup>, which in contrast to single end points such as cell viability, allows researchers to assess multiple end points, better resembling the full drug response in PDOs. Furthermore, PDO drug screens are being optimized to become more high throughput<sup>48</sup>. These

developments will aid in automating PDO drug screens, decreasing the amount of PDOs needed and developing read-outs which more accurately represent the true drug response in PDOs compared to traditional read-outs.

## **Aims for future studies**

The clinical validity of using PDOs to predict treatment response should be confirmed in studies with a larger group of patients, ideally in a specific clinical setting for one tumour and treatment type. The desired predictive test qualities may vary for a given clinical setting, e.g. the amount of treatment options still available and the *a priori* chance that a patient will respond to a given treatment. The predictive value of PDOs may be tumour or treatment specific, given the conflicting results regarding oxaliplatin-containing treatment within different mCRC studies and ICI treatment for melanoma and CRC. Having results available for specific subgroups will give us further knowledge concerning the settings in which PDOs may offer predictive value for patients.

Subsequently, the use of more complex PDO models, such as co-cultures may be necessary to accurately predict treatment response for certain treatments where the tumour microenvironment – including the immune system – affects treatment sensitivity (e.g. immunotherapy)<sup>19,22,49</sup>. The available evidence suggests that co-cultures may not be necessary to predict treatment response for chemotherapeutics, since the discussed studies on mono-culture PDO models could predict chemotherapy response. However, more complex drug screen models may increase the predictive ability for treatments, with increasing anti-cancer agents targeting the tumour microenvironment as well as the tumour itself and the possibility to include effects of drug metabolism.

Ultimately, the clinical value of using PDO individualized tumour response testing should be proven by comparing clinical outcomes, such as progression-free survival or response rates, in randomized clinical trials comparing physician guided standard-of-care treatment versus assay-guided treatment derived from PDO drug screens<sup>50</sup>. Patients receiving assay-guided treatment should benefit clinically and ideally this benefit should be cost effective compared to standard-of-care treatment or, for example, genomic-guided treatment<sup>50,51</sup>.

## CONCLUSIONS

The currently available results offer an optimistic perspective that individualized tumour response testing using PDOs have clinical validity as a predictive biomarker for cancer patients. The pooled sensitivity and specificity for discriminating patients with a clinical response through PDO-based screening were 0.81 (95% C.I. 0.69-0.89) and 0.74 (95% C.I. 0.64-0.82), respectively, although this is an estimation since not all studies reported results which could be used for the pooled analysis and not all studies used a pre-defined index test when analyzing results. The pooled results are an indication of the overall performance of PDOs in predicting response across different tumour types and treatments for the available evidence. However they should be interpreted cautiously considering that PDOs may predict differently between tumour types and treatments. The current evidence is strongest for CRC patients, with larger studies showing a correlation between PDO-based drug screen results and systemic therapy/radiation treatment response and with smaller studies showing promising descriptive results for other tumour and treatment types. Associations were found for a broad variety of tumours, treatment types and encompassing several drug screen parameters, albeit, not consistently for all tumour types and treatments. Prior to being able to implement PDO-based drug screens in the clinic, the results should be validated in similar, larger patient cohorts. The current challenge is to prove that PDO-based individualized tumour response testing is feasible, by optimizing organoid establishment rates and time to obtaining PDO-based screening results. The results regarding clinical validity of PDOs as a predictive biomarker are promising and ultimately the clinical utility should be proven by demonstrating that PDO-based individualized tumour response testing is cost effective and offers clinical benefit for patients. If PDOs can be established for the majority of patients within a feasible time frame, this potential predictive biomarker can facilitate personalized medicine for a group of patients for whom there is a great need for valid predictive biomarkers.

## REFERENCES

1. Ferlay J., Colombet M., Soerjomataram I., Dyba T., Randi G., Bettio M., et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. *Eur J Cancer*. (2018);**103**:356–87.
2. Sawyers C.L. The cancer biomarker problem. *Nature*. (2008);**452**(7187):548–52.
3. Letai A. Functional precision cancer medicine-moving beyond pure genomics. *Nat Med*. (2017);**23**(9):1028–35.
4. Meric-Bernstam F., Brusco L., Shaw K., Horombe C., Kopetz S., Davies M.A., et al. Feasibility of large-scale genomic testing to facilitate enrollment onto genomically matched clinical trials. *JCO*. (2015);**33**(25):2753–62.
5. van der Velden D.L., Hoes L.R., van der Wijngaart H., van Berge Henegouwen J.M., van Werkhoven E., Roepman P., et al. The Drug Rediscovery protocol facilitates the expanded use of existing anticancer drugs. *Nature*. (2019);**574**(7776):127–31.
6. De Souza N. Organoids. *Nat Methods*. (2018);**15**(1):23.
7. Li M., Izpisua Belmonte J.C. Organoids - Preclinical models of human disease. *NEJM*. (2019);**380**(6):569–79.
8. Sato T., Stange D.E., Ferrante M., Vries R.G.J., Van Es J.H., Van Den Brink S., et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology*. (2011);**141**(5):1762–72.
9. Van De Wetering M., Francies H.E., Francis J.M., Bounova G., Iorio F., Pronk A., et al. Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell*. (2015);**161**(4):933–45.
10. Pampaloni F., Reynaud E.G., Stelzer E.H.K. The third dimension bridges the gap between cell culture and live tissue. *Nat Rev Mol Cell Biol*. (2007);**8**(10):839–45.
11. Vlachogiannis G., Hedayat S., Vatsiou A., Jamin Y., Fernández-mateos J., Khan K., et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science*. (2018);**359**(6378):920–6.
12. Simon R. Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology. *Per Med*. (2010);**7**(1):33–47.
13. Driehuis E., Kretzschmar K., Clevers H. Establishment of patient-derived cancer organoids for drug-screening applications. *Nat Protoc*. (2020);**15**:3380–3409.
14. Sachs N., de Ligt J., Kopper O., Gogola E., Bounova G., Weeber F., et al. A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. *Cell*. (2018);**172**(1–2):373–386. e10.
15. Kopper O., de Witte C.J., Löhmußaar K., Valle-Inclán J.E., Hami N., Kester L., et al. An organoid platform for ovarian cancer captures intra- and interpatient heterogeneity. *Nat Med*. (2019);**25**(5):838–49.
16. Miao Y., Ha A., de Lau W., Yuki K., Santos A.J.M., You C., et al. Next-Generation Surrogate Wnts Support Organoid Growth and Deconvolute Frizzled Pleiotropy In Vivo. *Cell Stem Cell*. (2020);**27**(5):840–851.e6.
17. Sharick J.T., Walsh C.M., Sprackling C.M., Pasch C.A., Pham D.L., Esbona K., et al. Metabolic Heterogeneity in Patient Tumor-Derived Organoids by Primary Site and Drug Treatment. *Front Oncol*. (2020);**10**(May):1–17.

18. Phan N., Hong J.J., Tofig B., Mapua M., Elashoff D., Moatamed N.A., et al. A simple high-throughput approach identifies actionable drug sensitivities in patient-derived tumor organoids. *Commun Biol.* (2019);2(1):1–11.
19. Votanopoulos K.I., Forsythe S., Sivakumar H., Mazzocchi A., Aleman J., Miller L., et al. Model of Patient-Specific Immune-Enhanced Organoids for Immunotherapy Screening: Feasibility Study. *Ann Surg Oncol.* (2019);27(6):1956–67.
20. Mazzocchi A.R., Rajan S.A.P., Votanopoulos K.I., Hall A.R., Skardal A. In vitro patient-derived 3D mesothelioma tumor organoids facilitate patient-centric therapeutic screening. *Sci Rep.* (2018);8(1):1–12.
21. Ooft S.N., Weeber F., Dijkstra K.K., McLean C.M., Kaing S., van Werkhoven E., et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Sci Transl Med.* (2019);11(513):eaay2574.
22. Chalabi M., Fanchi L.F., Dijkstra K.K., Van den Berg J.G., Aalbers A.G., Sikorska K., et al. Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat Med.* (2020);26(April).
23. Tiriac H., Belleau P., Engle D.D., Plenker D., Deschênes A., Somerville T.D.D., et al. Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. *Cancer Discov.* (2018);8(9):1112–29.
24. Fallahi-Sichani M., Honarnejad S., Heiser L.M., Gray J.W., Sorger P.K. Metrics other than potency reveal systematic variation in responses to cancer drugs. *Nat Chem Biol.* (2013);9(11):708–14.
25. Huang S., Pang L. Comparing statistical methods for quantifying drug sensitivity based on in vitro dose-response assays. *Assay Drug Dev Technol.* (2012);10(1):88–96.
26. Yao Y., Xu X., Yang L., Zhu J., Wan J., Shen L., et al. Patient-Derived Organoids Predict Chemoradiation Responses of Locally Advanced Rectal Cancer. *Cell Stem Cell.* (2020);26(1):17–26.
27. Li X., Francies H.E., Secrier M., Perner J., Miremadi A., Galeano-Dalmau N., et al. Organoid cultures recapitulate esophageal adenocarcinoma heterogeneity providing a model for clonality studies and precision therapeutics. *Nat Commun.* (2018);9(1):2983.
28. Narasimhan V., Wright J.A., Churchill M., Wang T., Lannagan T.R.M., Vrbanc L., et al. Medium-throughput drug screening of patient-derived organoids from colorectal peritoneal metastases to direct personalized therapy. *CCR.* (2020);26:3662–70.
29. Ganesh K., Wu C., O'Rourke K.P., Szeglin B.C., Zheng Y., Sauvé C.E.G., et al. A rectal cancer organoid platform to study individual responses to chemoradiation. *Nat Med.* (2019);25(10):1607–14.
30. Jacob F., Salinas R.D., Zhang D.Y., Nguyen P.T.T., Schnoll J.G., Wong S.Z.H., et al. A Patient-Derived Glioblastoma Organoid Model and Biobank Recapitulates Inter- and Intra-tumoral Heterogeneity. *Cell.* (2020);180(1):188–204.e22.
31. Steele N.G., Chakrabarti J., Wang J., Biesiada J., Holokai L., Chang J., et al. An Organoid-Based Preclinical Model of Human Gastric Cancer. *Cmgh.* (2019);7(1):161–84.
32. de Witte C.J., Espejo Valle-Inclan J., Hami N., Löhmussaar K., Kopper O., Vreuls C.P.H., et al. Patient-Derived Ovarian Cancer Organoids Mimic Clinical Response and Exhibit Heterogeneous Inter- and Inpatient Drug Responses. *Cell Rep.* (2020);31(11):107762.

33. Yao Y., Xu X., Yang L., Zhu J., Wan J., Shen L., et al. Patient-Derived Organoids Predict Chemoradiation Responses of Locally Advanced Rectal Cancer. *Cell Stem Cell*. (2020);**26**(1):17–26.e6.
34. Hey S.P., Gerlach C. V., Dunlap G., Prasad V., Kesselheim A.S. The evidence landscape in precision medicine. *Sci Transl Med*. (2020);**12**(540):1–5.
35. Driehuis E., Kolders S., Spelier S., Löhmußaar K., Willems S.M., Devriese L.A., et al. Oral Mucosal Organoids as a Potential Platform for Personalized Cancer Therapy. *Cancer Discov*. (2019);**9**(7):852–71.
36. Varbanov H.P., Ortiz D., Höfer D., Menin L., Galanski M., Keppler B.K., et al. Oxaliplatin reacts with DMSO only in the presence of water. *Dalt Trans*. (2017);**46**(28):8929–32.
37. Hinohara K., Polyak K. Intratumoral Heterogeneity: More Than Just Mutations. *Trends Cell Biol*. (2019);**29**(7):569–79.
38. Bruun J., Kryeziu K., Eide P.W., Moosavi S.H., Eilertsen I.A. Patient-derived organoids from multiple colorectal cancer liver metastases reveal moderate intra-patient pharmacotranscriptomic heterogeneity Translational relevance. *CCR*. (2020);**26**:4107–19.
39. Verissimo C.S., Overmeer M., Ponsioen B., Drost J., Verissimo C.S., Mertens S., et al. Targeting mutant RAS in patient-derived colorectal cancer organoids by combinatorial drug screening. *Elife*. (2016);**15**(5):e18489.
40. Altman D.G., McShane L.M., Sauerbrei W., Taube S.E. Reporting recommendations for tumor marker prognostic studies (REMARK): Explanation and elaboration. *BMC Med*. (2012);**10**(51):1–39.
41. Slobodan Vukicevic, Hynda K. Kleinman, Frank P. Luyten, Anita B. Roberts, Nanette S. Roche A.H.R. Identification of multiple active growth factors in basement membrane matrigel suggests caution in interpretation of cellular activity related to extracellular matrix components. *Exp Cell Res*. (1992);**202**(1):1–8.
42. Hernandez-Gordillo V., Kassis T., Lampejo A., Choi G.H., Gamboa M.E., Gnecco J.S., et al. Fully synthetic matrices for in vitro culture of primary human intestinal enteroids and endometrial organoids. *Biomaterials*. (2020);**254**(May):120125.
43. Ng S., Tan W.J., Pek M.M.X., Tan M.H., Kurisawa M. Mechanically and chemically defined hydrogel matrices for patient-derived colorectal tumor organoid culture. *Biomaterials*. (2019);**219**(July):119400.
44. Gjorevski N., Sachs N., Manfrin A., Giger S., Bragina M.E., Ordóñez-Morán P., et al. Designer matrices for intestinal stem cell and organoid culture. *Nature*. (2016);**539**(7630):560–4.
45. Sorrentino G., Rezakhani S., Yildiz E., Nuciforo S., Heim M.H., Lutolf M.P., et al. Mechano-modulatory synthetic niches for liver organoid derivation. *Nat Commun*. (2020);**11**(1):1–10.
46. Schuster B., Junkin M., Kashaf S.S., Romero-Calvo I., Kirby K., Matthews J., et al. Automated microfluidic platform for dynamic and combinatorial drug screening of tumor organoids. *Nat Commun*. (2020);**11**(1):1–12.
47. Renner H., Grabos M., Becker K.J., Kagermeier T.E., Wu J., Otto M., et al. A fully automated high-throughput workflow for 3D-based chemical screening in human midbrain organoids. *Elife*. (2020);**9**:1–39.
48. Du Y., Li X., Niu Q., Mo X., Qui M., Ma T., et al. Development of a miniaturized 3D organoid culture platform for ultra-high-throughput screening. *J Mol Cell Biol*. (2020);**12**(8):630–43.

49. Son B., Lee S., Youn H., Kim E., Kim W., Youn B. The role of tumor microenvironment in therapeutic resistance. *Oncotarget*. (2017);**8**(3):3933–45.
50. Wilhelm-Benartzi C.S., Mt-Isa S., Fiorentino F., Brown R., Ashby D. Challenges and methodology in the incorporation of biomarkers in cancer clinical trials. *Crit Rev Oncol Hematol*. (2017);**110**:49–61.
51. Polley M.Y.C., Freidlin B., Korn E.L., Conley B.A., Abrams J.S., McShane L.M. Statistical and practical considerations for clinical evaluation of predictive biomarkers. *J Natl Cancer Inst*. (2013);**105**(22):1677–83.
52. Team R.C. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2018. p. <https://www.R-project.org/>.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Literature search strategy

Search terms & databases used	<p><b>Pubmed:</b> (Organoid*[tiab] OR Tumoroid*[tiab] OR PDO[tiab] OR “patient-derived organoid”[tiab]) AND (predict*[tiab] OR correlate*[tiab] OR personalized[tiab] OR precision[tiab]) AND (cancer[tiab] OR neoplasm[tiab] OR tumor[tiab]) NOT review[tiab] No filters.</p> <p><b>Embase:</b> (organoid*:ab,ti OR tumoroid*:ab,ti OR pdo:ab,ti OR ‘patient-derived organoid’:ab,ti) AND (predict*:ab,ti OR correlate*:ab,ti OR personalized:ab,ti OR precision:ab,ti) AND (cancer:ab,ti OR neoplasm:ab,ti OR tumor:ab,ti) NOT review:ab,ti Filters: source from article, article in press, data papers, letter or short survey.</p>
Inclusion criteria	<p>Original research articles, reporting results on: <b>Population:</b> patients with cancer, with clinical response data available for a given treatment. <b>Intervention:</b> organoids established using tissue obtained from the tumour (without xenograft models in the establishment process), on which drug screens were performed <b>Control:</b> clinical response of the patient to a given treatment <b>Outcome:</b> compare the <i>in vitro</i> organoid drug screen response with the clinical response of the patient to a given treatment &gt;1 patient results comparing the PDO-based drug screen result and clinical outcome must be reported.</p>
Results search	<p>Pubmed: 385 results Embase: 286 results <b>Total: 671 results</b></p> <p>Remove duplicates: -268 results <b>Total: 403 results</b> Screen title/abstract: -374 results <b>Total: 29 results</b> Screen full-text article: -13 results (all excluded due to not reporting comparison with patient-level response for a given treatment) <b>Identified studies for systematic review: 16</b> + 1 results (from cross-referencing / citations of included studies) <b>Included studies in systematic review: 17.</b></p>

Abbreviations: (tiab or ab,ti) title abstract.

**Supplementary Table 2.** Drug screen experimental designs & authentication of tumour organoids

Study (Author, year)	Quality control of tumour organoids	Drug screen model	Timeline drug screen	Drug exposure duration (days)	End point of drug screen	Blinding of drug screen and/or clinical parameters
Ooft, 2019 (21)	DNA sequencing (discard if identity scores <0.9).	Embedded in Geltrex (1:2 ratio of medium:matrix).	Day 1: 4-day old PDOs; baseline read-out; drug exposure. Day 6: end-point.	6	Cell viability (CellTiter-Glo 3D). Calculate growth rate inhibition metrics (GR).	Drug screens were performed by 2 researchers (2 <sup>nd</sup> was blinded for clinical parameters). No other blinding reported.
Chalabi, 2020 (22)	DNA sequencing (discard if identity scores <0.9).	PDO lymphocyte co-culture (patient- derived TILs) in suspension.	Day 0: 1 day prior stimulate PDOs with IFN- $\gamma$ . Day 1: single cell PDOs + PBLs (1:20 ratio), co- stimulate with IL-2 + anti- CD-28 (plate-bound); drug exposure. Every 2-3 days: refresh medium (IL-2 + drugs). Day 14: end-point.	14	Quantification of IFN- $\gamma$ production by CD8+ T-cells.	No blinding performed.
Ganesh, 2019 (29)	Histopathology confirmation of morphology. DNA sequencing.	Embedded in Matrigel.	<u>Drug screen:</u> Day 1: 2-3 day old PDOs; drug exposure. Day 3: refresh medium with drugs. Day 6: end-point. <u>Radiation:</u> Day 1: irradiate PDOs. Day 8-13: end-point.	Chemotherapy: 6 Radiation: 8-13	Cell viability (CellTiter-Glo).	Investigators performing radiation on PDOs were blinded for patient characteristics. No other blinding reported.

Study (Author, year)	Quality control of tumour organoids	Drug screen model	Timeline drug screen	Drug exposure duration (days)	End point of drug screen	Blinding of drug screen and/or clinical parameters
Yao, 2020 (26)	Histopathology confirmation of morphology. DNA sequencing.	Embedded in Matrigel.	<i>Drug screen:</i> <u>Day 1:</u> PDOs (100 uM diameter) drug exposure. <u>Day 3:</u> refresh medium with drugs. Every subsequent 3 days: refresh medium. <u>Day 24:</u> end-point. <i>Irradiation:</i> <u>Day 1:</u> PDOs (100 uM diameter) irradiated. Every 3 days: refresh medium. <u>Day 24:</u> end-point.	24	PDO size (day 24 versus 0, imaging).	No blinding reported.
Narasimhan, 2020 (28)	Histopathology confirmation of morphology. DNA sequencing. Xenograft to confirm tumorigenicity.	Suspension (5% Matrigel).	<u>Day 1:</u> PDOs; drug exposure. <u>Day 6:</u> end-point.	6	Cell viability (CellTiter-Glo 2D).	No blinding reported.
Vlachogiannis, 2018 (11)	Histopathology confirmation of morphology. DNA sequencing.	Embedded in Matrigel.	<u>Day 1:</u> 3-day old PDOs; drug exposure. Every 2 days: refresh medium with drugs. <u>Day 9-11</u> (growth rate- dependent); end-point.	9-11	Cell viability (CellTiter-Glo).	Researchers were blinded to the patients' response.

Study (Author, year)	Quality control of tumour organoids	Drug screen model	Timeline drug screen	Drug exposure duration (days)	End point of drug screen	Blinding of drug screen and/or clinical parameters
Steele, 2019 (31)	Niche-dependency. RNA sequencing. Xenograft to confirm tumorigenicity.	Unclear.	<i>Monotherapy:</i> <u>Day 1:</u> drug exposure. <u>Day 2</u> (48 hrs): end-point. <i>Combination (used for correlation):</i> <u>Day 1:</u> drug exposure (use IC <sub>50</sub> per PDO). <u>Every 24 hrs:</u> flow cytometry of single cells.	<i>Monotherapy:</i> 2 <i>Combination:</i> 4	<i>Monotherapy:</i> MTS assay. <i>Combination:</i> cell viability via flow cytometry.	No blinding reported.
Tiriac, 2018 (23)	DNA sequencing (discard if does not have any known pathogenic mutations).	Suspension (10% Matrigel)	<u>Day 1:</u> 1-day old PDOs; drug exposure. <u>Day 5:</u> end-point.	5	Cell viability (CellTiter-Glo).	No blinding reported.
Sharick, 2020 (17)	Xenograft to confirm tumorigenicity.	Embedded in Matrigel.	<u>Day 1:</u> drug exposure. <u>Day 2:</u> remove gemcitabine, paclitaxel, SN-38 and oxaliplatin (not other drugs). <u>Day 3:</u> end-point.	3 (some cases 5 or 7)	Optical metabolic imaging (OMI) index, which reflects the metabolic state of cells (relative to average control values).	No blinding reported.
Li, 2018 (27)	Niche dependency. Histopathology confirmation of morphology. DNA & RNA sequencing.	Suspension (25% BME-2).	<u>Day 1:</u> plate PDO suspension. <u>Day 2:</u> drug exposure. <u>Day 7:</u> end-point.	6	Cell viability (CellTiter-Glo).	No blinding reported.

Study (Author, year)	Quality control of tumour organoids	Drug screen model	Timeline drug screen	Drug exposure duration (days)	End point of drug screen	Blinding of drug screen and/or clinical parameters
Driehuis, 2019 (34)	Niche-dependency. Histopathology confirmation of morphology. DNA & RNA sequencing. Xenograft to confirm tumorigenicity.	<i>Drug screen:</i> Suspension (5% BME). <i>Radiation:</i> Embedded in BME.	<i>Drug screen:</i> <u>Day 1:</u> 2-day old PDOs; drug exposure. <u>Day 5:</u> end-point. <i>Radiation:</i> <u>Day 1:</u> 2-day old PDOs; irradiated in water bath, medium change. <u>Day 4:</u> end-point.	Drug screen: 5 Radiation: 4	Cell viability (CellTiter-Glo 3D).	No blinding reported.
Sachs, 2018 (14)	Histopathology confirmation of morphology. DNA (exclude from analysis if cross- contamination seen) & RNA sequencing.	Suspension (5% BME), bottom layer BME.	<u>Day 1:</u> 5-7 day old PDOs; drug exposure. <u>Day 5:</u> end-point.	5	Cell viability (CellTiter-Glo 3D).	No blinding reported.
Phan, 2019 (18)	Histopathology confirmation of morphology. Xenograft to confirm tumorigenicity.	Embedded in Matrigel, plated as mini-ring.	<u>Day 1:</u> 2-day old PDOs; drug exposure. <u>Day 2,3:</u> refresh medium with drugs. <u>Day 4:</u> end-point.	3	PDO size and count (brightfield imaging). Cell viability (CellTiter-Glo 3D).	No blinding reported.
De Witte, 2020 (32)	Histopathology confirmation of morphology. DNA sequencing.	Suspension (2% BME), bottom layer BME.	<u>Day 1:</u> PDOs; drug exposure. <u>Day 5:</u> end-point.	5	Cell viability (CellTiter-Glo 2D).	Pathological assessment was blinded. No other blinding reported.

Study (Author, year)	Quality control of tumour organoids	Drug screen model	Timeline drug screen	Drug exposure duration (days)	End point of drug screen	Blinding of drug screen and/or clinical parameters
Votanopoulos, 2019 (19)	Histopathology confirmation of morphology.	Embedded in ECM-mimicking HA/collagen-based hydrogel. Immune-enhanced PDOs OR PDOs + immune cells (1:1 ratio).	Day 1: 7-day old PDOs (from establishment); drug exposure. Day 3 (72h): end-point.	3	Cell viability (CellTiter-Glo 3D). Live/Dead staining (calcein-AM and ethidium) and confocal imaging.	Partial blinding: Clinical information was not shared with the laboratory, with the exception of type of tumour and type of prior treatments.
Mazzocchi, 2018 (20)	Histopathology confirmation of morphology.	Embedded in HA + gelatin hydrogel in tumour-on-chip microfluidic device.	Day 1 (growth dependent): drug exposure. Day 7 or 14: end-point.	7 or 14	Live/Dead staining (calcein-AM and ethidium) via immunofluorescence.	No blinding reported.
Jacob, 2020 (30)	Histopathology confirmation of morphology. DNA & RNA sequencing. Xenograft to confirm tumorigenicity.	Suspension.	Day 1: irradiation + temozomide exposure. Day 3.5: refresh medium. Day 7: end-point.	7	Cell viability staining (Ki67 + DAPI) via immunofluorescence.	No blinding reported.

*Abbreviations:* BME (basement membrane extract), CD8+ T-cells (cytotoxic T-cells), CD-28 (cluster of differentiation 28), ECM (extracellular matrix), DNA (deoxyribonucleic acid), Gy (gray), HA (hyaluronic acid), IL-2 (interleukin-2), IFN- $\gamma$  (interferon-gamma), kV (kilovolt), mA (milliampere), MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulphophenyl)-2H-tetrazolium)), PBLs (peripheral blood lymphocytes), PDOs (patient-derived organoids), RNA (ribonucleic acid), TILs (tumour infiltrating lymphocytes).

Supplementary Table 3. Clinical association with drug screen - descriptive results

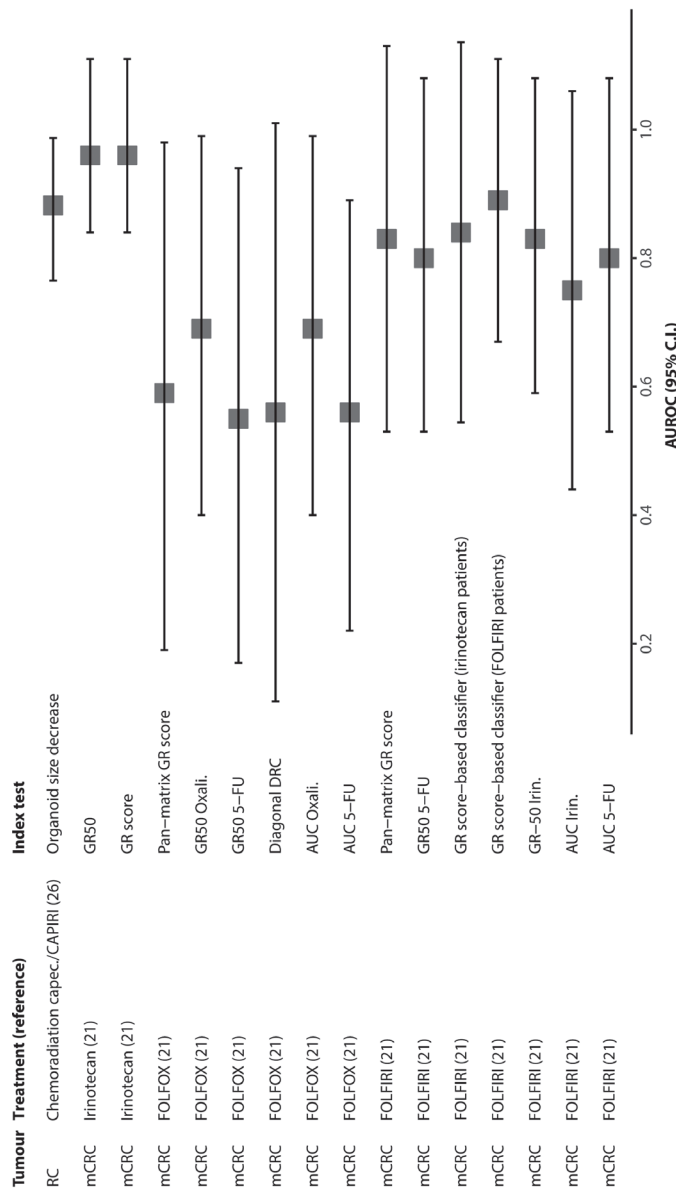
Tumour type [Publication]	Treatment	# patients/ treatment	# events (clinical non- responders)	PDO drug screens & clinical response parameters associated		Descriptive results
				Index test (PDO)	Reference test (Clinic)	
mCRC [Narasimhan (28)]	FOLFIRI	4	4	AUC.	PD versus PR/SD/CR.	Not possible to compare since all patients had PD.
	Regorafenib	1	1	AUC.	PD versus PR/SD/CR.	AUC resistant & patient failed to respond.
	Vandetanib (PDO- guided treatment)	1	1	AUC.	PD versus PR/SD/CR.	AUC with response, but no clinical response (discordant).
	Gemcitabine (PDO-guided treatment)	1	0	AUC.	PD versus PR/SD/CR.	AUC sensitive & patient had PR to gemcitabine- capecitabine (3 month evaluation).
	TAS-102 (mCRC)	4	2	Not specified.	Not specified.	1 patient with mixed response had 8-fold difference in GI <sub>50</sub> between SD versus PD metastasis-derived PDO. In 3 other patients (2 with SD, 1 resistant), PDO response matched clinical response.
mGIC [Vlachogiannis (11)]	Cetuximab (mCRC)	4	3	Not specified.	Not specified.	In 2 primary resistant patients & 1 PD patient, PDOs showed no response. In 1 patient with slow growing metastasis, PDO had marginal response.
	Paclitaxel (mGC)	3	3	Not specified.	Not specified.	In patient with 2 PDOs (prior and after PD), the PDO obtained prior to PD had GI50 which was ¼ of GI50 of PDO obtained after PD and which matched GI50 of 2 other patients with PD.
	5-FU + cisplatin (mGOC)	2	1	Not specified.	Not specified.	In a chemosensitive versus refractory patient, GI <sub>50</sub> differed by 10-fold.

Tumour type [Publication]	Treatment	# patients/ treatment	# events (clinical non- responders)	PDO drug screens & clinical response parameters associated		Descriptive results
				Patients	Index test Reference (PDO) test (Clinic)	
Breast cancer [Sharick (17)]	AC-T (breast cancer; neoadjuvant).	3	1	3	OMI index	The patient with clinical RCB III (extensive residual disease) had the lowest OMI index (Glass $\Delta$ 1.97) compared to the patients with clinical RCB I/II (minimal/ moderate disease) with OMI index of 3.76 and 6.23, respectively. The OMI index decreased upon AC-T exposure in the PDOs (Glass's $\Delta$ <1.9, $p$ <0.0001), while the OMI index heterogeneity altered to varying degrees.
Breast cancer [Sachs (14)]	Tamoxsifen.	2	1	2	Not specified.	Location of DRC (relative to others) matched clinical response.

Results concerning the association or predictive value between PDO drug screens and clinical response are displayed for studies reporting only descriptive results, along with the publication, tumour type & stage, parameters compared, treatment and number of patients per treatment.

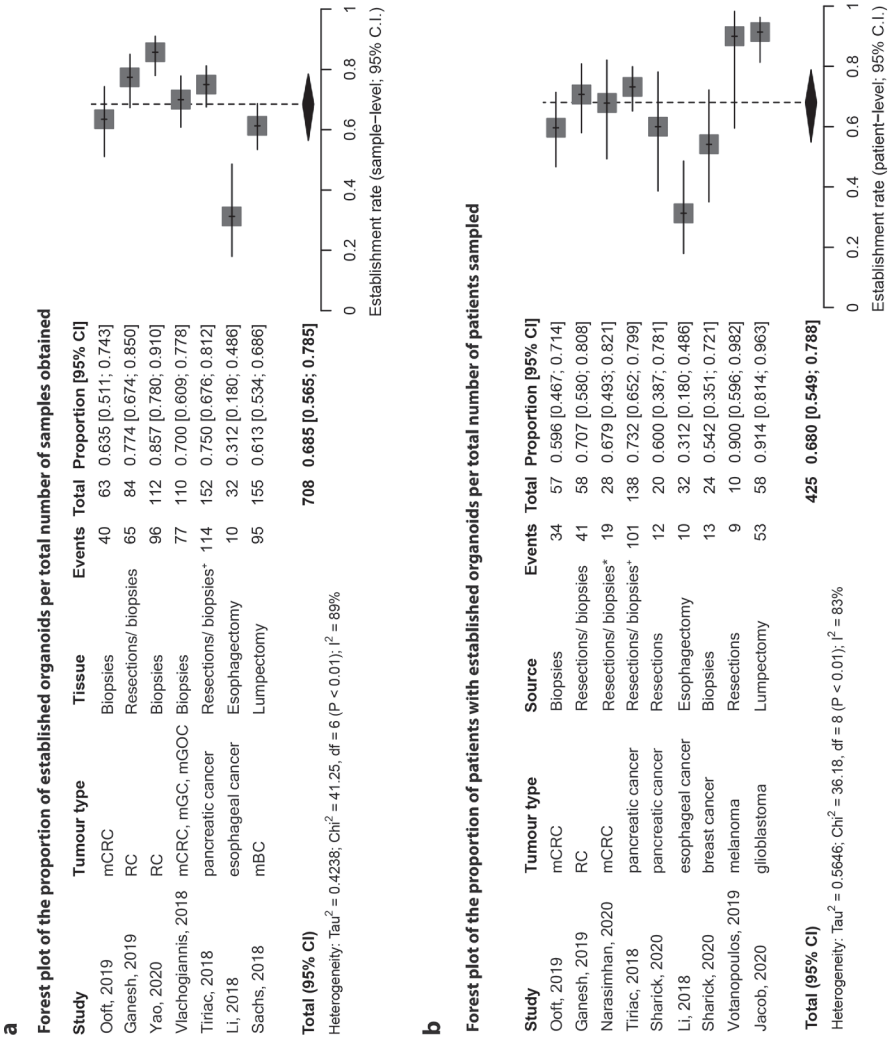
*Abbreviations:* 5-FU (5-fluorouracil), AC-T (doxorubicin + cyclophosphamide + paclitaxel), CAPIRI (capecitabine + irinotecan), AUC (area under the curve), CR (complete response), FOLFIRI (5-FU + irinotecan), GI<sub>50</sub> (concentration that inhibits growth of cancer cells by 50%), mCRC (metastatic colorectal cancer), mGC (metastatic gastric cancer), mGIC (metastatic gastrointestinal cancer), mGOC (metastatic gastroesophageal cancer), OMI (optical metabolic imaging), PD (progressive disease), PDO (patient-derived organoid), PR (partial response), RCB (residual cancer burden grade), SD (stable disease), TAS-102 (trifluridine/tipiracil).

Supplementary Figure 1. Forest plot of reported AUROC per study, treatment and index test



Suppl. Fig. 1. A forest plot for the reported AUROC per study (with 95% C.I.), treatment and index test (results were digitized from the figures from Ooft *et al.*). Abbreviations: 5-FU (5-fluorouracil), AUROC (area under the receiver operator curve), capec. (capecitabine), CAPIRI (capecitabine + irinotecan), C.I. (confidence interval), DRC (drug response curve), FOLFIRI (5-fluorouracil + irinotecan), FOLFOX (5-fluorouracil + oxaliplatin), GR (growth rate metrics), GR50 (50% GR), irin. (irinotecan), mCRC (metastatic colorectal cancer), oxali. (oxaliplatin), RC (rectal cancer). The analysis was performed in R (Version 3.6.1) using the 'mada' package <sup>31</sup>.

Supplementary Figure 2. Meta-analysis and forest plots of reported organoid establishment rates per study



► Suppl. Fig. 2. A random effects meta-analysis of the reported organoid establishment rates per study (listed in Supplementary Table 4) was performed using a generalized linear mixed model. a) Forest plot for the proportion of established organoids per total number of samples obtained (illustrated by the red square, with 95% C.I. error bars); b) Forest plot for the proportion of patients with established organoids per total number of patients sampled (illustrated by the red square, with 95% C.I. error bars).

*Abbreviations:* C.I. (confidence interval), mBC (metastatic breast cancer), mCRC (metastatic colorectal cancer), mGC (metastatic gastric cancer), mGOC (metastatic gastroesophageal cancer), RC (rectal cancer). \*Specimens were also obtained via rapid autopsy procedure. \*Operative specimens at staging laparoscopy, cytoreductive reductive surgery and HIPEC, or percutaneous biopsies. The establishment rates for Sharick *et al.* (2020) differ from the published rates, due to the study including patients with pancreatic intraepithelial neoplasia ( $n=2$ ), which were excluded from this analysis. The analysis was performed in R (Version 3.6.1) using the “binom”, “tidyverse”, “metafor” and “meta” packages<sup>51</sup>.

**Supplementary Table 4.** Reported organoid establishment rates

Study (author, year)	No. samples with established organoids	No. samples for organoid culture	Sample-level establishment rate (%)	No. patients with established organoids	No. patients with samples for organoid culture	Patient-level establishment rate (%)	Tumour type	Tissue	Comments
Ooft, 2019	40	63	0,63	34	57	0,60	mCRC	Biopsies.	
Ganesh, 2019	65	84	0,77	41	58	0,71	RC	Resections / biopsies.	
Yao, 2020	96	112	0,86				RC	Biopsies.	
Narasimhan, 2020				19	28	0,68	mCRC	Resections / biopsies (via CRS/HIPEC).	
Vlachogiannis, 2018	77	110	0,70		71		mCRC, mGC, mGOC	Biopsies.	
Tiriac, 2018	114	152	0,75	101	138	0,73	pancreatic cancer	Resections / biopsies (also included rapid autopsies).	75% (all organoid attempts, 72% fine-needle biopsies to 78% resection).

Study (author, year)	No. samples with established organoids	No. samples established for organoid culture	Sample-level establishment rate (%)	No. patients with established organoids	No. patients with samples for organoid culture	Patient-level establishment rate (%)	Tumour type	Tissue	Comments
Sharick, 2020				12	20	0,60	pancreatic cancer	Resections.	Differs from publication. Published rates included 2x pancreatic intraepithelial neoplasia (are not malignant).
Li, 2018	10	32	0,31	10	32	0,31	esophageal cancer	Esophagectomy.	
Driehuis, 2019							HNSCC	Resections / biopsies.	~ 60%, no further data available.
Sharick, 2020				13	24	0,54	breast cancer	Resections.	
Sachs, 2018	95	155	0,61				mBC	Lumpectomy.	Through improvements, latest rate >80%.
Votanopoulos, 2019				9	10	0,90	melanoma	Resections.	
Jacob, 2020				53	58	0,91	glioblastoma	Lumpectomy.	

The organoid establishment rates (per samples obtained for organoid culture and per patient with samples available for organoid culture) were extracted from the original studies. For studies not listed in this table, an organoid establishment rate was not reported.

*Abbreviations:* HNSCC (head & neck squamous cell carcinoma), mBC (metastatic breast cancer), mCRC (metastatic colorectal cancer), mGC (metastatic gastric cancer), mGOC (metastatic gastric-esophageal cancer), no. (number) and RC (rectal cancer).



# COLORECTAL CANCER ORGANOIDS REFLECT PATIENT RESPONSE TO SYSTEMIC TREATMENT

Manuscript in preparation.

Emerens Wensink<sup>†</sup>  
Lidwien Smabers<sup>†</sup>  
Carla Verissimo  
Esmee Koedoot  
Katerina-Chara Pitsa  
Maarten Huisman  
Celia Higuera Barón  
Ricardo Korporaal  
Emma Teal  
Mayke Doorn  
Inez van Weersch  
Hamdy Warda  
Edwin van Oosten  
Roshni Nair  
Liselot Valkenburg  
Geert Cirkel  
Anneta Brousal  
Jorieke Salij  
René Overmeer  
Manon Braat  
Sjoerd Elias  
Robert Vries  
Onno Kranenburg  
Miriam Koopman  
Sylvia Boj<sup>‡</sup>  
Jeanine Roodhart<sup>‡</sup>

<sup>†</sup> Both authors share equal contribution.

<sup>‡</sup> Both authors share last authorship.

## ABSTRACT

### Background

One of the holy grails in cancer research is being able to predict to which treatment a patient will respond, which would enable personalized treatment and increase survival of patients. Personalized treatment in oncology has been largely based on genomic biomarkers. However, for most anticancer agents, especially chemotherapy, no genetic markers are available. As a result, oncologists are currently unable to predict treatment response for individual patients. A promising predictive biomarker is *in vitro* response testing using patient-derived organoids (PDOs). We examined if PDOs predict response to standard-of-care agents in metastatic colorectal cancer (mCRC) patients.

### Methods

Tissue was obtained for PDO establishment from 18 patients with mCRC who gave informed consent via the HUB-Cancer Biobank and OPTIC study prior to starting a novel line of systemic treatment for mCRC. Drug screens were performed by exposing PDOs to a concentration range of 5-FU, irinotecan/5-FU, irinotecan or panitumumab for five days. Drug sensitivity was measured by cell viability. Growth rate (GR) inhibition metrics were used for drug response curve (DRC) fitting (area under the DRC ( $GR_{AUC}$ ),  $GR_{max}$  and  $GR_{50}$ ) to quantify organoid sensitivity. Organoid sensitivity was compared to the patient's response, which was determined by % change in the size of the target lesions on response scans (RECIST) after central revision. We compared *in vitro* sensitivity for all organoids based on *RAS/BRAF* mutational status and primary tumour location.

### Results

Associations were seen between organoid response and patient response for all treatment types. The best association was found for  $GR_{AUC}$  and % change in tumour size. A classifier based on normalized  $GR_{AUC}$  could predict patient response to chemotherapy correctly for 13/16 patients, with a sensitivity of 77.8% and specificity of 85.7%. PDOs show differential sensitivity to treatments. PDOs from patients previously exposed to chemotherapy showed increased resistance to chemotherapy but not to panitumumab. Organoids derived from left-sided colon *RAS/BRAF*-wildtype tumours were more sensitive to panitumumab versus *RAS/BRAF*-mutant and wildtype right-sided or rectal tumours.

## Conclusions

Our results demonstrate that organoid response correlates with patient response to standard-of-care treatment, suggesting that organoids can serve as a predictive biomarker and guide personalized cancer treatment. Associations were seen for all treatment types. The differential sensitivity of organoids to treatments implies that organoid drug screening facilitates treatment prioritizing, to improve patient response. Interestingly, we confirmed panitumumab resistance for organoids from right-sided and rectal *RAS/BRAF*-wildtype tumours. We observed varying organoid sensitivities to panitumumab for left-sided colon *RAS/BRAF*-wildtype tumours, suggesting that organoids may further aid in guiding anti-EGFR treatment. Our findings should be confirmed for other available treatments and validated in a larger prospective trial, such as the ongoing OPTIC trial, to facilitate the implementation of organoids in clinical treatment.

## INTRODUCTION

For metastatic colorectal cancer (mCRC) patients, more advanced systemic treatment options are available than ever before. In addition to different types of chemotherapy which form the backbone of treatment, targeted treatment options for molecular subgroups are increasingly used, including EGFR-targeting therapy for *RAS/BRAF*-wildtype, encorafenib/cetuximab for *BRAF*-mutant and inhibitors for *KRASG12C*-mutant mCRC tumours<sup>1</sup>. Unfortunately, treatment response in mCRC patients varies from 10%-90%, depending on patient and tumour characteristics, treatment type and treatment line<sup>2-5</sup>. One of the holy grails in cancer research is being able to predict to which treatment a patient will respond, which would enable personalized treatment and increase survival of patients. Considerable research has been done in the field of predictive biomarkers, with a focus on genetic biomarkers. Genetic biomarkers are, however, not available for chemotherapy and do not adequately predict treatment response for targeted treatments<sup>6</sup>. As a result, oncologists are currently unable to predict treatment response or estimate the contribution of the different agents in combination therapy for individual patients. A promising predictive biomarker is *in vitro* response testing using patient-derived organoids (PDOs).

PDOs are (cancer) stem-cell derived, 3D self-organizing and proliferating structures comprised of epithelial cells representing their corresponding tumour<sup>7-10</sup>. PDOs are well suited for drug screening due to their long-term stability, applicability for high throughput screens and inherent heterogeneity<sup>10,11</sup>. Therefore, organoids represent a novel promising biomarker which can predict treatment response for various cancer and treatment types and can be used to identify new treatment types<sup>9</sup>. In previous studies with CRC patients, organoid response correlated to patient response<sup>9,12-19</sup>. More research is needed to confirm the predictive value of organoids. It is still unclear if organoids can predict response over multiple treatment lines and if primary tumour organoids can predict response for treatment in the metastatic setting. Besides, the effect of mutational status and primary tumour location (sidedness) on organoid drug sensitivity is not widely understood.

In this study, we investigate the association of PDOs with patient response in a cohort of 18 mCRC patients treated with systemic therapy, from whom 19 established PDOs were analyzed. We optimize the organoid drug screening methods and analyze how strong the associations between clinical response and organoid response are for predefined drug

response curve (DRC) metrics. We assess if organoids can be used to predict treatment response in patients for standard-of-care systemic chemotherapies and targeted treatment in different treatment lines after PDO establishment. We also examine how mutational status and primary tumour location affect drug sensitivity for standard-of-care systemic chemotherapies and targeted treatment. Our results will be validated in our ongoing prospective study, the OPTIC trial, which establishes PDOs for mCRC patients and examines if PDO response is associated with patient response<sup>20</sup>.

## METHODS

### Study design

We included patients based on the following criteria: were diagnosed with mCRC, CRC organoids were generated (from primary tumour or metastatic tissue), tissue for organoid derivation was obtained prior to receiving a new line of standard-of-care systemic therapy for metastatic disease and detailed clinical data (including treatments given, response scans and vital status) was known.

Tissue samples were collected during surgery or fine-needle biopsies within the Biobanking protocol HUB-Cancer (TCBIO #12-093) or within the prospective clinical trial Organoids to Predict Therapy Response In Colorectal Cancer (OPTIC<sup>20</sup> #17-356), which were approved by the medical ethical committee of the University Medical Center Utrecht (UMCU). Written informed consent was obtained prior to study inclusion and the study was performed in accordance with the Declaration of Helsinki.

### Clinical data collection

Clinical data was collected from the electronic patient database by two data managers (GEW, LS) who were blinded from the drug screen response data. The following variables were collected: gender, date of metastatic disease diagnosis, number and location of metastases upon diagnosis of metastatic disease, mutational status of the tumour (including *RAS* and *BRAF* status) and sidedness of the primary tumour (defined as right-sided (coecum-transverse colon), left-sided (splenic flexure-sigmoid) and rectosigmoid/rectal. Class I and II *BRAF* mutations were considered pathogenic, but not Class III<sup>21</sup>. *RAS* mutations included *KRAS* amplifications. Treatment information was collected, including primary tumour resection, local treatment of metastases (including metastasectomy or ablation), adjuvant chemotherapy (type

treatment received) and the type and duration of treatment for each systemic palliative therapy line given. To assess if a patient was chemotherapy naïve prior to organoid establishment, any given systemic treatment (including adjuvant) was counted as chemotherapy exposure prior to organoid establishment.

Response CT scans for each patient's treatment line were revised according to RECIST 1.1 criteria by a radiologist who was blinded from the drug response data (MB). The radiologist determined the best RECIST response (complete response (CR), partial response (PR), stable disease (SD) or progressive disease (PD)) during treatment. The size of the target lesions for the baseline and best response scans were measured by the radiologist according to the RECIST criteria, from which the percentage size change of the target lesions were calculated. If a patient had multiple scans with a partial response, the scan with the deepest decrease in percentage size change was used for the percentage size change of target lesions measurement. Progression-free survival for each treatment line was determined as the days between the start of a given treatment and the date of progressive disease on the response scan. If a patient received local treatment for metastatic target lesions, the progression-free survival date was censored for the date of the local treatment.

## **Organoid culture and in vitro drug screens**

Patient-derived tumour organoids isolation was performed as described by van de Wetering *et al.* (2015)<sup>10</sup>. PDOs were passaged using mechanical dissociation by pipetting and enzymatic dissociation with TrypLE Express (Gibco, Breda, Netherlands, #12604021) for 5-10 minutes at 37°C, washed using organoid culturing medium and re-plated by embedding in a solution of ice-cold Matrigel® (70%) and organoid culturing medium (Supplementary Table 1) and plated as droplets on a pre-warmed 6-well plate. After solidification of the Matrigel®, organoid culturing medium was added to the plates. Medium was refreshed three times per week.

Organoid identity was confirmed using single-nucleotide polymorphism (SNP) array targeting 60 SNPs using TagMan OpenArray technology in combination with the QuantStudio 12K Flex Real-Time PCR System at the Utrecht Sequencing Facility (USEQ). The genetic distance of tumour DNA versus healthy blood DNA was calculated. PDOs with a genetic distance <5 were included in the drug screens.

All drug screens and derivation of drug response curve metrics were performed by the Foundation Hubrecht Organoid Technology. The timeline of the drug screen

was: day -1 PDOs were sheared and re-plated. Day 0, organoids were harvested by incubating with 1 mg/mL Dispase II (Life Technologies Europe B.V., Zuid-Holland, the Netherlands, #17105041) for 30 minutes at 37°C, washed twice using organoid culturing medium with Rho-kinase inhibitor (10  $\mu$ M, Abmole Bioscience, Brussels, Belgium, #M1817), filtered using a 100  $\mu$ m mesh filter to collect the organoids which have passed through the filter. Organoids were resuspended in organoid screening medium with Rho-kinase inhibitor (5  $\mu$ M, Abmole Bioscience, Brussels, Belgium, #M1817) with 5% Matrigel® for a final concentration of 125,000 PDOs/20 ml. The organoid screening medium comprised of the organoid culture medium with the following adjustments: 1 ng/ml human recombinant EGF (Sigma-Aldrich, #A9165) and 10 ng/ml heregulin (Peprotech, London, UK, #100-03) to avoid interference by epidermal growth factor (EGF) with panitumumab sensitivity. Using an automated Multidrop™ Combi Reagent Dispenser, 40  $\mu$ l of PDO suspension was dispensed in clear-bottomed, black-walled 384-well plates with ultra low-attachment coating (Corning, Zuid-Holland, the Netherlands, #4588). A ten-point concentration range of the treatments (Supplementary Table 2), the positive control (staurosporin) and negative control (1% Dimethyl sulfoxide, Phosphate Buffered Saline or combination depending on the solvent used), were dispensed in technical triplicates using a Tecan Fluent dispenser or Tecan D300E dispenser (latter for SN-38). The following commonly used CRC treatments were screened: 5-fluorouracil (5-FU), SN-38 (active metabolite of irinotecan), combination treatment 5-FU + SN-38, and panitumumab. Organoid response to bevacizumab could not be tested, since its main effect is on vascularization. On day 5, readouts were obtained by quantifying cell viability using CellTiter-Glo 3D (Promega, #G9681, 40  $\mu$ l/well) with a Tecan Spark plate reader. A baseline readout was measured on day 0 in a separate plate with PDOs without treatment. Drug screens were performed in at least two duplicate experiments on different days.

Growth rate inhibition (GR) metrics values were calculated and four parameter log-logistic curves were fit using the DRC package in R<sup>22</sup>. GR analysis calculates drug response using normalized growth rate inhibition in treated versus untreated conditions, which results in parameters that are normalized to a single cell division and thus independent of cell division rate<sup>23</sup>. GR values range from 1 to -1, with 0 to 1 for partial growth inhibition, 0 for complete inhibition of cell growth, and 0 to -1 for cell death. The GR<sub>AUC</sub> (area under the curve of the measured values of the drug response curve (DRC)), GR<sub>50</sub> (concentration of 50% growth rate inhibition) and

$GR_{max}$  (GR value at the highest concentration) were calculated. For several PDOs, the  $GR_{50}$  was calculated using a two or three parameter log-logistic curve (setting the organoid lower and upper plateau at a  $GR=-1$  and  $GR=1$ ), since the measured lower or upper plateau far exceeded -1 or 1, respectively, for certain treatment types. To allow comparison between different treatment types, with different concentration ranges, normalized values for  $GR_{AUC}$  and  $GR_{50}$  were calculated using the maximum and minimum measured parameter per treatment. Quality of the drug screen assays were assessed using the  $Z'$ -factor and strictly standardized mean difference (SSMD). Biological replicates were not included in the analysis if the runs had a  $Z'$ -factor  $<0.4$ , a SSMD  $<3.0$  or if there were doubts concerning the results (e.g. dispensing error).

Organoid screens for oxaliplatin-based treatments (5-FU and oxaliplatin) are ongoing and thus not included in this manuscript.

## Statistical methods

Patient response was measured using the percentage size change of the target lesions for each treatment line given. Organoid response parameters were chosen prior to analysis based on literature and physiological relevance and was measured using  $GR_{AUC}$ ,  $GR_{max}$  and  $GR_{50}$  per treatment. Once organoid response and clinical response databases were complete, the data was unblinded and analyzed. Associations between clinical and organoid response were examined using scatterplots, and where applicable, using Spearman correlation tests (for two continuous variables), Mann-Whitney U-test (for 2 groups) or Kruskal-Wallis test (for  $>2$  groups).

For systemic chemotherapy, a diagnostic receiver operator characteristic (ROC) curve was generated using the parameters with the best observed association. In order to assess if organoid drug screening results (normalized  $GR_{AUC}$ ) could predict observed patient response (decrease in tumour size) for a given treatment, a logistic model with as outcome dichotomized change in tumour size (decrease versus stable/increase in size) and as predictor continuous normalized  $GR_{AUC}$  was created to estimate the predicted observed patient outcome (decrease or stable/increase in tumour size). Patients receiving chemotherapy (5-FU monotherapy, irinotecan monotherapy and 5-FU/irinotecan combination treatment) were included in the analysis.

Analysis was performed in R (version 4.0.3) with the GRmetrics (version 1.16.0), DRC (version 3.0-1) and pROC (version 1.18.0) packages.

## RESULTS

### Patient characteristics

The study cohort consists of 18 patients with mCRC who were treated with systemic treatment for metastatic disease. Tissue was derived for organoid establishment before the patient received a new standard-of-care systemic treatment (**Figure 1**). Nineteen organoids were included, three from primary tumours and sixteen from liver metastases (**Table 1**). Most patients were chemotherapy naïve when the PDO was established (14/18).

**Figure 1. Schematic overview of the study**

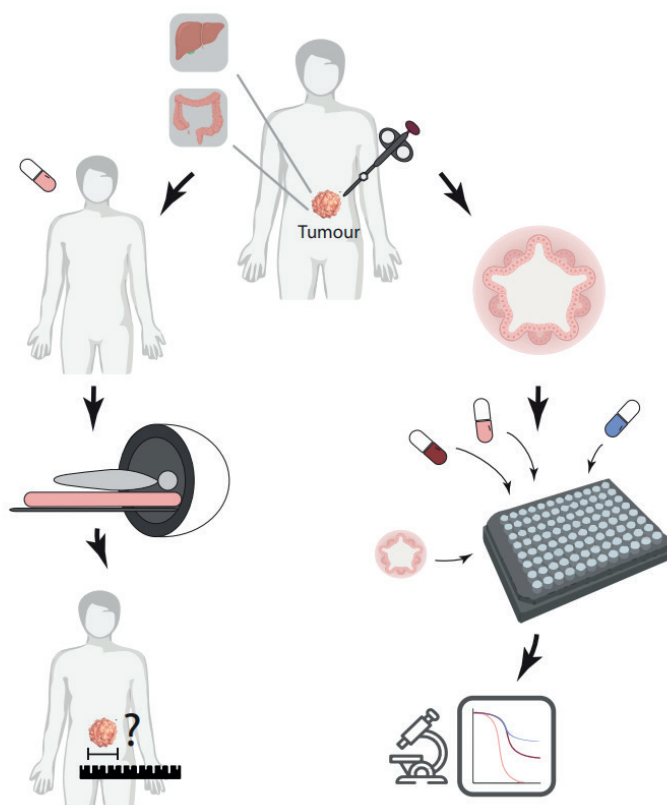


Fig 1. Tissue from metastatic and primary CRC tumours were obtained prior to starting a new line of systemic treatment. Organoids were cultured and screened for standard-of-care treatments while the patient received systemic treatment. For all patients, organoid and patient response were compared for treatment lines given after the organoid was established.

Table 1. Overview of evaluable treatments per patient

Patient ID	Evaluable treatment 1	Response 1	Evaluable treatment 2	Response 2	Evaluable treatment 3	Response 3	Mutational status*	PDO origin
16	5-FU	SD	Panitumumab	↑ PD			BRAF mutant	primary
6	5-FU	SD					RAS mutant	metastatic
4	5-FU	SD	5-FU & irinotecan	SD			RAS mutant	metastatic
7	5-FU	SD					RAS mutant	metastatic
18	5-FU	↑ PD	5-FU & oxaliplatin	SD			Wildtype	metastatic
11	5-FU	↑ PD					Wildtype	metastatic
1	Irinotecan	SD					RAS mutant	metastatic
2	5-FU & irinotecan	SD					RAS mutant	metastatic
5	5-FU & oxaliplatin	↓ PR	5-FU & irinotecan	↑ PD			RAS mutant	primary
3	5-FU & oxaliplatin	↑ PD					RAS mutant	primary
9	5-FU & oxaliplatin	SD					BRAF mutant	metastatic
13	5-FU & oxaliplatin	SD	Irinotecan	SD			RAS mutant	metastatic
17	5-FU & oxaliplatin	↓ PR	5-FU & irinotecan	SD			RAS mutant	metastatic
14	5-FU & oxaliplatin	↓ PR	5-FU & irinotecan	↓ PR			RAS mutant	metastatic
12	5-FU & oxaliplatin	↓ PR	5-FU & irinotecan	SD	Panitumumab	SD	Wildtype	metastatic
10	5-FU & oxaliplatin	↓ PR	Panitumumab	SD			Wildtype	metastatic
15	5-FU & oxaliplatin	↓ PR	5-FU & irinotecan	SD	Panitumumab	SD	Wildtype	metastatic
8	5-FU & oxaliplatin	↓ PR	Irinotecan	↑ PD	Panitumumab	SD	Wildtype	metastatic

The evaluable treatments, for the first, second and third line after organoid establishment are shown for each patient, along with the patient's best RECIST response during treatment. The mutational status and organoid tissue origin are also displayed. All patients had a proficient mismatch repair (pMMR) tumour. For patient ID 10, 2 organoids were included which were established from different liver metastases. \*Mutational status for RAS and BRAF status.

Abbreviations: 5-FU (5-fluorouracil), ID (identification), PD (progressive disease), PDO (patient-derived organoid), PR (partial response), SD (stable disease).

The characteristics of the included patients are heterogeneous and reflect the general mCRC population. Mutational status of included patients vary with ten *RAS*-mutant, two *BRAF*-mutant and six *RAS/BRAF*-wildtype tumours and with regards to the location of the original primary tumour, twelve patients had a left-sided tumour, five a right-sided tumour and one multiple primary tumour locations (**Supplementary Table 3**). Evaluable treatments in the first line of treatment after PDO establishment include 5-FU monotherapy ( $n=6$ ) and irinotecan-based treatment ( $n=2$ ). Later-line treatments include irinotecan-based treatment ( $n=8$ ) and panitumumab ( $n=5$ ). The median PFS for first line treatment is 307 days (95% C.I. 127-not reached).

## Organoid drug screen quality

The drug screen quality was analyzed by examining the Z'-factor, correlation in drug screen metrics among the replicates and Bland-Altman plot for the difference in normalized growth rate metric between replicates (**Supplementary Figure 2**). The average Z'-factor for all organoids was 0.60, indicating a good quality drug screen. There was a high correlation among the calculated  $GR_{AUC}$ ,  $GR_{max}$ , and  $GR_{50}$  for each treatment in separate runs (Spearman correlation  $>0.90$ ). The variation between duplicate assays of each PDO was minimal (**Supplementary Figure 2A**).

The drug response curves (DRC) for each treatment screened and the derived parameters are displayed in **Supplementary Figure 3** and **Supplementary Table 4**. The  $GR_{AUC}$  and  $GR_{max}$  derived from the DRCs were correlated, while the other parameters ( $GR_{AUC}$  versus  $GR_{50}$  and  $GR_{max}$  versus  $GR_{50}$ ) were not correlated (**Supplementary Figure 4**).

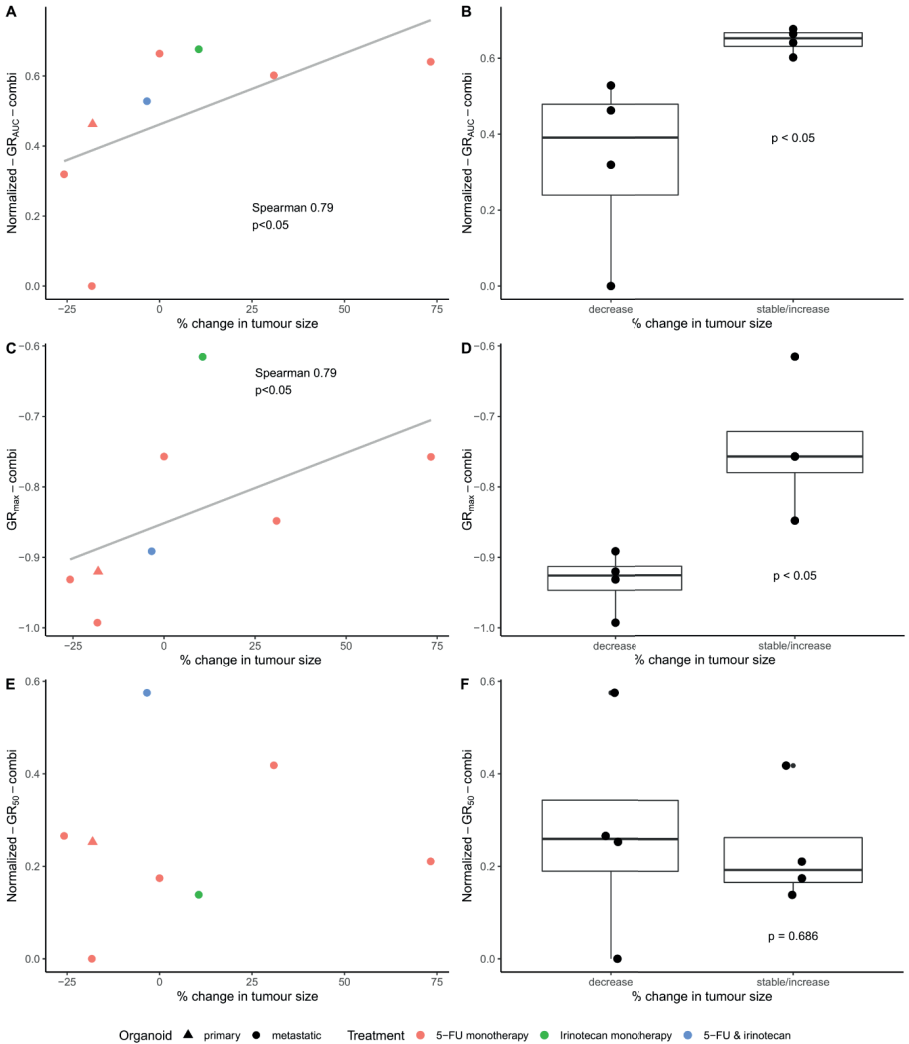
## Organoid response is associated with patient response

### *Combined analysis for different treatment types*

A combined analysis was performed of evaluable treatment lines for eight patients receiving 5-FU monotherapy (capecitabine,  $n=6$ ), irinotecan monotherapy ( $n=1$ ) or 5-FU/irinotecan combination therapy ( $n=1$ ). Organoid response determined by either normalized  $GR_{AUC}$  or  $GR_{max}$  correlated with the change in tumour size in the patient, with a Spearman's correlation of 0.79,  $p < 0.05$  (same correlation coefficient for normalized  $GR_{AUC}$  and  $GR_{max}$ ; **Figure 2**). The median normalized  $GR_{AUC}$  in patients with a decrease in tumour size was 0.39 (Interquartile Range (IQR) 0.24 to 0.48) versus 0.65 (IQR 0.63 to 0.67,  $p < 0.05$  for the Mann-Whitney U-test) in patients with

stable/increase in tumour size. Similarly, the median  $GR_{max}$  was -0.93 in patients with a decrease in tumour size (IQR -0.95 to -0.91) versus -0.76 in patients with a stable/increase in tumour size (IQR -0.78 to -0.72,  $p < 0.05$ ). There was no clear association between the normalized  $GR_{50}$  with the observed response in patients ( $p = 0.69$ ).

**Figure 2.** Decrease in tumour size during treatment is associated with *ex vivo* organoid response



**Fig 2.** The association between organoid (normalized  $GR_{AUC}$ ,  $GR_{max}$  and normalized  $GR_{50}$ ) and patient response (observed % change in target lesion size during treatment) is shown for the evaluable treatments given in the first line after PDO establishment (5-FU, irinotecan and irinotecan-based treatment). A) Scatterplot of normalized  $GR_{AUC}$  versus % size change of the

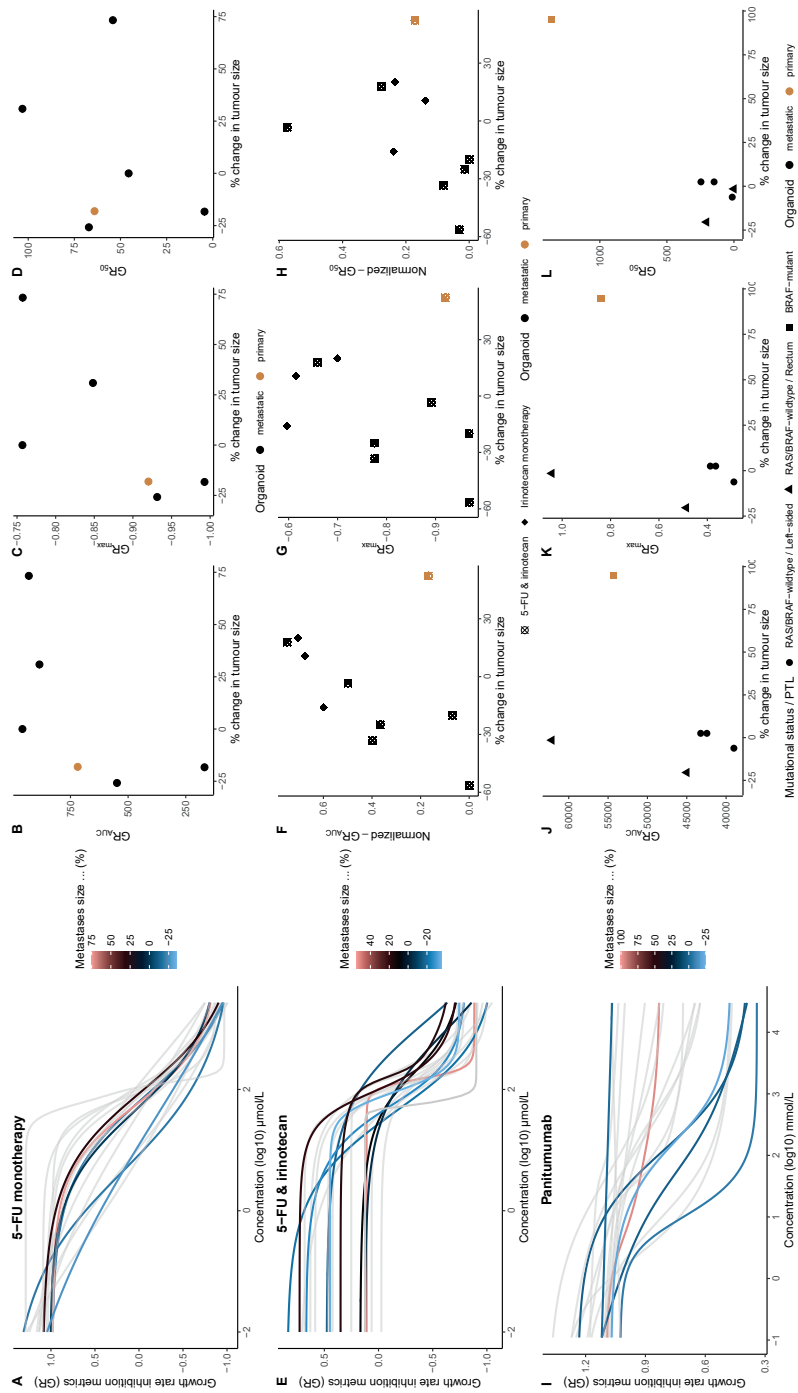
target lesions with a regression line, Spearman's correlation and corresponding  $p$ -value. **B)** Boxplot of normalized  $GR_{AUC}$  for patients with decrease in tumour size versus stable/increase in tumour size during treatment, with  $p$ -values for a Mann-Whitney U-test. Similarly, plots **C** and **D** display the results for normalized  $GR_{max}$  versus change in tumour size and plots **E** and **F** for normalized  $GR_{50}$  versus change in tumour size.

**Abbreviations:**  $GR_{AUC}$  (area under the curve for a normalized growth rate inhibition drug response curve),  $GR_{max}$  (maximum normalized growth rate inhibition at the highest concentration) and  $GR_{50}$  (the concentration of drug at which the normalized growth rate inhibition is 0.5).

### *Analysis per treatment type*

The association between organoid response and the observed response in patients was examined per treatment type, including later line treatment, for 5-FU monotherapy, irinotecan-based treatment (irinotecan monotherapy and 5-FU/irinotecan combination treatment) and panitumumab (**Figure 3**). For six patients receiving 5-FU monotherapy (all in the first line), a decrease in tumour size during treatment was associated with organoid response ( $GR_{AUC}$ ,  $GR_{max}$  and  $GR_{50}$ , **Figure 3A-D**). For ten patients receiving irinotecan-based treatment, the change in tumour size during treatment was associated with organoid response in most patients (normalized  $GR_{AUC}$ ,  $GR_{max}$  and normalized  $GR_{50}$ , **Figure 3E-H**). Eight of the ten patients received irinotecan-based treatment in the second line after organoid establishment, since the organoid was established prior to the patient having been exposed to first line systemic treatment. PDO response and clinical response were similar for patients who were sensitive during irinotecan-based treatment and for whom the PDOs were sensitive at baseline. Similarly, PDO response and clinical response were concurrent for patients who were resistant during irinotecan-based treatment and for whom the PDOs were resistant at baseline. However, there was one patient who was resistant to irinotecan-based treatment but who had a sensitive PDO at baseline, likely since the patient had acquired resistance to irinotecan-based treatment during first line systemic treatment, which the PDO had not been exposed to. For irinotecan-based treatment, we compared organoid response for combination screening (exposing organoids to an anchor concentration of SN-38 and a ten-point concentration range of 5-FU) versus screening each agent separately (exposing organoids to 5-FU in one condition and SN-38 in a second condition). The measured  $GR_{AUC}$  for combination screening (5-FU + SN-38) was highly correlated with the measured  $GR_{AUC}$  for 5-FU monotherapy (Spearman  $\rho$  0.83,  $p < 0.05$ ), but less strongly for SN-38 monotherapy (Spearman  $\rho$  0.38,  $p = 0.11$ , **Supplementary Figure 5**). For five patients (6 PDOs) receiving panitumumab treatment, the decrease in tumour size during treatment was associated with organoid response in most patients ( $GR_{AUC}$ ,  $GR_{max}$  and  $GR_{50}$ , **Figure 3I-L**).

**Figure 3.** Association between organoid response and patient response for 5-FU, irinotecan-based and panitumumab treatments



► Fig 3. The association between observed % change in target lesion size during treatment and organoid response ( $GR_{AUC}$ ,  $GR_{max}$ , and  $GR_{50}$ ) is shown for 3 treatment types (5-FU monotherapy, irinotecan-based treatment and panitumumab) across multiple treatment lines. The organoid origin (primary versus metastatic lesion) are indicated by orange and black, respectively. The DRC curves of organoid sensitivity are shown for 5-FU (A), 5-FU and irinotecan combination treatment (E) and panitumumab (I). Scatterplots of % size change during 5-FU treatment versus  $GR_{AUC}$ ,  $GR_{max}$ , and  $GR_{50}$ , respectively in plots B-D are shown. Similarly, plots F-H, display % size change during irinotecan-based treatment versus  $GR_{AUC}$ ,  $GR_{max}$ , and  $GR_{50}$  organoid sensitivity for the patient's given treatment (irinotecan monotherapy indicated by a diamond shape or irinotecan-combination treatment indicated by a cross shape). Plots J-L show % size change during panitumumab treatment versus  $GR_{AUC}$ ,  $GR_{max}$ , and  $GR_{50}$  organoid sensitivity. Mutational status and primary tumour location are indicated by the shapes (circle left-sided *RAS/BRAF*-wildtype, triangle rectal *RAS/BRAF*-wildtype and square *BRAF*-mutant). Normalized  $GR_{AUC}$  and  $GR_{50}$  was analyzed for irinotecan-based treatments, due to differences in concentrations used for irinotecan monotherapy versus irinotecan combination treatment.

**Abbreviations:**  $GR_{AUC}$  (area under the curve for a normalized growth rate inhibition drug response curve),  $GR_{max}$  (maximum normalized growth rate inhibition at the highest concentration) and  $GR_{50}$  (the concentration of drug at which the normalized growth rate inhibition is 0.5).

In an exploratory assessment, we examined if a classifier based on the organoid normalized  $GR_{AUC}$  could predict patient response to chemotherapy (decrease in tumour size versus stable/increase in tumour size during treatment). A logistic regression model based on organoid normalized  $GR_{AUC}$  for 5-FU monotherapy and irinotecan-based treatment correctly classified the patient response for 13/16 patients (**Supplementary Figure 6**). The area under the ROC curve was 0.905 (based on a threshold for normalized  $GR_{AUC}$  below 0.599), with a sensitivity of 77.8% and specificity of 85.7%.

## Drug sensitivity in all organoids based on tumour characteristics

All organoids in the cohort were screened for commonly used treatments in CRC, allowing comparison of drug sensitivity for patients with different tumour characteristics, including prior chemotherapy exposure, mutational status, and primary tumour location ('sidedness'). When examining the sensitivity of organoids to treatment, a wide range of responses was seen (**Figure 4**). Some organoids display a generally resistant (e.g. PDO13) or sensitive (e.g. PDO14) phenotype to all treatments, while other organoids display a mixed treatment response (e.g. PDO6). PDOs generated from tumours exposed to chemotherapy prior to PDO establishment (including adjuvant chemotherapy) are generally more resistant to chemotherapy, but not to panitumumab, than chemotherapy naïve PDOs.

**Figure 4.** Heatmap for all PDOs and treatments

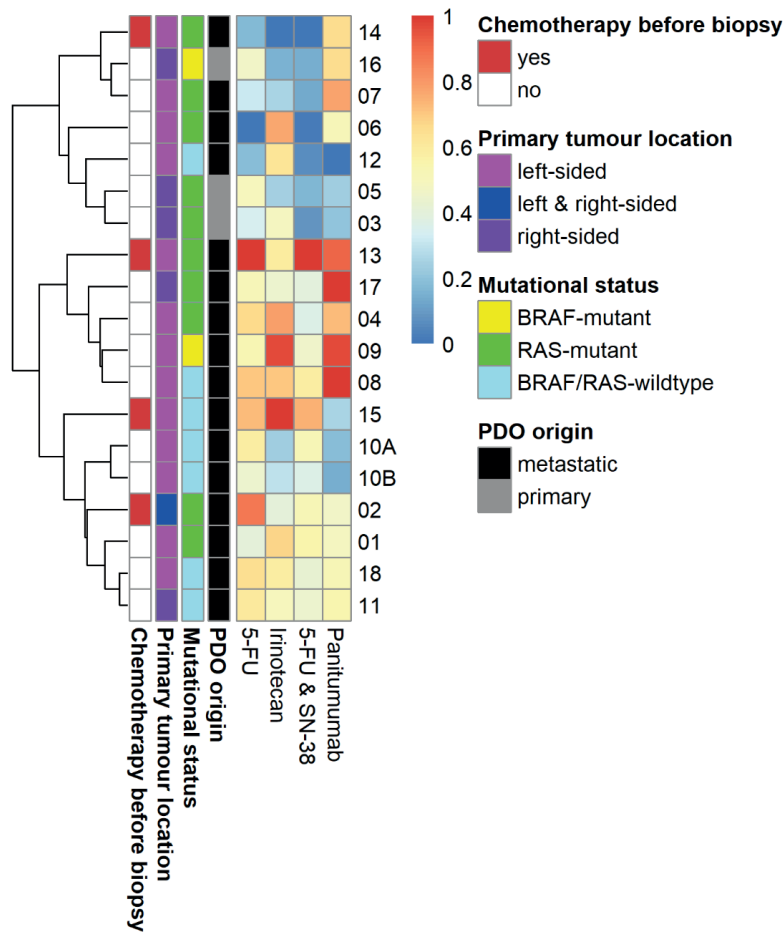
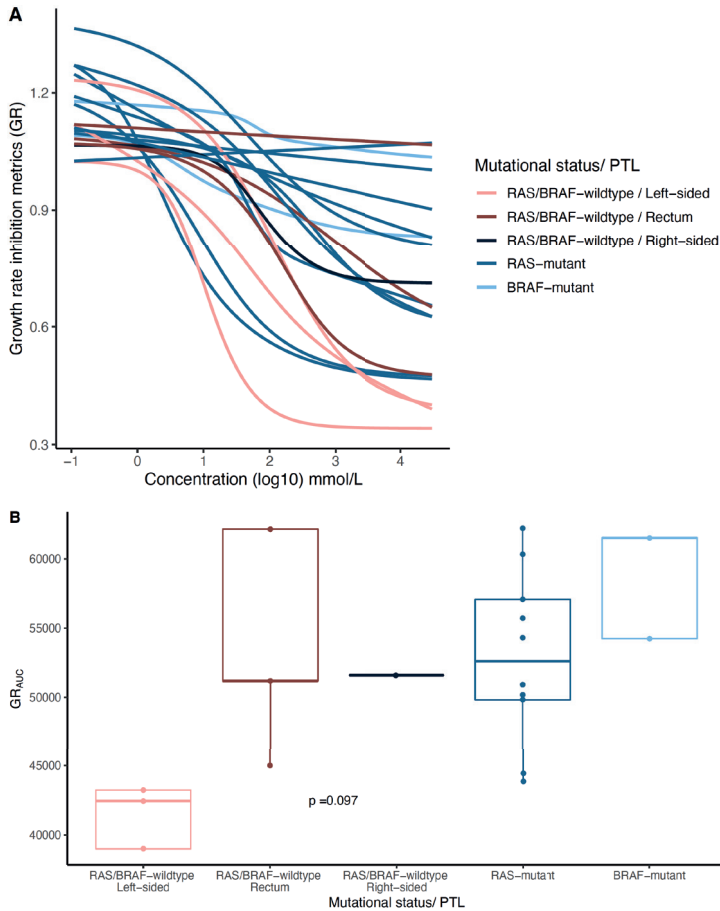


Fig 4. Clustered heatmap of the normalized  $GR_{AUC}$  with characteristics of the patient indicated in the first four columns for chemotherapy exposure prior to PDO establishment (adjuvant or first line), the primary tumour location (left- versus right-sided), mutational status (*BRAF*-mutant, *RAS*-mutant and *RAS/BRAF*-wildtype) and PDO origin (metastatic and primary tumour). The normalized  $GR_{AUC}$  are illustrated as a heatmap with a column for each treatment type examined. *Abbreviations:* 5-FU (5-flouruoracil), GR (normalized growth rate inhibition), PDO (patient-derived organoid), SN-38 (metabolite for irinotecan).

PDOs obtained from patients with a left-sided colon *RAS/BRAF*-wildtype tumour were more sensitive to panitumumab ( $GR_{AUC}$ ) compared to PDOs obtained from patients with right-sided or rectal *RAS/BRAF*-wildtype tumours, *RAS*-mutant tumours and *BRAF*-mutant tumours (**Figure 5**,  $p = 0.097$  for the Kruskal-Wallis test). Interestingly,

we observed varying organoid sensitivities to panitumumab within the group of left-sided colon *RAS/BRAF*-wildtype tumours. Neither mutational status nor tumour-sidedness related to organoid sensitivity to chemotherapeutic treatments (**Figure 4**).

**Figure 5.** Sensitivity for panitumumab based on primary tumour location and mutational status



**A)** The growth rate drug response curves for all organoids for panitumumab treatment is displayed, categorized per *RAS/BRAF* mutational status and primary tumour location (PTL) of the corresponding patient's tumour. **B)** Displays boxplots of the  $GR_{AUC}$  of the DRC for panitumumab sensitivity for organoids categorized according to the tumour's mutational status and primary tumour location (*RAS/BRAF*-wildtype and left-sided colon; *RAS/BRAF*-wildtype and rectal; *RAS/BRAF*-wildtype and right-sided; *RAS*-mutant and *BRAF*-mutant) along with the Kruskal-Wallis  $p$ -value.

**Abbreviations:** DRC (drug response curve), GR (normalized growth rate inhibition),  $GR_{AUC}$  (area under the curve for a normalized growth rate inhibition drug response curve), PTL (primary tumour location).

## DISCUSSION

Our study shows that organoid response correlates well with patient response to 5-FU, irinotecan-based and EGFR-targeting treatment. The associations between organoid and patient response were most consistently seen for the DRC parameters  $GR_{AUC}$  and  $GR_{max}$ , and less consistently for  $GR_{50}$ . Organoid response for panitumumab and irinotecan-based treatments, often given in the second or third line of treatment after PDO establishment, were also associated with clinical response. In our study, organoid sensitivity ( $GR_{AUC}$ ) predicted change in tumour size during chemotherapy with a sensitivity of 77.8% and specificity of 85.7%. In all organoids screened, neither mutational status nor sidedness related to organoid sensitivity to chemotherapeutic treatments. PDOs obtained from patients with a left-sided colon *RAS/BRAF*-wildtype tumour were more sensitive to panitumumab compared to PDOs obtained from patients with right-sided or rectal *RAS/BRAF*-wildtype, *RAS*-mutant and *BRAF*-mutant tumours. Our results encompass a cohort of patients with varying treatment types, mutational status, and treatment responses, offering insight into results for a wide range of standard-of-care treatments for mCRC patients.

Our findings confirm the reported results from previous studies in which associations were found for PDO response and clinical response in mCRC patients<sup>9,12–17</sup>. Although no studies reported results specifically for 5-FU monotherapy treatment, one study reported that the DRC AUC of the PDO was associated with PFS in seven rectal cancer patients who received 5-FU monotherapy or oxaliplatin-based combination treatment<sup>12</sup>. Irinotecan-based treatment was assessed in several studies<sup>13–15</sup>, of which one reported results separately for irinotecan-based treatment<sup>14</sup>. In mCRC patients receiving irinotecan-based treatment, several PDO drug response metrics (including  $GR_{AUC}$ ) were associated with patient response (RECIST response of the target lesion and PFS)<sup>14</sup>, which is in line with our results. For EGFR-targeting treatment, one study reported that PDO response (unclear which parameter was measured) matched the clinical response in three of the four mCRC patients who received cetuximab treatment<sup>16</sup>. As opposed to the previous study, in which half of the organoids were established after the patient had progressed during cetuximab treatment<sup>16</sup>, our results are based upon organoids established prior to EGFR-targeting treatment exposure. For all chemotherapy treatments, we were able to predict change in tumour size using organoid sensitivity ( $GR_{AUC}$ ) with a sensitivity of 77.8% and specificity of 85.7%. Our predictive values agree with the previously reported predictive value for patient response to standard-of-care mCRC treatment<sup>14,15</sup>.

In contrast to previous studies, where patient response was measured using RECIST response and PFS<sup>12,14–17</sup>, we measured clinical response using the percentage change in tumour size during treatment. The percentage change in tumour size has two main advantages: it allows comparison for multiple treatment types which can have inherent variations in PFS and it avoids loss of information through categorization of RECIST response. The association between patient and organoid response was highest when measuring organoid response using the DRC parameters  $GR_{AUC}$  and  $GR_{max}$  but was weak or not present for  $GR_{50}$ . Different organoid response parameters have been used in previous studies, including GR metrics, AUC and  $IC_{50}/GR_{50}$ <sup>12–15</sup>; however, it is not yet clear which parameter has the strongest association with clinical response. The information captured by each DRC parameter varies, with  $GR_{50}$  representing drug potency,  $GR_{max}$  drug efficacy and  $GR_{AUC}$  representing potency and efficacy combined as the cumulative effect of the drug<sup>24</sup>. The DRC parameters capture distinct information and may vary systematically with the preclinical model and drug being tested<sup>24</sup>. There is a strong focus in the literature on using potency ( $IC_{50}$  or  $GR_{50}$ ) to quantify organoid response, nonetheless, using AUC may be beneficial by encompassing both potency and efficacy<sup>24</sup>. Clinically, the second half of the DRC may capture more relevant information for patient response than the concentration at which a drug achieves 50% of the maximal response ( $IC_{50}/GR_{50}$ ), since the latter part of the curve illustrates the maximal level of cell death achieved and the heterogeneous treatment resistant population present at the highest tested concentration. An added benefit is that the AUC and maximum effect can be determined without curve fitting, which avoids curve fitting artefacts. A strength in our study is the systematic comparison of DRC parameters reflecting potency and efficacy for different treatment types and lines, to capture the strongest association with clinical response. Likewise, drug screen medium composition, read-outs and combination screen layout vary widely among studies, and the ideal set-up is as of yet unknown<sup>9,25</sup>. A standardized screening and analysis protocol would ideally allow comparison across studies. Furthermore, using growth rate metrics to analyze organoid response avoided bias by the proliferation rate of organoids<sup>23</sup>.

Our results demonstrate that organoids show a differential response to panitumumab based on the primary tumour location, in addition to *RAS/BRAF* mutational status. Organoids from patients with a left-sided colon *RAS/BRAF*-wildtype tumour were the most sensitive to panitumumab in our cohort. Our results are concurrent with patient response in large clinical trials<sup>26,27</sup>. *RAS/BRAF*-wildtype organoids from rectal tumours were less sensitive to panitumumab compared to left-sided colon tumours, which although reported<sup>28</sup>, is generally not analyzed separately for rectal and left-sided colon tumours<sup>26,27</sup>. Importantly, we observe varying sensitivity to panitumumab within organoids from left-sided colon

*RAS/BRAF*-wildtype tumours. Considering the low numbers of patients per mutational status and primary tumour location, our results should be confirmed in a larger cohort of patients. Our results suggest that patient-derived organoids may further aid in guiding anti-EGFR treatment compared to molecular pathology and primary tumour location alone.

To enable personalized treatment, PDO drug screening may guide the treatment of patients with mCRC by selecting the most effective treatment and avoiding over-treatment. In a later line setting, when treatments often have a low response rate, PDO drug screening results can be used to select a treatment to increase the chance of response. In earlier line settings, when combination treatments are often given, PDOs may be used to avoid over-treatment by comparing if monotherapy may be sufficiently effective compared to combination treatment.

Our study is limited by the small number of patients examined. However, the study was designed to optimize the drug screening methods, analysis strategies and to explore how strong the associations between organoid response and clinical response were for predefined DRC metrics. Our results will be validated in our ongoing prospective study, the OPTIC trial, which establishes PDOs for mCRC patients and examines if PDO response is associated with patient response<sup>20</sup>. The OPTIC trial will enable us to further develop organoids as a predictive biomarker for clinical response in mCRC patients by defining threshold values for organoid sensitivity/resistance and analyzing the diagnostic power for different treatment types and lines. Trials, such as the OPTIC, may also offer clarity regarding the feasibility of using PDOs as a biomarker in the clinic by examining if PDO establishment rates and drug screen results are available for most patients in a clinically relevant, short timeframe.

## CONCLUSIONS

Our results demonstrate that organoid response correlates with patient response to commonly given CRC treatment, suggesting that organoids can serve as a predictive biomarker and guide personalized cancer treatment. Associations were seen for all treatment types, including 5-FU monotherapy, irinotecan-based treatment and panitumumab. The association was highest when measuring organoid response using the DRC parameters  $GR_{AUC}$  and  $GR_{max}$ . Interestingly, we observed varying organoid sensitivities to panitumumab for left-sided colon *RAS/BRAF*-wildtype tumours, suggesting that organoids may further aid in guiding anti-EGFR treatment. Our findings will be confirmed for other treatments and prospectively validated in a larger trial, such as the ongoing OPTIC trial, to facilitate the implementation of organoids in clinical treatment.

## REFERENCES

1. Biller L.H., Schrag D. Diagnosis and treatment of metastatic colorectal cancer: A review. *JAMA*. (2021);**325**(7):669–85.
2. Yamazaki K., Nagase M., Tamagawa H., Ueda S., Tamura T., Murata K., et al. Randomized phase III study of bevacizumab plus FOLFIRI and bevacizumab plus mFOLFOX6 as first-line treatment for patients with metastatic colorectal cancer (WJOG4407G). *Ann Oncol*. (2016);**27**(8):1539–46.
3. Venook A.P., Niedzwiecki D., Lenz H.J., Innocenti F., Fruth B., Meyerhardt J.A., et al. Effect of first-line chemotherapy combined with cetuximab or bevacizumab on overall survival in patients with KRAS wild-type advanced or metastatic colorectal cancer: a randomized clinical trial. *JAMA*. (2017);**317**(23):2392–401.
4. Kwakman J.J.M., Simkens L.H.J., van Rooijen J.M., van de Wouw A.J., Ten Tije A.J., Creemers G.J.M., et al. Randomized phase III trial of S-1 versus capecitabine in the first-line treatment of metastatic colorectal cancer: SALTO study by the Dutch Colorectal Cancer Group. *Ann Oncol*. (2017);**28**(6):1288–93.
5. Overman M.J., Lonardi S., Yeung K., Wong M., Lenz H., Gelsomino F., et al. Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair–Deficient/Microsatellite Instability–High Metastatic Colorectal Cancer. *JCO*. (2018);**36**(8):773–9.
6. Sawyers C.L. The cancer biomarker problem. *Nature*. (2008);**452**(7187):548–52.
7. De Souza N. Organoids. *Nat Methods*. (2018);**15**(1):23.
8. Li M., Izpisua Belmonte J.C. Organoids - Preclinical models of human disease. *NEJM*. (2019);**380**(6):569–79.
9. Wensink G.E., Elias S.G., Mullenders J., Koopman M., Boj S.F., Kranenburg O.W., et al. Patient-derived organoids as a predictive biomarker for treatment response in cancer patients. *npj Precis Oncol*. (2021);**5**(30).
10. Van De Wetering M., Francies H.E., Francis J.M., Bounova G., Iorio F., Pronk A., et al. Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell*. (2015);**161**(4):933–45.
11. Sato T., Stange D.E., Ferrante M., Vries R.G.J., Van Es J.H., Van Den Brink S., et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology*. (2011);**141**(5):1762–72.
12. Ganesh K., Wu C., O'Rourke K.P., Szeglin B.C., Zheng Y., Sauvé C.E.G., et al. A rectal cancer organoid platform to study individual responses to chemoradiation. *Nat Med*. (2019);**25**(10):1607–14.
13. Cho Y.W., Min D.W., Kim H.P., An Y., Kim S., Youk J., et al. Patient-derived organoids as a preclinical platform for precision medicine in colorectal cancer. *Mol Oncol*. (2022);**16**(12):2396–412.
14. Ooft S.N., Weeber F., Dijkstra K.K., McLean C.M., Kaing S., van Werkhoven E., et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Sci Transl Med*. (2019);**11**(513):eaay2574.
15. Wang T., Pan W., Zheng H., Zheng H., Wang Z., Li J.J., et al. Accuracy of Using a Patient-Derived Tumor Organoid Culture Model to Predict the Response to Chemotherapy Regimens in Stage IV Colorectal Cancer: A Blinded Study. *Dis Colon Rectum*. (2021);**64**:833–50.

16. Vlachogiannis G., Hedayat S., Vatsiou A., Jamin Y., Fernández-mateos J., Khan K., et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science*. (2018);**359**(6378):920–6.
17. Yao L., Zao X.-L., Pan X.-F., Zhang H.-G., Wang F.-J., Qiao P.-F. Application of tumoroids derived from advanced colorectal cancer patients to predict individual response to chemotherapy. *J Chemother*. (2022);**14**:1–13.
18. Yao Y., Xu X., Yang L., Zhu J., Wan J., Shen L., et al. Patient-Derived Organoids Predict Chemoradiation Responses of Locally Advanced Rectal Cancer. *Cell Stem Cell*. (2019);**26**(1):17–26.
19. Narasimhan V., Wright J.A., Churchill M., Wang T., Lannagan T.R.M., Vrbanac L., et al. Medium-throughput drug screening of patient-derived organoids from colorectal peritoneal metastases to direct personalized therapy. *CCR*. (2020);**26**:3662–70.
20. Netherlands Trial Register [Internet]. Available from: <https://www.trialregister.nl/trial/7305>
21. Sahin I.H., Klostergaard J. BRAF Mutations as Actionable Targets: A Paradigm Shift in the Management of Colorectal Cancer and Novel Avenues. *JCO Oncol Pract*. (2021);**17**(12):723–30.
22. Ritz C., Baty F., Streibig J.C., Gerhard D. Dose-Response Analysis Using R. (2015);1–13.
23. Hafner M., Niepel M., Chung M., Sorger P.K. Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. *Nat Methods*. (2016);**13**(6):521–7.
24. Fallahi-Sichani M., Honarnejad S., Heiser L.M., Gray J.W., Sorger P.K. Metrics other than potency reveal systematic variation in responses to cancer drugs. *Nat Chem Biol*. (2013);**9**(11):708–14.
25. Driehuis E., Kretzschmar K., Clevers H. Establishment of patient-derived cancer organoids for drug-screening applications. *Nat Protoc*. (2020);**15**(10):3380–409.
26. Tejpar S., Stintzing S., Ciardiello F., Tabernero J., Van Cutsem E., Beier F., et al. Prognostic and predictive relevance of primary tumor location in patients with ras wild-type metastatic colorectal cancer retrospective analyses of the CRYSTAL and FIRE-3 trials. *JAMA Oncol*. (2017);**3**(2):194–201.
27. Holch J.W., Ricard I., Stintzing S., Modest D.P., Heinemann V. The relevance of primary tumour location in patients with metastatic colorectal cancer: A meta-analysis of first-line clinical trials. *Eur J Cancer*. (2017);**70**:87–98.
28. Seligmann J.F., Elliott F., Richman S., Hemmings G., Brown S., Jacobs B., et al. Clinical and molecular characteristics and treatment outcomes of advanced right-colon, left-colon and rectal cancers: data from 1180 patients in a phase III trial of panitumumab with an extended biomarker panel. *Ann Oncol*. (2020);**31**(8):1021–9.

## SUPPLEMENTARY MATERIAL

**Supplementary Table 1.** Composition of organoid culture medium

Component	Source (catalogue number)	Concentration
Advanced (DMEM/F12) medium	Gibco (12634-010)	1x (500 ml total)
HEPES Buffer	Lonza (17737E)	10 mM
Penicillin/Streptomycin	Gibco (15070-063)	50 U/ml
GlutaMAX	Gibco (35050-038)	2 mM
N-acetylcysteine (NAC)	Sigma-Aldrich (A9165)	1.25 mM
A83-01	SignalChem (A09-900-05)	500 nM
B27	Invitrogen (17504-044)	1x
Human recombinant EGF	PeptroTech EC Ltd (A10187)	50 ng/ml
Gastrin	Sigma-Aldrich (G9145)	5 nM
Noggin-Fc conditioned medium	U-Protein Express BV (N002)	1% vol/vol
Recombinant human R-spondin-3 conditioned medium	Bio-Techne Ltd (3500-RS/CFMTO)	250 ng/ml
SB202190	Gentaur (A1632)	10 $\mu$ M
Nicotinamide	Merck Life Science (N0636-100G)	10 mM
Primocin	InvivoGene SAS (ANT-PM-2)	50 $\mu$ g/ml
Wnt surrogate	U-Protein Express BV (N001)	0.25 nM

The composition of the organoid culture medium is described.

*Abbreviations:* DMEM/F12 (Dulbecco's Modified Eagle Medium/Ham's F-12), EGF (epidermal growth factor), ml (milliliter), mM (millimolar), nM (nanomolar), ng/ml (nanogram per milliliter), U/ml (Units per milliliter),  $\mu$ g/ml (microgram per milliliter),  $\mu$ M (micromolar), vol/vol (volume/volume).

**Supplementary Table 2.** Chemotherapies and targeted treatments used in drug screens

<b>Drug</b>	<b>Source (catalogue number)</b>	<b>Target</b>	<b>Concentration range (μM)</b>
5-FU	15596885 (UMCU)	DNA synthesis	0.01 – 2700
SN-38 (irinotecan)	Selleck Chemicals GmbH (S4908)	Topoisomerase I	0.00001 – 1.8
5-FU + SN-38 (0.01)	15596885 (UMCU) + Selleck Chemicals GmbH (S4908)	DNA synthesis + Topoisomerase I	Anchor SN-38: 0.01 5-FU: 0.01 – 2700
5-FU + SN-38 (0.0001)	15596885 (UMCU) + Selleck Chemicals GmbH (S4908)	DNA synthesis + Topoisomerase I	Anchor SN-38: 0.0001 5-FU: 0.01 – 2700
Panitumumab	15343561 (UMCU)	EGFR	0.11 – 30,000
Staurosporine	Merck Life Science N. V. (37095)	Protein kinase inhibitor	2
DMSO	VWR (ICNA0219605525)	Solvent	0.2

The compounds used in the organoid drug screens is described. Concentration ranges were optimized to ensure a full drug response curve (DRC) would be obtained. For 5-FU and SN-38 combination treatment, two anchor concentrations of SN-38 were tested with a ten-point concentration range of 5-FU, with highly correlated drug response curve metrics for the two combination conditions. As such, we report the results for the 5-FU + SN-38 combination condition with an anchor concentration of 0.01 μM.

*Abbreviations:* 5-FU (5-Fluorouracil), DMSO (dimethyl sulfoxide), EGFR (epidermal growth factor receptor), SN-38 (active metabolite for irinotecan), μM (micromolar), UMCU (University Medical Center Utrecht).

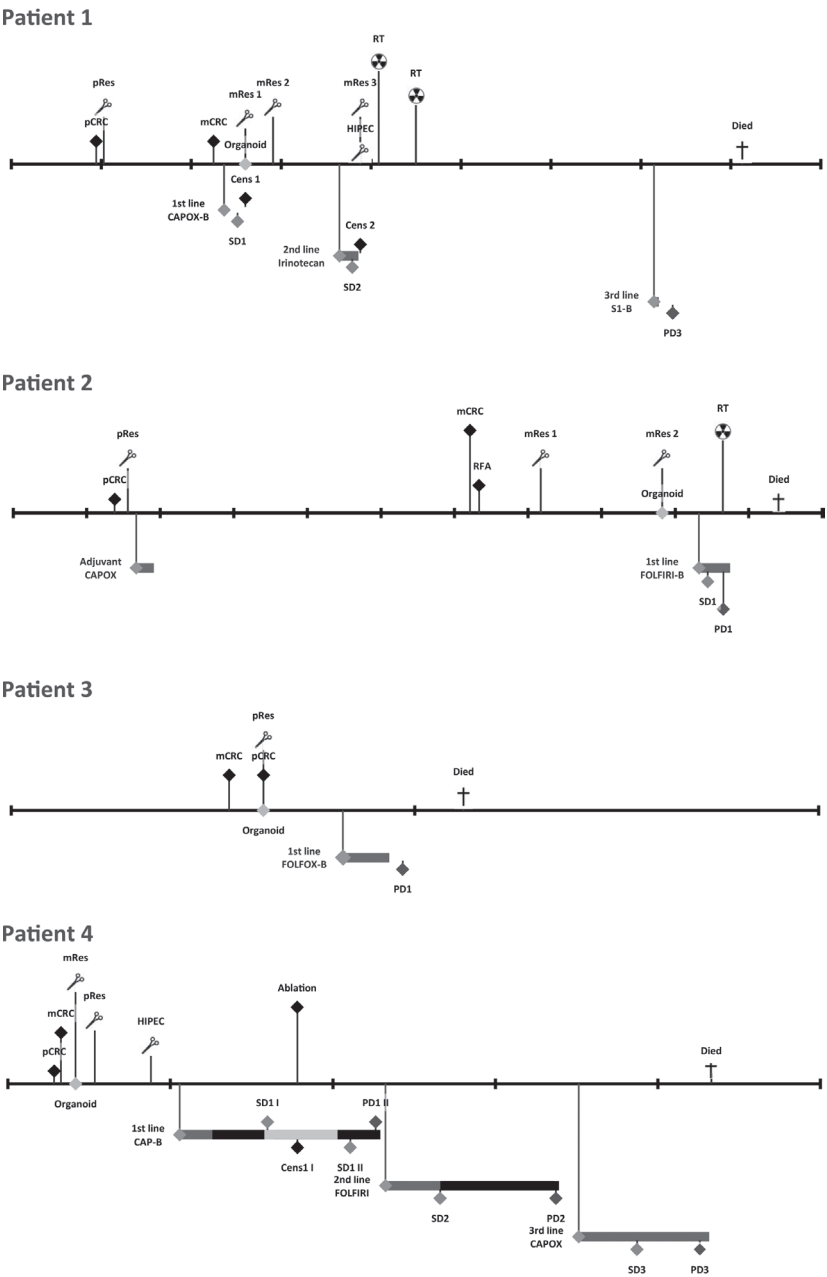
**Supplementary Table 3.** Baseline characteristics of the cohort of patients

	Overall <i>n</i> =18
Age	
Median [Min, Max]	66.5 [45.0, 81.0]
Primary tumour location	
Left-sided (splenic flexure-sigmoid)	7 (38.9%)
Multiple primary tumours (with different sidedness)	1 (5.6%)
Rectum (rectosigmoid/rectal)	5 (27.8%)
Right-sided (coecum-transverse colon)	5 (27.8%)
Primary tumour resection	
No	2 (11.1%)
Yes	16 (88.9%)
Adjuvant chemotherapy	
No	15 (83.3%)
Yes	3 (16.7%)
Liver-only disease	
No	18 (100%)
Mutational status	2 (11.1%)
<i>BRAF</i> -mutant	
<i>RAS</i> -mutant	10 (55.6%)
<i>RAS/BRAF</i> -wildtype	6 (33.3%)
Treatment line(s) prior to PDO establishment	
No	17 (94.4%)
Yes	1 (5.6%)
PDO origin	
Metastatic	15 (83.3%)
Primary	3 (16.7%)
PDO specimen	
Biopsy	4 (22.2%)
Resection	14 (77.8%)

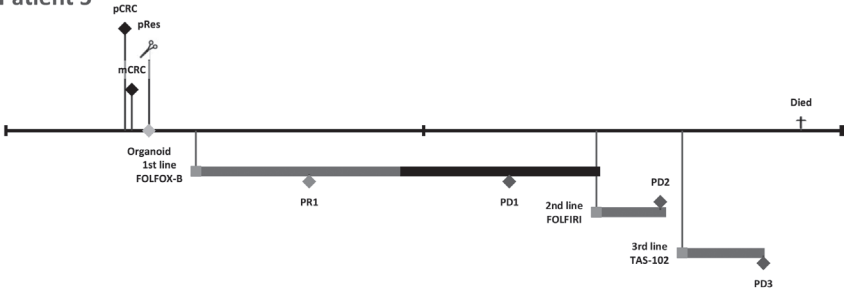
The baseline characteristics of the patient cohort. Treatment line prior to PDO establishment regards a first line systemic treatment line, excluding adjuvant treatment.

*Abbreviations:* max (maximum), min (minimum), n (count), PDO (patient-derived organoid).

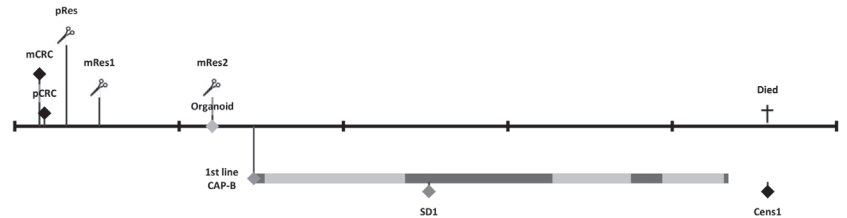
Supplementary Figure 1. Fisher's plots demonstrating patient histories



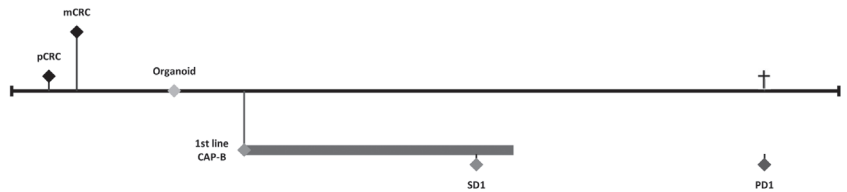
Patient 5



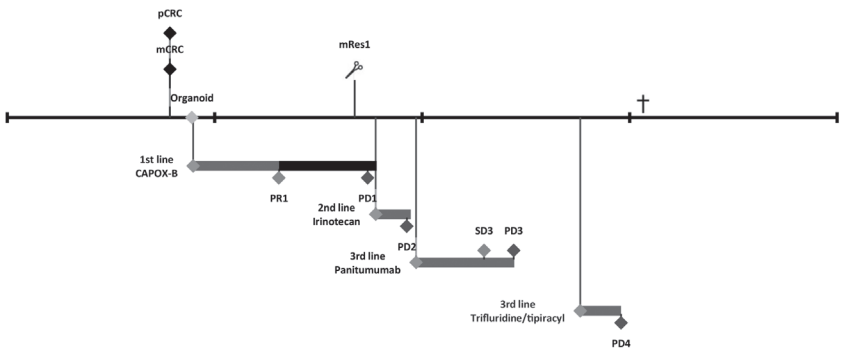
Patient 6



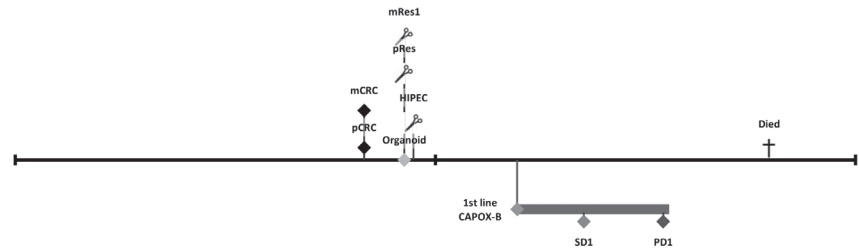
Patient 7



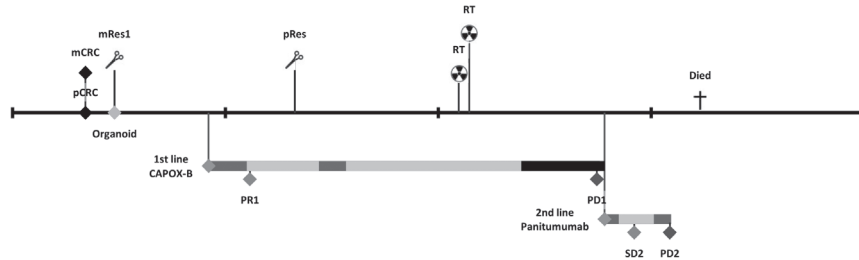
Patient 8



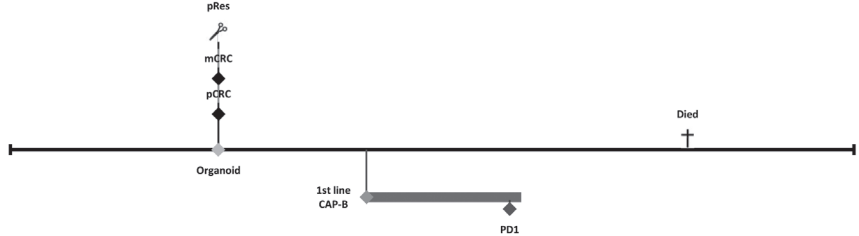
Patient 9



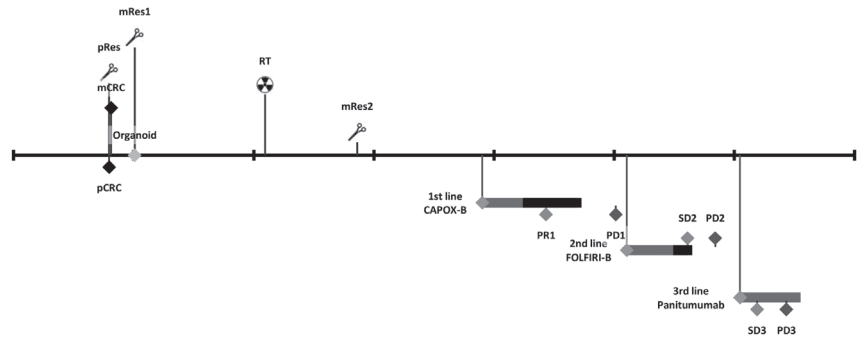
Patient 10



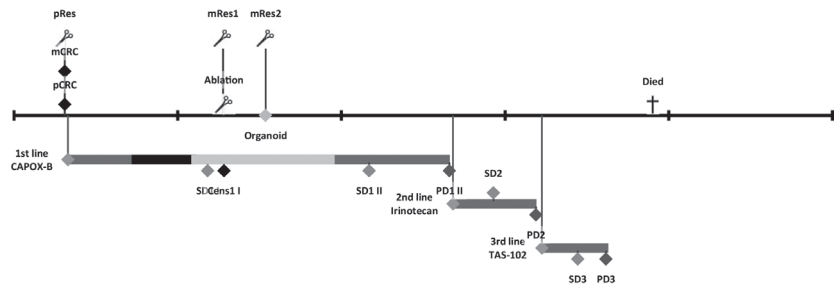
Patient 11



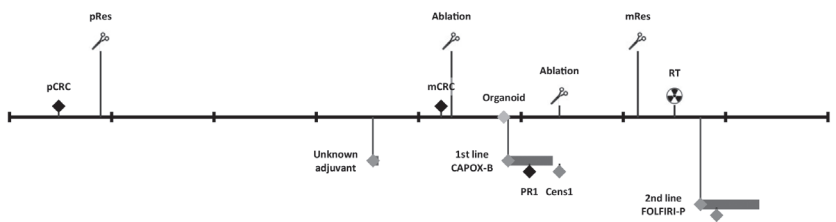
Patient 12



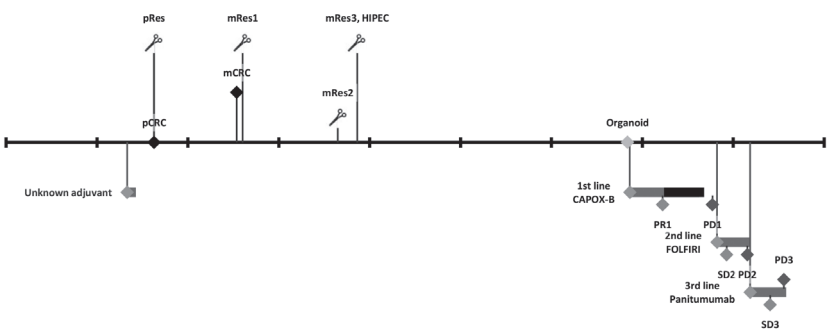
Patient 13



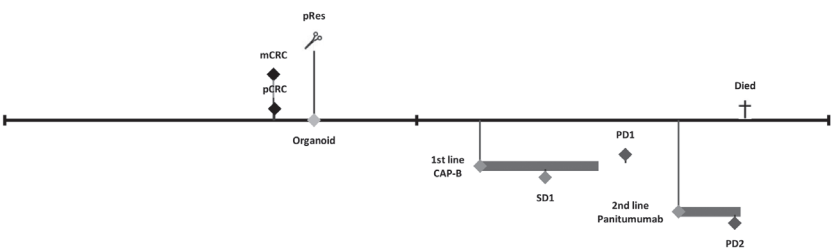
Patient 14



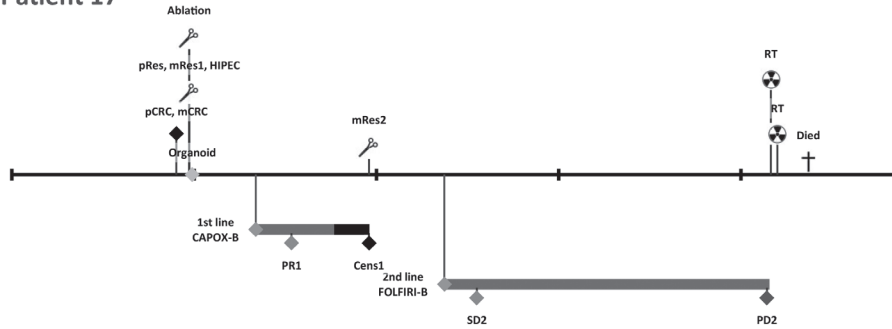
Patient 15



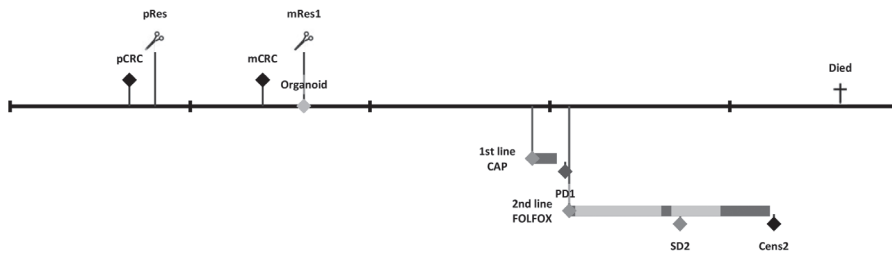
Patient 16



**Patient 17**



**Patient 18**



**Legend**

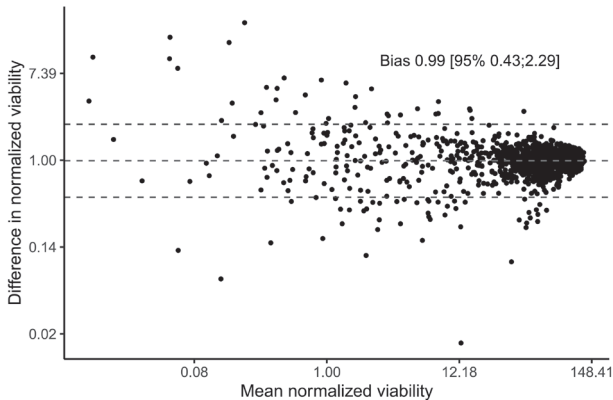
- ◆ Diagnosis primary or metastatic colorectal cancer
- ◆ Sample collected for organoid establishment
- ✂ Local treatment: resection or ablation
- ☒ Local treatment: radiotherapy
- ◆ Start of new treatment line
- Induction treatment
- Maintenance treatment
- Treatment pause
- ◆ Treatment response: stable disease or partial response
- ◆ Treatment response: progressive disease
- † Death

Suppl. Fig. 1. A timeline illustrating the clinical course for individual patients is shown. The distance between two vertical bars indicates one year.

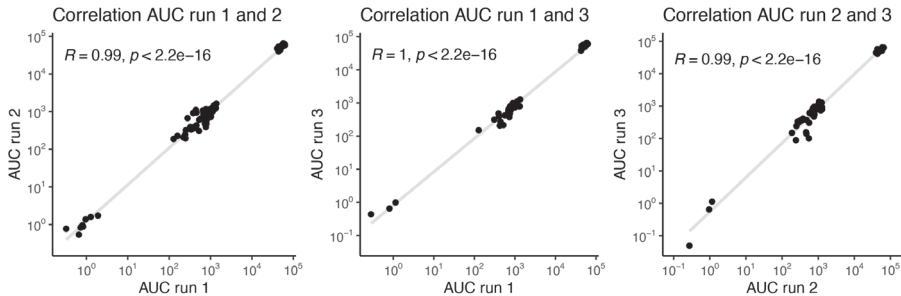
**Abbreviations:** B (bevacizumab), CAP (capecitabine), CAPOX (capecitabine & oxaliplatin), Cens (date of progression was censored), FOLFIRI (5-fluoro-uracil & irinotecan), FOLFOX (5-fluoro-uracil & oxaliplatin), HIPEC (hyperthermic intraperitoneal chemotherapy), mCRC (diagnosis metastatic colorectal cancer), mRes (metastatic resection), pCRC (diagnosis primary colorectal cancer), PD (progressive disease), PR (partial response), pRes (primary tumour resection), SBRT (stereotactic body radiation therapy), SD (stable disease), TT (trifluridine/tipiracil).

**Supplementary Figure 2.** Quality control analysis of the drug screens, illustrating the difference between duplicate assays and Z'-factor

**A.**



**B.**



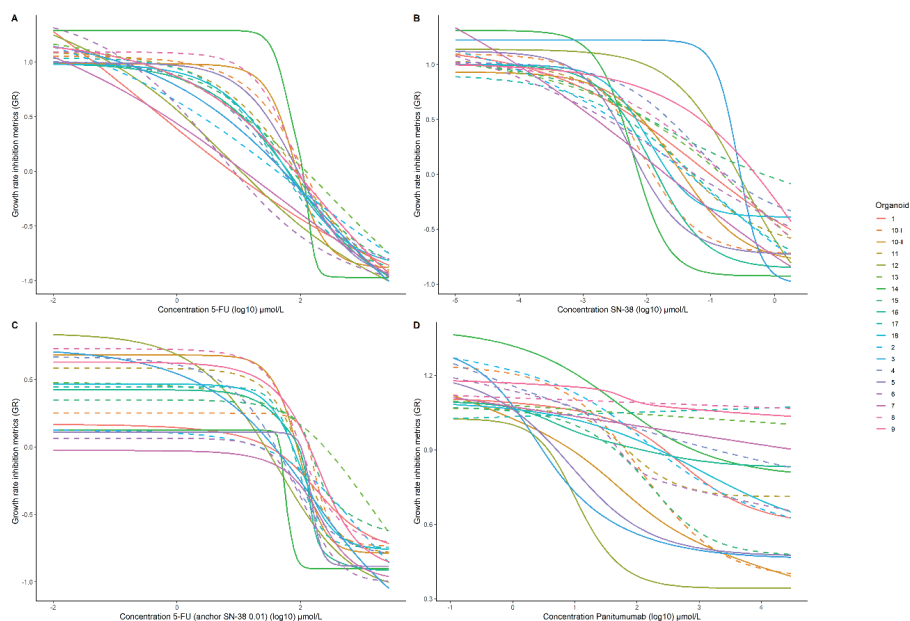
**C.**

Patient	Z'-factor			
	DMSO	PBS	PBS-DMSO	PBS high
16	0,52	0,70	0,73	0,59
05	0,54	0,65	0,71	0,54
03	0,60	0,67	0,69	0,53
13	0,56	0,55	0,49	0,56
09	0,55	0,60	0,48	0,57
12	0,43	0,65	0,43	0,49
06	0,59	0,62	0,57	0,58
01	0,66	0,52	0,55	0,55
18	0,62	0,68	0,60	0,53
04	0,51	0,60	0,57	0,51
10-I	0,61	0,70	0,68	0,71
10-II	0,66	0,64	0,69	0,56
02	0,49	0,66	0,58	0,66
17	0,64	0,62	0,71	0,56
11	0,60	0,56	0,51	0,64
14	0,70	0,69	0,70	0,55
08	0,64	0,60	0,68	0,56
15	0,68	0,66	0,64	0,51
07	0,43	0,74	0,71	0,45

► Suppl. Fig. 2. A) Bland-Altman plot of the mean normalized viability ( $(\text{viability}_{\text{experimental condition}} - \text{viability}_{\text{positive control}}) / (\text{viability}_{\text{negative control}} - \text{viability}_{\text{positive control}})$ ) of the biological replicates on the x-axis versus the difference in normalized viability between replicates on the y-axis. The middle dashed line indicates the average difference observed between the replicates and the outer dashed lines are the standard deviation of the differences. The bias and limits of agreement are indicated on the plot. B) Scatterplots show the correlation between the three different technical replicates of drug screening data. Each data point represents the AUC value for an individual organoid. C) The Z'-factor is shown for different compounds used as the negative control, per organoid line. For one patient (10), two organoids were created.

*Abbreviations:* DMSO (dimethyl sulfoxide), PBS (phosphate buffered saline).

### Supplementary Figure 3. Individual DRCs for each PDO, per treatment



Suppl. Fig. 3. The normalized growth rate inhibition (GR) drug response curves for all organoids are displayed per treatment type. A) 5-FU, B) SN-38, C) 5-FU & SN-38 and D) panitumumab. SN-38 is a metabolite for irinotecan. To differentiate between the organoids, a combination of dashed or solid lines and colours are used.

**Supplementary Table 4.** Drug response curve parameters

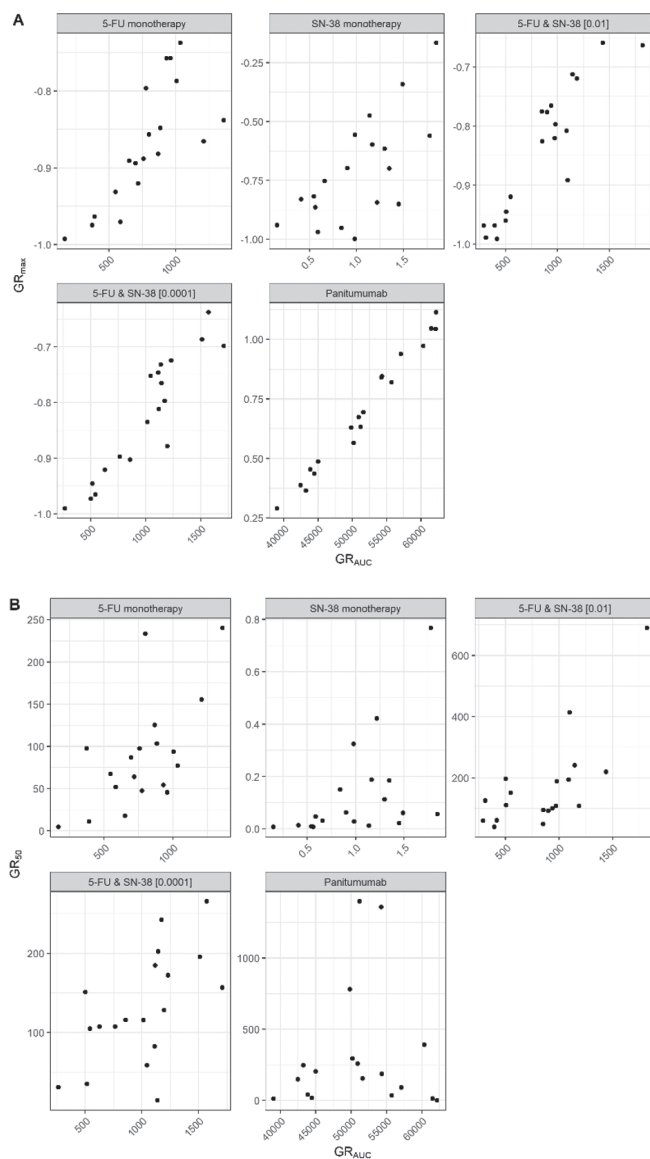
Organoid ID	Treatment	GR <sub>AUC</sub>	GR <sub>max</sub>	GR <sub>50</sub>
1	5-FU & SN-38 (conc. 0.0001)	1137.93	-0.73	14.58
1	5-FU & SN-38 (conc. 0.01)	1145.89	-0.71	241.19
1	5-FU monotherapy	651.96	-0.89	17.68
1	Panitumumab	50188.01	0.57	294.62
1	SN-38 monotherapy	1.30	-0.62	0.11
2	5-FU & SN-38 (conc. 0.0001)	1196.38	-0.88	128.58
2	5-FU & SN-38 (conc. 0.01)	1097.82	-0.89	413.64
2	5-FU monotherapy	1207.28	-0.87	155.44
2	Panitumumab	49838.35	0.63	781.80
2	SN-38 monotherapy	0.84	-0.95	0.15
3	5-FU & SN-38 (conc. 0.0001)	542.42	-0.97	104.88
3	5-FU & SN-38 (conc. 0.01)	418.96	-0.99	60.73
3	5-FU monotherapy	584.81	-0.97	51.84
3	Panitumumab	43853.67	0.45	40.70
3	SN-38 monotherapy	0.98	-1.00	0.32
4	5-FU & SN-38 (conc. 0.0001)	1044.50	-0.75	58.60
4	5-FU & SN-38 (conc. 0.01)	850.37	-0.78	47.89
4	5-FU monotherapy	958.35	-0.76	45.58
4	Panitumumab	55724.49	0.82	34.52
4	SN-38 monotherapy	1.49	-0.34	0.06
5	5-FU & SN-38 (conc. 0.0001)	764.68	-0.90	107.28
5	5-FU & SN-38 (conc. 0.01)	547.58	-0.92	150.61
5	5-FU monotherapy	757.77	-0.89	97.52
5	Panitumumab	44442.92	0.44	18.38
5	SN-38 monotherapy	0.56	-0.86	0.01
6	5-FU & SN-38 (conc. 0.0001)	263.76	-0.99	30.97
6	5-FU & SN-38 (conc. 0.01)	311.50	-0.99	125.42
6	5-FU monotherapy	168.67	-0.99	4.47
6	Panitumumab	50927.42	0.67	258.07
6	SN-38 monotherapy	1.45	-0.85	0.02
7	5-FU & SN-38 (conc. 0.0001)	628.31	-0.92	107.33
7	5-FU & SN-38 (conc. 0.01)	501.23	-0.96	196.81
7	5-FU monotherapy	548.33	-0.93	67.13
7	Panitumumab	57077.60	0.94	91.64
7	SN-38 monotherapy	0.59	-0.97	0.05
8	5-FU & SN-38 (conc. 0.0001)	1512.72	-0.69	195.42
8	5-FU & SN-38 (conc. 0.01)	1185.35	-0.72	108.35
8	5-FU monotherapy	1005.32	-0.79	93.88
8	Panitumumab	62153.01	1.04	0.69
8	SN-38 monotherapy	1.35	-0.70	0.19

Organoid ID	Treatment	GR <sub>AUC</sub>	GR <sub>max</sub>	GR <sub>50</sub>
9	5-FU & SN-38 (conc. 0.0001)	1143.64	-0.76	202.47
9	5-FU & SN-38 (conc. 0.01)	979.72	-0.80	188.10
9	5-FU monotherapy	799.13	-0.86	233.13
9	Panitumumab	61514.37	1.05	14.88
9	SN-38 monotherapy	1.78	-0.56	0.77
10	5-FU & SN-38 (conc. 0.0001)	1118.99	-0.81	184.60
10	5-FU & SN-38 (conc. 0.0001)	1014.03	-0.84	115.74
10	5-FU & SN-38 (conc. 0.01)	1087.75	-0.81	194.65
10	5-FU & SN-38 (conc. 0.01)	853.33	-0.83	94.87
10	5-FU monotherapy	867.97	-0.88	125.20
10	5-FU monotherapy	697.03	-0.89	86.82
10	Panitumumab	43227.19	0.37	246.12
10	Panitumumab	42442.28	0.39	148.19
10	SN-38 monotherapy	0.54	-0.82	0.01
10	SN-38 monotherapy	0.66	-0.75	0.03
11	5-FU & SN-38 (conc. 0.0001)	1173.58	-0.80	242.35
11	5-FU & SN-38 (conc. 0.01)	971.95	-0.82	108.39
11	5-FU monotherapy	884.52	-0.85	103.13
11	Panitumumab	51602.46	0.69	153.43
11	SN-38 monotherapy	0.98	-0.56	0.03
12	5-FU & SN-38 (conc. 0.0001)	515.90	-0.95	35.39
12	5-FU & SN-38 (conc. 0.01)	394.42	-0.97	39.05
12	5-FU monotherapy	392.26	-0.96	11.06
12	Panitumumab	39009.93	0.29	12.89
12	SN-38 monotherapy	1.22	-0.84	0.42
13	5-FU & SN-38 (conc. 0.0001)	1708.98	-0.70	156.80
13	5-FU & SN-38 (conc. 0.01)	1820.44	-0.66	690.25
13	5-FU monotherapy	1358.05	-0.84	240.29
13	Panitumumab	60334.44	0.97	390.42
13	SN-38 monotherapy	1.17	-0.60	0.19
14	5-FU & SN-38 (conc. 0.0001)	501.60	-0.97	151.28
14	5-FU & SN-38 (conc. 0.01)	289.35	-0.97	59.68
14	5-FU monotherapy	374.97	-0.97	97.67
14	Panitumumab	54316.85	0.84	187.00
14	SN-38 monotherapy	0.15	-0.94	0.01
15	5-FU & SN-38 (conc. 0.0001)	1571.38	-0.64	265.52
15	5-FU & SN-38 (conc. 0.01)	1436.70	-0.66	219.51
15	5-FU monotherapy	1034.74	-0.74	77.21
15	Panitumumab	44996.91	0.49	205.04
15	SN-38 monotherapy	1.85	-0.17	0.06
16	5-FU & SN-38 (conc. 0.0001)	857.93	-0.90	116.00

Organoid ID	Treatment	GR <sub>AUC</sub>	GR <sub>max</sub>	GR <sub>50</sub>
16	5-FU & SN-38 (conc. 0.01)	505.16	-0.95	110.42
16	5-FU monotherapy	718.67	-0.92	64.07
16	Panitumumab	54248.57	0.84	1358.25
16	SN-38 monotherapy	0.41	-0.83	0.01
17	5-FU & SN-38 (conc. 0.0001)	1113.59	-0.75	82.68
17	5-FU & SN-38 (conc. 0.01)	898.61	-0.78	92.13
17	5-FU monotherapy	776.63	-0.80	47.42
17	Panitumumab	62215.51	1.11	>30,000
17	SN-38 monotherapy	0.90	-0.70	0.06
18	5-FU & SN-38 (conc. 0.0001)	1231.23	-0.72	172.18
18	5-FU & SN-38 (conc. 0.01)	938.76	-0.77	100.48
18	5-FU monotherapy	930.64	-0.76	54.13
18	Panitumumab	51192.44	0.63	1398.48
18	SN-38 monotherapy	1.14	-0.47	0.01

The parameters extracted from the drug response curves are shown for all organoids and all treatments.

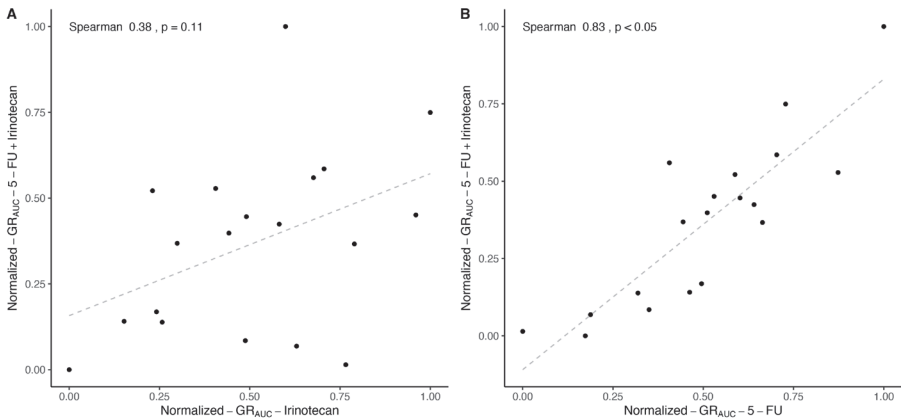
*Abbreviations:* 5-FU (5-fluorouracil), GR<sub>AUC</sub> (area under the curve for a normalized growth rate inhibition drug response curve), GR<sub>max</sub> (maximum normalized growth rate inhibition at the highest concentration) and GR<sub>50</sub> (the concentration of drug at which the normalized growth rate inhibition is 0.5).

**Supplementary Figure 4.** Correlation between different drug response curve parameters

Suppl. Fig. 4. Scatterplots show the correlation between  $GR_{AUC}$  and  $GR_{max}$  (A) and  $GR_{AUC}$  and  $GR_{50}$  (B). One outlier in  $GR_{50}$  was removed from the panitumumab plot.

**Abbreviations:** 5-FU (5-fluorouracil),  $GR_{AUC}$  (area under the curve for a normalized growth rate inhibition drug response curve),  $GR_{max}$  (maximum normalized growth rate inhibition at the highest concentration) and  $GR_{50}$  (the concentration of drug at which the normalized growth rate inhibition is 0.5), SN-38 (active metabolite of irinotecan).

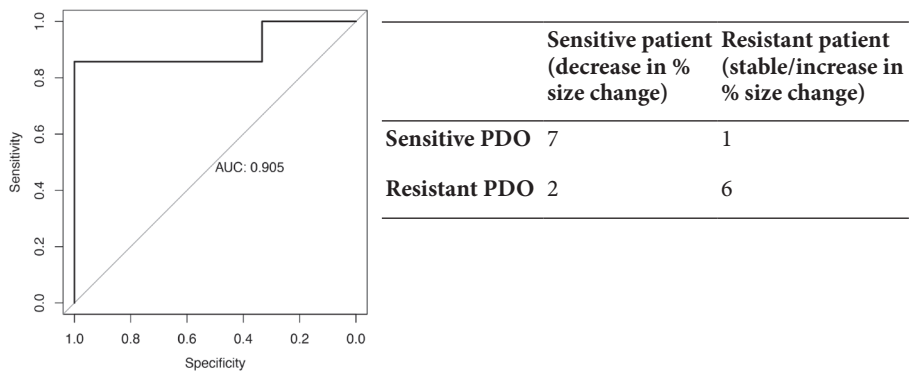
**Supplementary Figure 5.** Correlation between monotherapy versus combination drug response curve parameters



Suppl. Fig. 5. Scatterplots show the correlation between the normalized  $GR_{AUC}$  for 5-FU & irinotecan combination treatment and irinotecan monotherapy (A) and 5-FU (B).

*Abbreviations:* 5-FU (5-fluorouracil),  $GR_{AUC}$  (area under the curve for a normalized growth rate inhibition drug response curve).

**Supplementary Figure 6.** Predicting patient response based on organoid response for chemotherapy



Suppl. Fig. 6. A receiver operating characteristic (ROC) curve of  $GR_{AUC}$  values for chemotherapy regimens tested in PDOs and the corresponding cross table for calculating sensitivity, specificity and predictive values are displayed. The AUC indicates the area under the ROC-curve.

*Abbreviations:* % (percentage), AUC (area under the (Receiver Operator Characteristic) curve),  $GR_{AUC}$  (area under the curve for a normalized growth rate inhibition drug response curve), PDO (patient-derived organoid).





# **PREDICTION MODELS FOR PATIENTS WITH METASTATIC COLORECTAL CANCER**

**PART III**





# EXTERNAL VALIDATION OF TWO ESTABLISHED CLINICAL RISK SCORES PREDICTING OUTCOME AFTER LOCAL TREATMENT OF COLORECTAL LIVER METASTASES IN A NATIONWIDE COHORT

*Cancers*, 2022, 14(10), 2356.

Emerens Wensink<sup>†</sup>  
Karen Bolhuis<sup>†</sup>  
Marloes Elferink  
Marinde Bond  
Willemieke Dijksterhuis  
Remond Fijneman  
Onno Kranenburg  
Inne Borel Rinkes  
Miriam Koopman  
Rutger-Jan Swijnenburg  
Geraldine Vink  
Jeroen Hagendoorn  
Cornelis Punt  
Sjoerd Elias<sup>‡</sup>  
Jeanine Roodhart<sup>‡</sup>

<sup>†</sup> Both authors share equal contribution.

<sup>‡</sup> Both authors share last authorship.

## ABSTRACT

Optimized surgical techniques and systemic therapy have increased the number of patients with colorectal liver metastases (CRLM) eligible for local treatment. To increase postoperative survival, we need to stratify patients to customize therapy. Most clinical risk scores (CRSs) which predict prognosis after CRLM resection were based on the outcome of studies in specialized centers, and this may hamper the generalizability of these CRSs in unselected populations and underrepresented subgroups. We aimed to externally validate two CRSs in a population-based cohort of patients with CRLM. A total of 1105 patients with local treatment of CRLM, diagnosed in 2015/2016, were included from a nationwide population-based database. Survival outcomes were analyzed. The Fong and more recently developed GAME CRS were externally validated, including in pre-specified subgroups ( $\leq 70$ / $>70$  years and with/without perioperative systemic therapy). The three-year DFS was 22.8%, and the median OS in the GAME risk groups (high/moderate/low) was 32.4, 46.7, and 68.1 months, respectively ( $p < 0.005$ ). The median OS for patients with versus without perioperative therapy was 47.6 (95% C.I. 39.8- 56.2) and 54.9 months (95% C.I. 48.8-63.7), respectively ( $p = 0.152$ ), and for below/above 70 years, it was 54.9 (95% C.I. 49.3-64.1) and 44.2 months (95% C.I. 37.1-54.3), respectively ( $p < 0.005$ ). The discriminative ability for OS of Fong CRS was 0.577 (95% C.I. 0.554-0.601), and for GAME, it was 0.596 (95% C.I. 0.572-0.621), and was comparable in the subgroups. In conclusion, both CRSs showed predictive ability in a population-based cohort and in predefined subgroups. However, the limited discriminative ability of these CRSs results in insufficient preoperative risk stratification for clinical decision-making.

## INTRODUCTION

Approximately 30% of patients with colorectal cancer (CRC) develop liver metastases (CRLM)<sup>1</sup>. Currently, local treatment of CRLM (e.g., resection or tumour ablation) offers the only chance for long-term survival, with 5-year overall survival (OS) rates of up to 55%<sup>2-4</sup>. Surgical techniques continue to evolve, with two-stage resections including associating liver partition and portal vein ligation for staged hepatectomy (ALPPS); and laparoscopic liver resections, including minor/major resections, robotic hepatectomy, anatomic resections, parenchymal sparing strategies, and minimally invasive procedures for simultaneous resections of liver metastases and primary CRC<sup>5,6</sup>. Improved surgical procedures, more lenient resection criteria, and optimization of

induction systemic therapy have increased the number of patients with CRLM that are considered technically resectable<sup>7,8</sup>. However, relapse after liver resection occurs in up to 75% of patients<sup>9–11</sup>, and a subgroup of patients have no long-term OS benefit, due to aggressive tumour biology. This underscores the urgent need to improve risk-stratification prior to surgery<sup>12</sup>.

An ideal clinical risk score (CRS) for these patients should identify patients with a high risk of early recurrence after surgery in order to prevent major surgery with associated risk of perioperative morbidity and mortality. Among earlier CRSs for patients with CRLM<sup>2,13,14</sup>, the Fong score—developed in 1999<sup>13</sup>—is still used most frequently to predict prognosis after liver resection<sup>14</sup>. The Fong CRS incorporated lymph node status, CEA value, disease-free interval (DFI), and size and number of liver metastases<sup>13</sup>. However, essential validation efforts of these earlier CRSs are scarce<sup>15–17</sup>, especially in populations receiving modern systemic therapies, improved surgical, and ablative treatment options<sup>2,13,18,19</sup>.

Novel CRSs<sup>14,20–24</sup> have been proposed with their own strengths and limitations, including the modified clinical score (m-CS)<sup>20</sup>, Liverpool score<sup>23</sup>, comprehensive evaluation of relapse risk score (CERR)<sup>22</sup>, alternative clinical score (a-CS)<sup>24</sup>, and the Genetic And Morphological Evaluation (GAME) score<sup>14</sup>. The GAME score incorporates recalibrated tumour markers such as KRAS mutational status, extrahepatic disease presence, and Tumour Burden Score (TBS). The TBS is suggested to better correlate with OS compared to separate information on the number and size of metastases<sup>25</sup>. The GAME score outperformed the Fong score in two single-institution patient cohorts but lacks external validation in more unselected patient cohorts.

Overall, the generalizability of these CRSs to routine care remains questionable. The scores were developed in single and/or specialized liver centers and validated in other specialized centers, potentially not reflecting results in a general population of patients with CRLM<sup>17,26</sup>. Furthermore, important subgroups were underrepresented in the development and validation cohorts such as elderly patients, who represent 50% of the CRC population and who are increasingly offered local liver treatment, as long-term survival can also be achieved in these patients undergoing resection of CRLM<sup>27–29</sup>. Lastly, geographical differences in treatment guidelines might influence cohort characteristics and, therefore, risk score performance. For example, the GAME score was developed and validated in the United States of America, with the majority of

patients receiving perioperative systemic therapy according to local guidelines<sup>30</sup>, while other guidelines do not recommend standard (neo)adjuvant systemic therapy<sup>31,32</sup>.

The aim of this study was to evaluate the generalizability and clinical validity of two CRSs, the widely used Fong score and the more recent GAME score, in a nationwide population-based cohort of patients after local treatment of CRLM. Furthermore, we validated both CRSs in two pre-specified subgroups: with/without modern perioperative systemic therapy and age below/above 70 years.

## **MATERIALS AND METHODS**

### **Population-Based Cohort**

All patients initially diagnosed with CRC between 1 January 2015 and 31 December 2016 and who underwent local treatment (resection and/or local ablation) for CRLM were identified in the Netherlands Cancer Registry (NCR). The NCR is a population-based registry with clinical data of all newly diagnosed cancer patients in the Netherlands, based on notification of newly diagnosed malignancies in the Netherlands by the national automated pathological archive (PALGA<sup>33</sup>) or national registry of hospital discharge. PALGA comprises all patients with histologically confirmed cancer in the Netherlands. Patients with extrahepatic metastases before resection, R2 liver resections, appendix carcinoma, concomitant local liver treatments other than resection or ablation, and inadequate follow-up information were excluded. The research protocol and use of this data was approved by the Netherlands Comprehensive Cancer Organisation (IKNL). Written informed consent was not applicable according to national legislation. The study was performed in accordance with the Declaration of Helsinki.

### **Clinical Data**

Pseudonymized clinical data were retrieved from the NCR and PALGA, including age, sex, American Joint Committee on Cancer (AJCC) tumour status (T-status), nodal status (N-status; N0, N1, and N2), location of primary tumour (left, right, rectum), DFI between detection of primary tumour and metastases, size and number of metastases, serum carcinoembryonic antigen (CEA) level (ug/L) prior to liver resection, type of local treatment, resection margin status (R0 was defined as a microscopically tumour free surgical margin), and RAS/BRAFV600E mutational status. TBS<sup>25</sup> was calculated. A major resection was defined as resection of  $\geq 4$  liver segments<sup>34</sup>, synchronous disease as a DFI of  $\leq 6$  months<sup>35</sup>,

and perioperative systemic therapy as any systemic therapy administered within 100 days before and/or after local treatment of CRLM and initiated prior to progression of disease after resection. No distinction could be made between neo-adjuvant or induction systemic therapy in the NCR data, because intention of treatment was not registered. However, the Dutch guidelines for CRC<sup>31</sup> recommend not to administer perioperative systemic therapy in initially resectable CRLM contrary to the NCCN guidelines<sup>30</sup>. Thus, patients who have received preoperative systemic treatment are assumed to have undergone induction treatment for initially unresectable or potentially resectable CRLM. All assumptions regarding systemic treatment can be found in Supplementary Table 1.

## Overall Survival and Disease-Free Survival

Follow-up data for recurrences were collected from medical records by trained data managers from the IKNL until May 2020, and vital status was obtained by linkage with the municipal population registry on 31 January 2021. OS was defined as the date of first CRLM resection/ablation till the date of vital status. Disease-free survival (DFS) was defined as the date of first CRLM resection/ablation till date of a DFS event, which was defined as recurrence of disease or death, whichever occurred first, or censored on last date of DFS. If the follow-up for recurrences was shorter than the follow-up for vital status, all vital status follow-up beyond the last follow-up for recurrences was discarded for assessment of DFS. All survival assumptions are included in Supplementary Table 1.

## RAS and BRAFV600E Mutational Status

Tumour KRAS (codons 12, 13, 61, 117, and 146), NRAS (codons 12, 13, and 61) and BRAF V600E mutational status, as ascertained during routine clinical care, were retrieved from the NCR and PALGA<sup>33</sup>. As mutational status is generally only determined clinically if there is an indication for (palliative) systemic treatment, this information was not available for all patients. To further complement the RAS/BRAFV600E mutational status of the cohort, we aimed to sequence >170 available tumour tissues (the first 171 available of 250 requested) by Sequenom Massarray<sup>36</sup>. We specifically selected these 250 patients, as they had the lowest predicted chance of having a clinically assessed mutational status according to their clinicopathological profile (based on a logistic regression propensity score for mutational status with 16 clinicopathological variables). We used this strategy to improve the chance of successful multiple imputation and of accommodating the missing at random assumption (see below).

## Statistical Analysis and Handling of Missing Data

The study population was described using standard descriptive statistics, overall, according to systemic treatment, and according to age, using median values and interquartile interval (IQI) for continuous variables and frequencies and percentages for categorical variables. Differences between systemic treatment and age groups were statistically tested by the Mann–Whitney U test or the Fisher’s Exact Test. All reported  $p$ -values are two-sided and  $p < 0.05$  was considered statistically significant.

To handle missing data in the context of survival analysis, we performed multiple imputation by using a substantive model compatible fully conditional specification (SMC-FCS) approach<sup>37</sup>, assuming missingness at random. The substantive model was a Cox proportional hazards model for OS which contained the following variables: T-status, N-status, KRAS mutational status, number and size of liver metastases, CEA, systemic perioperative treatment type, sidedness of the primary tumour, age, DFI, R-status, GAME CRS, Fong CRS, and TBS (with the last 3 being passively imputed in the model). We generated 53 imputed datasets based on the percentage of patients with at least one missing key variable.

Kaplan–Meier (KM) curves were created for OS and DFS. Using the multiple imputed dataset, pooled statistics were obtained by using Rubin’s rules, including number at risk for given time points, log-rank subgroup comparison, and survival estimates with confidence intervals (using log–log transformation prior to pooling for the latter two)<sup>38,39</sup>.

## External Validation of CRSs

The GAME<sup>14</sup> and Fong score<sup>13</sup> were externally validated following the TRIPOD guidelines sections pertinent to external validation studies<sup>40</sup>. Predictive performances were assessed by measures of calibration and discrimination. Calibration was evaluated by digitizing the originally published KM curves of scores by WebPlotDigitizer version 4.4<sup>41</sup> and plotted together with the observed KM curves of the NCR cohort. Discrimination was calculated by Harrell’s concordance index (C-index) across each imputed dataset and pooled by using Rubin’s rules. The C-index reflects the ability of the model to differentiate between patients who do and do not experience an event, with 0.5 representing a model without any discriminatory ability beyond chance and 1 perfect discrimination<sup>42</sup>.

Patients were assigned to low, moderate, or high CRS risk categories, as described previously<sup>14</sup>: low risk, 0–1 points; moderate risk, 2–3 points, and high risk, 4 or more points, with similar allocation for the GAME and Fong CRS points.

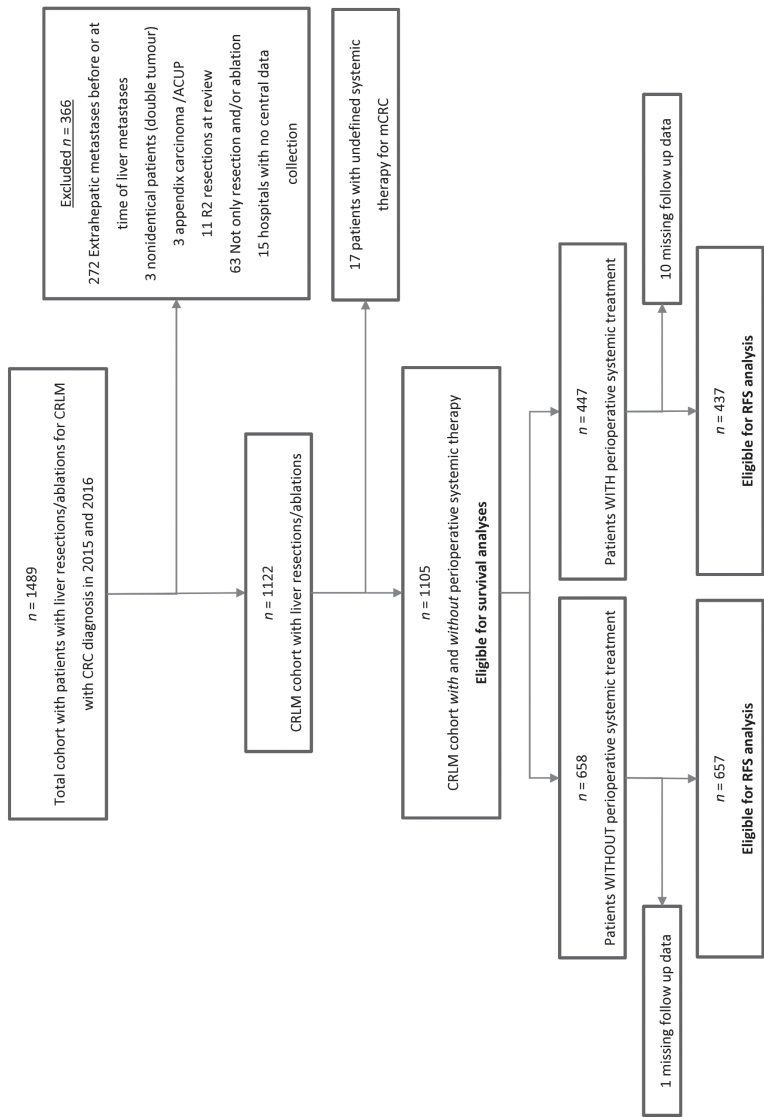
To analyze the overlap in risk groups following the two CRSs, a contingency table and heatmap were made. External validation was repeated for the following subgroups: perioperative systemic therapy (yes/no) and age ( $\leq 70$ / $>70$  years). An analysis was performed in IBM SPSS Statistics (Version 26) and R (Version 4.0.3 for Windows) with the mice (3.13.0), smcfcs (1.5.0), survival (3.2-7), and rms (6.2-0) packages.

## RESULTS

### Patient Characteristics

A total of 1105 patients fulfilled the eligibility criteria (1105/1489) (Figure 1). The cohort comprised 447 (40%) patients with and 658 (60%) patients without perioperative systemic therapy and 759 (69%) patients  $\leq 70$  and 346 (31%) patients  $>70$  years. Among patients with perioperative systemic treatment, 334 (75%) received preoperative-only, 54 (12%) postoperative-only, and 59 (13%) received pre- and postoperative systemic treatment. The patient characteristics are displayed in Table 1. The median age of patients was 66 years, with 690 (62%) males, and 823 (75%) patients had synchronous disease. (Table 1). Patients were treated in a total of 39 hospitals, with 45% of patients treated in academic, 44% in teaching, and 11% in regional hospitals.

**Figure 1.** Flowchart of population-based NCR patients with local liver treatment for CRLM included in the study.



Abbreviations: ACUP (adenocarcinoma with unknown primary), CRLM (colorectal liver metastases), mCRC (metastatic colorectal cancer), RFS (recurrence-free survival).

**Table 1.** Characteristics of total NCR cohort and patients with and without systemic therapy and below or above 70 years

	NCR Cohort (n = 1105)	Patients without Systemic Therapy (n = 658)	Patients with Systemic Therapy (n = 447)	p-value	Patients ≤ 70 Years (n = 759)	Patients > 70 Years (n = 346)	p-value
Age Median (IQI)	66 (59–72)	68 (61–74)	63 (56–70)	< 0.001	62 (56–66)	75 (72–78)	<0.001
Sex				0.60			<0.008
Male	690 (62)	415 (63)	275 (62)		454 (60)	236 (68)	
Female	415 (38)	243 (37)	172 (39)		305 (40)	110 (32)	
Side primary tumour				0.92			0.09
Right	261 (24)	154 (23)	107 (24)		167 (22)	94 (27)	
Left	473 (43)	285 (43)	188 (42)		324 (43)	149 (43)	
Rectum	371 (34)	219 (33)	152 (34)		268 (35)	103 (30)	
Chemoradiotherapy primary tumour				< 0.001			0.67
No	977 (88)	556 (85)	421 (94)		669 (88)	308 (89)	
Yes	128 (12)	102 (15)	26 (6)		90 (12)	38 (11)	
T-status primary tumour				0.01			0.97
1	27 (3)	21 (3)	6 (1)		19 (3)	8 (2)	
2	128 (12)	86 (13)	42 (10)		89 (12)	39 (11)	
3	757 (69)	455 (69)	302 (69)		516 (69)	241 (70)	
4	185 (17)	96 (15)	89 (20)		129 (17)	56 (16)	
Missing	8 (-)	0 (-)	8 (-)		6 (-)	2 (-)	
Nodal status primary tumour				0.22			0.71
N0	408 (37)	257 (39)	151 (34)		277 (37)	131 (38)	
N1	389 (35)	224 (34)	165 (37)		265 (35)	124 (36)	
N2	306(28)	177 (27)	129 (29)		216 (29)	90 (26)	
Missing	2 (-)	0 (-)	2 (-)		1 (-)	1 (-)	
Stage of disease at diagnosis				< 0.001			0.40
I	25 (2)	20 (3)	5 (1)		16 (2)	9 (3)	
II	102 (9)	87 (13)	15 (3)		65 (9)	37 (11)	
III	187 (17)	162 (25)	25 (6)		123 (16)	64 (19)	
IV	791 (72)	389 (59)	402 (90)		555 (73)	236 (68)	
Differentiation grade of CRC				0.12			0.77
Low	17 (2)	7 (1)	10 (3)		13 (2)	4 (1)	
Intermediate	936 (92)	577 (93)	359 (90)		642 (92)	294 (92)	
High	68 (7)	37 (6)	31 (8)		46 (7)	22 (7)	
Missing	84 (-)	37 (-)	47 (11)		58 (-)	26 (-)	

	<b>NCR Cohort (n = 1105)</b>	<b>Patients without Systemic Therapy (n = 658)</b>	<b>Patients with Systemic Therapy (n = 447)</b>	<b>p-value</b>	<b>Patients ≤ 70 Years (n = 759)</b>	<b>Patients &gt; 70 Years (n = 346)</b>	<b>p-value</b>
Time to metastases				< 0.001			0.20
Synchronous	823 (75)	412 (63)	411 (92)		574 (76)	249 (72)	
Metachronous	282 (25)	246 (37)	36 (8)		185 (24)	97 (28)	
Number of liver metastases				< 0.001			<0.001
Median (IQI)	2 (1–4)	1 (1–3)	3 (2–6)		2 (1–4)	2 (1–3)	
Missing	42	19	23		33	9	
CEA level				< 0.001			0.90
Median (IQI)	9 (3.4–36)	6.3 (3.0–21)	14 (4.4–74)		18 (4.0–413)	17 (4.7–168)	
Unknown	231	180	51		160	71	
Size largest liver metastasis, mm				0.002			0.30
Median (IQI)	25 (16–36)	23 (16–35)	27 (16–45)		25 (15–45)	26 (18–42)	
Missing	86	45	41		58	28	
Type of surgery				< 0.001			0.34
Wedge/segment resection only	589 (53)	416 (63)	173 (39)		400 (53)	189 (55)	
Local ablative therapy only	95 (9)	63 (9)	32 (7)		59 (8)	36 (10)	
Wedge/segment and local ablative therapy	189 (17)	90 (14)	99 (22)		134 (18)	55 (16)	
Hemihepatectomy with/without ablation/wedge (major resection)	232 (21)	89 (14)	143 (32)		166 (22)	66 (19)	
One- or two-stage				< 0.001			0.84
1-stage	1042 (94)	643 (98)	399 (89)		715 (94)	327 (95)	
2-stage	63 (6)	15 (2)	48 (11)		44 (6)	19 (6)	
R-status				0.07			0.37
R0	866 (78)	521 (79)	345 (77)		598 (79)	268 (78)	
R1	143 (13)	74 (11)	69 (15)		101 (13)	42 (12)	
Unknown because RFA/MWA	96 (9)	63 (10)	33 (7)		60 (8)	36 (10)	

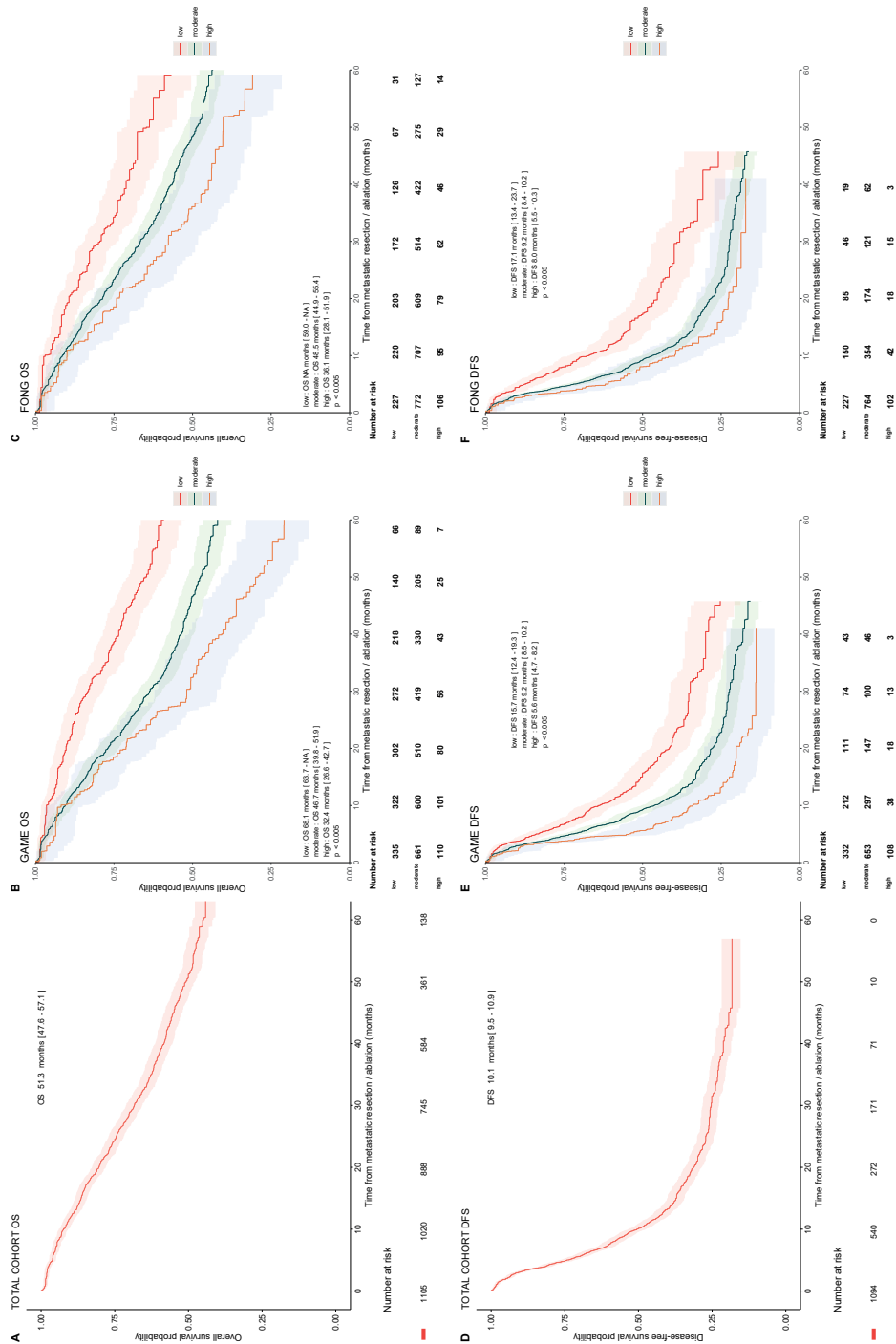
	NCR Cohort ( <i>n</i> = 1105)	Patients without Systemic Therapy ( <i>n</i> = 658)	Patients with Systemic Therapy ( <i>n</i> = 447)	<i>p</i> -value	Patients ≤ 70 Years ( <i>n</i> = 759)	Patients > 70 Years ( <i>n</i> = 346)	<i>p</i> -value
Tumour mutational status				0.36			0.93
<i>RAS</i> mutation	362 (51)	221 (53)	141 (48)		247 (50)	115 (52)	
<i>BRAF</i> <sup>V600E</sup> mutation	19 (3)	10 (2)	9 (3)		13 (3)	6 (3)	
<i>RAS</i> and <i>BRAF</i> <sup>V600E</sup> wildtype	335 (47)	188 (45)	147 (50)		233 (47)	102 (46)	
Missing ( <i>RAS</i> and/or <i>BRAF</i> status)	389 (-)	239 (-)	150 (-)		266 (-)	123 (-)	

*Abbreviations:* CEA (Carcinoembryonic Antigen), CRC (colorectal cancer), IQI (Interquartile range), MMR (mismatch repair), NCR (Netherlands cancer registry).

## Follow-up and OS and DFS Outcomes in Total Cohort

The median follow-up for OS and DFS was 53.7 and 35.0 months, with 556 (50%) and 807 (73%) documented events, respectively. The median OS was 51.3 months (95% C.I. 47.6-57.1), and the median DFS was 10.1 months (95% C.I. 9.5-10.9, Figure 2). One-, three-, and five-year OS rates were 89.9% (95% C.I. 88.2-91.7), 61.8% (95% C.I. 59.0-64.8), and 44.9% (95% C.I. 41.6-48.4), whereas the one- and three-year DFS rates were 43.1% (95% C.I. 40.2-46.1) and 22.8% (95% C.I. 20.2-25.8).

Figure 2. Overall survival and disease-free survival in cohort and subgroups



► Fig. 2. Kaplan–Meier analysis showing OS and DFS curves and 95% confidence intervals of the total cohort and for the risk categories following the GAME and Fong scores. OS for total cohort (A), and OS for GAME CRS risk groups (B), OS for Fong CRS risk groups (C). DFS for total cohort (D), DFS for GAME CRS risk categories (E), and DFS for Fong CRS risk categories (F).

## External Validation of GAME and Fong CRSs in Total Cohort

The study characteristics of the development cohorts of the GAME and Fong CRSs were compared to the NCR validation cohort (Table 2). The percentage of patients with adjuvant systemic therapy was 71% in the GAME cohort compared to 6% in our NCR cohort; the percentage was not reported for the Fong cohort. In the development cohort of GAME CRS, patients with extrahepatic disease were included, while these patients were excluded in the Fong cohort and the NCR cohort.

**Table 2.** Characteristics of original Fong and GAME CRS cohorts compared to Dutch NCR cohort used for external validation

	GAME	Fong	NCR
Number of patients (design/validation)	502/747	1001/-	-/1105
Country (design/validation)	USA/USA	USA/-	-/Dutch
Study design	Single center	Single center	Nation-wide multicenter
Patients with liver-only metastases, %	90	100	100
Handling of missing data	Patients excluded with <i>KRAS</i> status missing	NR	No patients excluded based on missing data
Available mutation status	<i>KRAS</i> codon 12, 13, and 61	-	<i>RAS/BRAF</i>
Primary endpoint	OS	OS	OS
Preoperative systemic therapy, %	67	NR	55
Adjuvant systemic therapy, %	71	NR	6
DFI < 12 months, %	74	49	84
Factors included in CRS, (points)	Nodal status (1) CEA > 20 (1) TBS < 9 (1) TBS ≤ 9 (2) <i>KRAS</i> mutation (1) Extrahepatic disease (2)	Nodal status (1) CEA > 200 (1) DFI < 1year (1) > 1 Liver tumour (1) Largest tumour > 5 cm (1)	-

**Abbreviations:** CEA (carcinoembryonic antigen), cm (centimeters), CRS (clinical risk score), DFI (disease-free interval), GAME (genetic and morphological evaluation score), NCR (Netherlands cancer registry), NR (not reported), OS (overall survival), TBS (tumour burden score), USA (United States of America).

The OS and DFS of the high, moderate, and low GAME and Fong risk groups are presented in Figure 2. The OS and DFS gradually decrease per point increase for both the GAME and Fong score (Supplementary Figure 1).

By analyzing the calibration of the CRSs, we see that the original survival curves of low- and high-risk GAME groups overlapped well with the corresponding curves in our validation cohort. The GAME moderate-risk group, however, showed a shorter median OS compared to the development cohort, 46.7 versus 60 months (Supplementary Figure 2).

Overall, the discriminative ability of the GAME versus the Fong score, as measured by the Harrell's C-index for OS, was weak, 0.596 (95% C.I. 0.572-0.621) versus 0.577 (95% C.I. 0.554-0.601), respectively. The C-indexes of OS and DFS and the pooled survival estimates per risk group and per given time-point are depicted in Table 3.

**Table 3.** Pooled Harrell's concordance index with 95% confidence intervals for 1-, 3-, and 5-year overall survival and disease-free survival outcomes for GAME and Fong risk scores and survival estimates at 1-, 3-, and 5 years for low-, moderate-, and high-risk groups according to GAME and Fong prediction model

	GAME Score	Survival Estimates GAME Risk Categories			Fong Score	Survival Estimates Fong Risk Categories		
	C-Index [95% CI]	Low (%)	Moderate (%)	High (%)	C-Index [95% CI]	Low (%)	Moderate (%)	High (%)
<b>OS</b>								
1-year	0.583 [0.531–0.636]	94	88	86	0.570 [0.521–0.619]	95	89	87
3-year	0.600 [0.573–0.627]	77	57	47	0.578 [0.552–0.604]	74	60	50
5-year	0.597 [0.573–0.621]	50	42	21	0.577 [0.554–0.601]	57	40	31
<b>DFS</b>								
1-year	0.585 [0.561–0.608]	57	39	27	0.586 [0.564–0.608]	60	39	34
3-year	0.579 [0.557–0.600]	30	21	14	0.581 [0.561–0.602]	32	20	17

*Abbreviations:* C-index (concordance index) , DFS (disease-free survival), OS (overall survival).

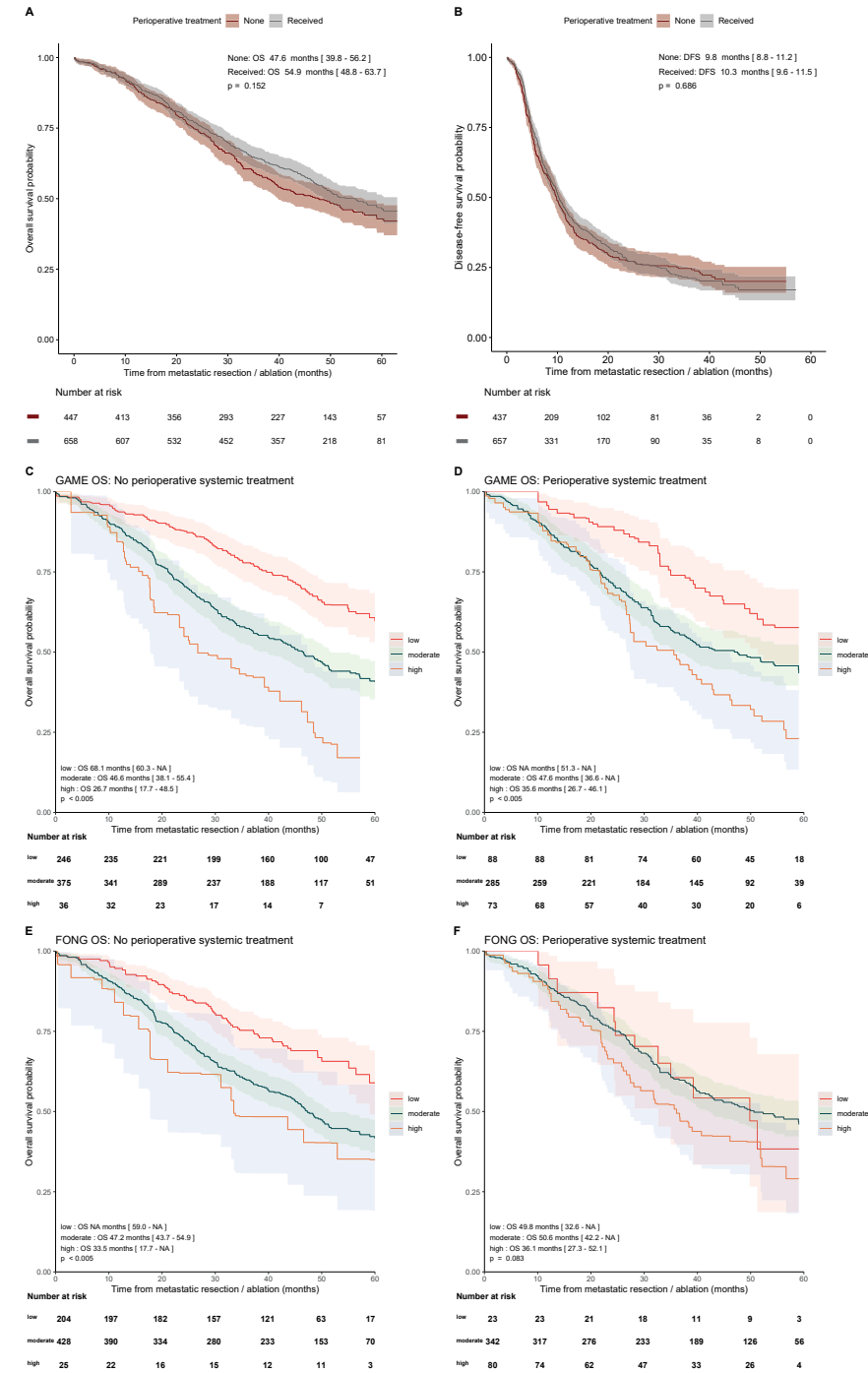
In a head-to-head comparison of the GAME and Fong CRSs, 730 patients (66.0%) were categorized in the same risk group in both prediction models. Only three patients (0.3%) showed major discordance (categorized as GAME high risk and Fong low risk). The frequency distributions among the Fong/GAME combination risk categories and corresponding survival curves are shown in Supplementary Figure 3.

## External Validation of GAME and Fong CRSs in Pre-Specified Subgroups

### With and without Perioperative Systemic Therapy

Although prognostic patient characteristics were unfavorable for patients with perioperative systemic therapy (Table 1), comparable survival outcomes were found in patients with and without perioperative systemic treatment, with a median OS of 47.6 (95% C.I. 39.8-56.2) and 54.9 months (95% C.I. 48.8-63.7;  $p = 0.152$ ) and median DFS of 9.8 (95% C.I. 8.8-11.2) versus 10.3 months (95% C.I. 9.6-11.5;  $p = 0.686$ ), respectively. GAME high-risk patients with perioperative systemic therapy had a longer median OS of 35.6 (95% C.I. 26.7-46.1) compared to patients without systemic therapy (median OS 26.7 months, 95% C.I. 17.7-48.5) (Figure 3) and a longer median DFS of 5.9 months (95% C.I. 4.8-10.9) versus 4.6 months (95% C.I. 3.9-10.0) (Supplementary Figure 4). A survival advantage for patients receiving perioperative systemic therapy was not evident in the low- and moderate-risk groups (Supplementary Figure 4). The GAME C-index for patients with and without peri-operative systemic therapy for OS was 0.590 (95% C.I. 0.554-0.626) versus 0.602 (95% C.I. 0.569-0.635), and the Fong C-index was 0.556 (95% C.I. 0.519-0.594) versus 0.593 (95% C.I. 0.563-0.624), respectively (Supplementary Table 2).

**Figure 3.** Kaplan–Meier analysis showing OS and DFS curves in patients with and without perioperative systemic therapy for the GAME and Fong risk categories

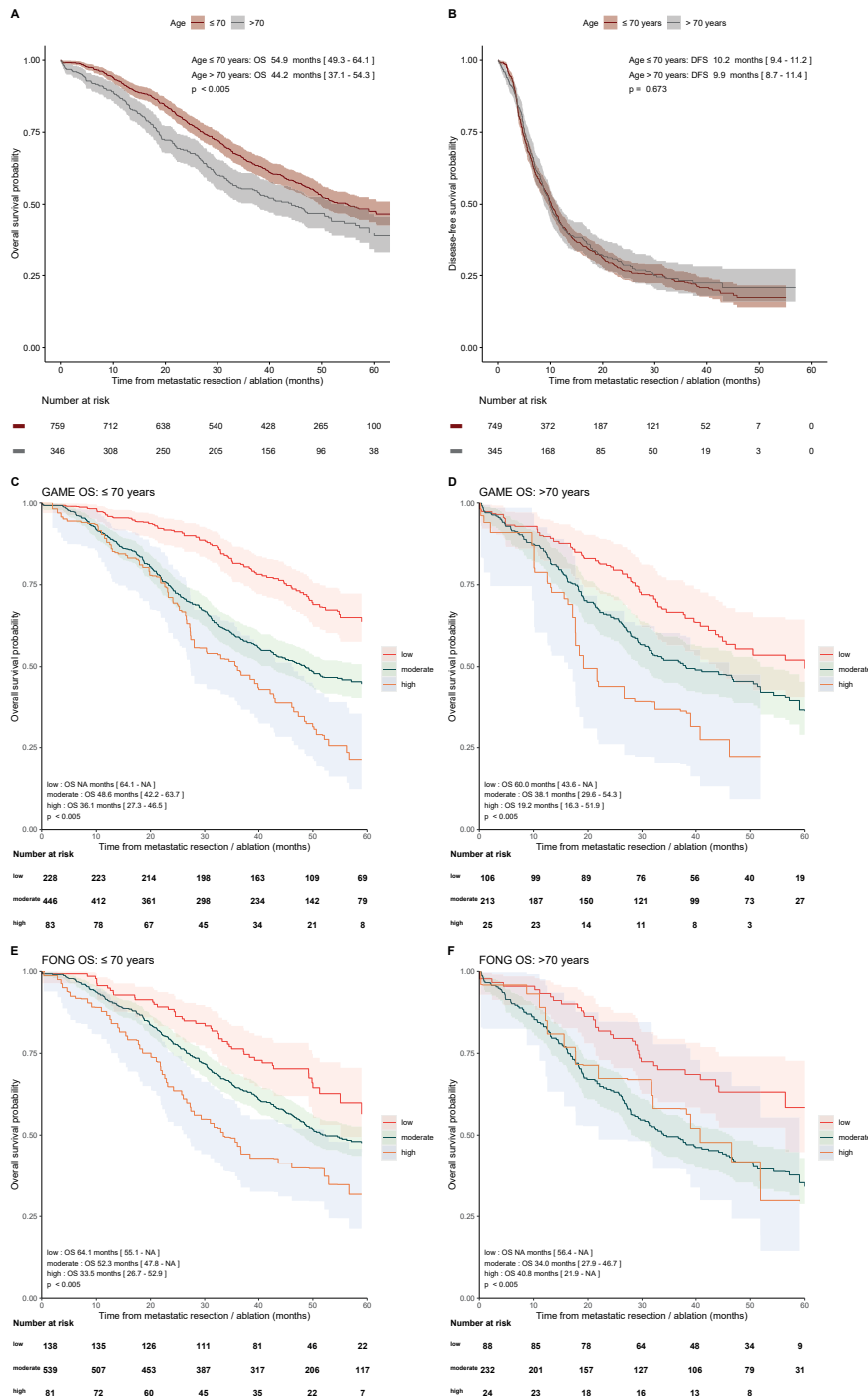


► Fig 3. (A) OS and (B) DFS in patients with and without perioperative systemic therapy. OS outcomes of the GAME risk categories were analyzed in the subgroup without (C) and with perioperative systemic therapy (D) and OS outcomes of the Fong risk categories in subgroups of patients without (E) and with perioperative systemic therapy (F).

### **Age $\leq$ 70 Years and $>$ 70 Years**

The median OS of 54.9 months (95 % CI 49.3–64.1) was higher in patients  $\leq$ 70 years compared to 44.2 months (95% C.I. 37.1–4.3) in patients  $>$ 70 years ( $p < 0.005$ ). The median DFS was similar for 10.2 months (95% C.I. 9.4–11.2) versus 9.9 months (95% C.I. 8.7–11.4;  $p = 0.673$ ) (Figure 4). The discriminative ability for OS of GAME CRS and Fong CRS was comparable in both age groups, with GAME C-indexes of 0.613 (95% C.I. 0.584–0.642) and 0.575 (95% C.I. 0.531–0.618) and Fong C-indexes of 0.583 (95% C.I. 0.554–0.612) and 0.589 (95% C.I. 0.548–0.630), respectively, for below/above 70 years. The C-indexes for one-, three-, and five-year OS and DFS of GAME versus Fong in predefined subgroups are shown in Supplementary Table 3.

**Figure 4.** Kaplan–Meier analysis showing OS and DFS curves in patients with age  $\leq 70$  years and  $>70$  years for the GAME and Fong risk categories



► Fig 4. (A) OS and (B) DFS in patients with age  $\leq 70$  years and  $>70$  years. Subsequently, the OS outcomes of the GAME risk categories were analyzed in the subgroup  $\leq 70$  years (C) and  $>70$  years (D) and of the Fong risk categories in subgroups of patients  $\leq 70$  years (E) and  $>70$  years (F).

## DISCUSSION

In this study, we externally validated and compared two established CRSs, the GAME and Fong score, for their ability to predict OS and DFS after resection of CRLM in the modern era in a real-life population-based cohort and in two pre-specified subgroups. Both CRSs showed predictive ability with a better performance of the GAME as compared to the traditional Fong CRS. The external validation in subgroups of both CRSs showed a comparable performance in patients with and without perioperative systemic therapy and in patients  $\leq 70$  and  $>70$  years. However, the overall predictive performance remained suboptimal, with a high prognostic uncertainty which limits its utility in clinical decision-making.

The GAME score was originally validated in a cohort of patients from specialized institutes, while the Fong score was not validated in the original paper. This could hamper their generalizability to real-life patients. In our real-life cohort, we found a similar C-index for the GAME and Fong score for OS as compared to the C-indexes published by Margonis *et al.*<sup>14</sup>. In our cohort, the GAME score outperformed the traditional Fong score. Both CRSs show discriminatory ability, but since C-indexes are 0.6 at most, a significant level of prognostic uncertainty remains. Furthermore, 25% of patients identified as “high-risk” according to the GAME score did achieve long-term survival, which exceeded five years, and this rate was even higher in the Fong high-risk group. This signifies that, although these CRSs might be used for risk counselling and managing expectations of patients, they cannot be used for clinical decision-making to select high-risk patients for whom surgery should be avoided or low-risk patient for whom extensive surgery may be justified.

To improve the prognostic performance of a CRS, categorizing variables should be avoided, and simplification of the CRS by a point system or classification in risk groups is not always desirable. While this strategy is performed to gain usability, it also results in the loss of information. One way to ensure model usability, while avoiding simplification, is to use a web calculator, along with a prediction model, which could be incorporated into electronic patient management systems for clinicians and patients<sup>24</sup>.

Evolving molecular research results in newly recognized tumour biomolecular prognostic markers and shows the heterogeneity of CRLM. The GAME CRS incorporated KRAS

codon 12, 13, and 61 only. However, BRAFV600E mutation is recognized to be a strong prognostic factor, as well which negatively influences post-resection survival outcomes. Other molecular markers are proposed as prognostic markers too, such as mutations in the SMAD family, TP53, and PIK3CA. In future practice, by incorporating novel biomarkers and integrating molecular subtypes, clinical risk stratification may be improved<sup>43</sup>.

Other recently published CRSs were not externally validated on our cohort for various reasons. The m-CS<sup>20</sup> simplified the traditional CRS and replaced two risk factors by RAS mutational status, and the Liverpool score<sup>23</sup> did not incorporate RAS mutation status in its CRS, which is recognized to be the most promising prognostic factor in patients with CRLM<sup>44–46</sup>. The Chinese CERR<sup>22</sup> included two variables (serum CA 19.9 and bilobar liver distribution of metastatic disease) which were not available in our cohort. For the a-CS<sup>24</sup>, discrepancies in the published survival outcomes and the web-based calculation tool of the a-CS complicate external validation.

When comparing the OS of our population-based cohort with the original GAME cohort, we found a lower median OS in our GAME moderate-risk group. Survival was similar in the GAME low- and high-risk groups. The difference in the moderate-risk group could potentially be influenced by treatment setting. The GAME cohort concerned a selected population treated in a tertiary center with potentially more (experimental) treatment options available, in contrast to our population-based cohort. We did not observe a survival difference between the moderate-risk groups in the subgroups with or without systemic treatment. Therefore, it is unlikely that the greater proportion of systemic therapy administered to the GAME cohort explains the survival differences in the moderate-risk group in our cohort versus the GAME cohort.

Furthermore, as our cohort consists of patients with and without perioperative systemic therapy, we could demonstrate additional interesting survival outcomes. Patients who received perioperative systemic therapy were found to have more prognostic unfavorable characteristics, while the median OS was similar in patients with and without perioperative systemic therapy. This could imply that, in these patients, systemic therapy compensates for the more unfavorable characteristics. This is supported by the findings that patients in the high-risk CRS groups showed a longer median OS and DFS in the subgroup with versus without systemic therapy. Our results are consistent with studies suggesting that high-risk patients with CRLM could benefit from (neo)-adjuvant therapy<sup>9,47–49</sup> and is supported by the negative results of the

EORTC 40983<sup>50</sup> study and the JCOG0603 study<sup>51</sup> for perioperative systemic treatment in, respectively, patients with low-risk disease with <4 CRLM and unselected patients with CRLM. Since the results of the treatment groups are based on retrospective data, this should be confirmed in prospective trials, randomizing high-risk CRLM patients between (neo-)adjuvant therapy or not. However, conducting a study such as this one has proved to be challenging<sup>52</sup>.

Another interesting finding is the OS difference in favor of patients  $\leq 70$  compared to  $>70$  years after resection. Since it did not concern disease-specific but overall survival, other factors, such as comorbidity, might have influenced the OS in this group. This is supported by the result that DFS did not differ between these two subgroups. This OS difference should therefore not be used as an argument against liver resection in patients above 70 years.

The external validation of the CRSs in this study met the TRIPOD guidelines' methodological criteria<sup>40</sup>. Additional strengths include validation of CRS in a real-life population-based cohort which is representative of the whole CRLM population and the near-complete follow-up. Furthermore, the proportion of missing RAS/BRAF mutational status was low, and this was achieved by additional mutational analysis. Selection bias was avoided in correction for missing data by including propensity score matching to identify patients for additional mutational analysis and by using multiple imputation. One limitation of this study is that the patients in our cohort were selected based on primary tumour diagnosis in 2015 and 2016. Thus, our cohort does not include patients with metachronous disease with a long DFI<sup>53</sup>. In addition, selection and information bias is unavoidable given the retrospective nature of the study, although we believe we minimized bias by using a population-based cohort and by handling missing data by multiple imputation. For the validation of the GAME score, mutation status as risk factor was scored by the detection of KRAS codon 12, 13, and 61 mutations, meaning that other RAS mutations were ignored to meet the exact GAME criteria, as proposed by Margonis *et al.* Lastly, the GAME score incorporated patients with extrahepatic disease as a risk factor. As these patients were excluded from our study, the GAME high-risk groups in our validation cohort did not include patients with a maximum of five risk factors.

## CONCLUSIONS

Two established CRSs, Fong and GAME, to predict outcome after CRLM resection were compared and externally validated in a real-life population-based cohort of patients with local treatment of CRLM, regardless of age or the administration of perioperative systemic therapy. Both CRSs showed predictive ability in the real-life cohort, with a better performance of the GAME as compared to the traditional Fong CRS. Although the novel CRS (GAME) outperformed the traditional CRS, the suboptimal predictive value of both CRSs limits the clinical utility of the CRSs. Surgical innovations increase the number of CRLM patients assessed as technically resectable, but high recurrence rates persist, and a significant group of patients has no long-term survival benefit of CRLM resection. Thus, there is still an unmet clinical need for a CRS with high discriminative ability that allows for a better stratification and counselling of patients before surgery and perioperative therapy in order to personalize therapy.

## ACKNOWLEDGMENTS

The authors thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry, as well as IKNL staff for scientific advice. We would like to acknowledge the NKI-AVL Core Facility Molecular Pathology & Biobanking (CFMPB) for supplying lab support and thank Willem Hoefakker and Pien van Diemen for mutational analysis and their contribution to the study. We would like to acknowledge the Sacha Swartttouw-Hijmans fund for the supporting grant.

## REFERENCES

1. van der Geest L.G.M., Lam-Boer J., Koopman M., Verhoef C., Elferink M.A.G., de Wilt J.H.W. Nationwide trends in incidence, treatment and survival of colorectal cancer patients with synchronous metastases. *Clin Exp Metastasis*. (2015);**32**(5):457–65.
2. Rees M., Tekkis P.P., Welsh F.K., O'Rourke T., John T.G. Evaluation of long-term survival after hepatic resection for metastatic colorectal cancer: a multifactorial model of 929 patients. *Ann Surg*. (2008);**247**(1):125–35.
3. Kanas G.P., Taylor A., Primrose J.N., Langeberg W.J., Kelsh M.A., Mowat F.S., et al. Survival after liver resection in metastatic colorectal cancer: review and meta-analysis of prognostic factors. *Clin Epidemiol*. (2012);**4**:283–301.
4. Andreou A., Aloia T.A., Brouquet A., Dickson P. V, Zimmitti G., Maru D.M., et al. Margin status remains an important determinant of survival after surgical resection of colorectal liver metastases in the era of modern chemotherapy. *Ann Surg*. (2013);**257**(6):1079–88.
5. Wakabayashi G., Cherqui D., Geller D.A., Buell J.F., Kaneko H., Han H.S., et al. Recommendations for Laparoscopic Liver Resection - A Report From the Second International Consensus Conference Held in Morioka. *Ann Surg*. (2015);**261**(4):619–29.
6. Rocca A., Cipriani F., Belli G., Berti S., Boggi U., Bottino V., et al. The Italian Consensus on minimally invasive simultaneous resections for synchronous liver metastasis and primary colorectal cancer: A Delphi methodology. *Updates Surg*. (2021);**73**(4):1247–65.
7. Kopetz S., Chang G.J., Overman M.J., Eng C., Sargent D.J., Larson D.W., et al. Improved survival in metastatic colorectal cancer is associated with adoption of hepatic resection and improved chemotherapy. *J Clin Oncol*. (2009);**27**(22):3677–83.
8. Bolhuis K., Kos M., van Oijen M.G.H., Swijnenburg R.J., Punt C.J.A. Conversion strategies with chemotherapy plus targeted agents for colorectal cancer liver-only metastases: A systematic review. *Eur J Cancer*. (2020);**141**:225–38.
9. Nordlinger B., Sorbye H., Glimelius B., Poston G.J., Schlag P.M., Rougier P., et al. Perioperative chemotherapy with FOLFOX4 and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC Intergroup trial 40983): a randomised controlled trial. *Lancet*. (2008);**371**(9617):1007–16.
10. de Jong M.C., Pulitano C., Ribero D., Strub J., Mentha G., Schulick R.D., et al. Rates and patterns of recurrence following curative intent surgery for colorectal liver metastasis: an international multi-institutional analysis of 1669 patients. *Ann Surg*. (2009);**250**(3):440–8.
11. Jones R.P., Jackson R., Dunne D.F., Malik H.Z., Fenwick S.W., Poston G.J., et al. Systematic review and meta-analysis of follow-up after hepatectomy for colorectal liver metastases. *Br J Surg*. (2012);**99**(4):477–86.
12. Imai K., Allard M.A., Benitez C.C., Vibert E., Sa Cunha A., Cherqui D., et al. Early Recurrence After Hepatectomy for Colorectal Liver Metastases: What Optimal Definition and What Predictive Factors? *Oncologist*. (2016);**21**(7):887–94.
13. Fong Y., Fortner J., Sun R.L., Brennan M.F., Blumgart L.H. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. *Ann Surg*. (1999);**230**(3):309–21.
14. Margonis G.A., Sasaki K., Gholami S., Kim Y., Andreatos N., Rezaee N., et al. Genetic And Morphological Evaluation (GAME) score for patients with colorectal liver metastases. *Br J Surg*. (2018);**105**(9):1210–20.

15. Zakaria S., Donohue J.H., Que F.G., Farnell M.B., Schleck C.D., Ilstrup D.M., et al. Hepatic resection for colorectal metastases: value for risk scoring systems? *Ann Surg.* (2007);**246**(2):183–91.
16. Creasy J.M., Sadot E., Koerkamp B.G., Chou J.F., Gonen M., Kemeny N.E., et al. Actual 10-year survival after hepatic resection of colorectal liver metastases: what factors preclude cure? *Surgery.* (2018);**163**(6):1238–44.
17. Mahar A.L., Compton C., Halabi S., Hess K.R., Weiser M.R., Groome P.A. Personalizing prognosis in colorectal cancer: A systematic review of the quality and nature of clinical prognostic tools for survival outcomes. *J Surg Oncol.* (2017);**116**(8):969–82.
18. Nordlinger B., Guiguet M., Vaillant J.C., Balladur P., Boudjema K., Bachellier P., et al. Surgical resection of colorectal carcinoma metastases to the liver. A prognostic scoring system to improve case selection, based on 1568 patients. Association Francaise de Chirurgie. *Cancer.* (1996);**77**(7):1254–62.
19. Iwatsuki S., Dvorchik I., Madariaga J.R., Marsh J.W., Dodson F., Bonham A.C., et al. Hepatic resection for metastatic colorectal adenocarcinoma: a proposal of a prognostic scoring system. *J Am Coll Surg.* (1999);**189**(3):291–9.
20. Brudvik K.W., Jones R.P., Giulante F., Shindoh J., Passot G., Chung M.H., et al. RAS Mutation Clinical Risk Score to Predict Survival After Resection of Colorectal Liver Metastases. *Ann Surg.* (2019);**269**(1):120–6.
21. Lang H., Baumgart J., Heinrich S., Tripke V., Passalacqua M., Maderer A., et al. Extended Molecular Profiling Improves Stratification and Prediction of Survival After Resection of Colorectal Liver Metastases. *Ann Surg.* (2019);**270**(5):799–805.
22. Chen Y., Chang W., Ren L., Chen J., Tang W., Liu T., et al. Comprehensive Evaluation of Relapse Risk (CERR) Score for Colorectal Liver Metastases: Development and Validation. *Oncologist.* (2020);**25**(7):e1031–41.
23. Dupré A., Berhane S., Chan A.W.H., Rivoire M., Chong C.C.N., Lai P.B.S., et al. Multicentre validation of a clinical prognostic score integrating the systemic inflammatory response to the host for patients treated with curative-intent for colorectal liver metastases: The Liverpool score. *Eur J Surg Oncol.* (2019);**45**(6):999–1004.
24. Paredes A.Z., Hyer J.M., Tsilimigras D.I., Moro A., Bagante F., Guglielmi A., et al. A Novel Machine-Learning Approach to Predict Recurrence After Resection of Colorectal Liver Metastases. *Ann Surg Oncol.* (2020);**27**(13):5139–47.
25. Sasaki K., Morioka D., Conci S., Margonis G.A., Sawada Y., Ruzzenente A., et al. The Tumor Burden Score: A New “Metro-ticket” Prognostic Tool For Colorectal Liver Metastases Based on Tumor Size and Number of Tumors. *Ann Surg.* (2018);**267**(1):132–41.
26. Meyer R.M. Generalizing the results of cancer clinical trials. *J Clin Oncol.* (2010);**28**(2):187–9.
27. Booth C.M., Nanji S., Wei X., Mackillop W.J. Management and Outcome of Colorectal Cancer Liver Metastases in Elderly Patients: A Population-Based Study. *JAMA Oncol.* (2015);**1**(8):1111–9.
28. Adam R., Frilling A., Elias D., Laurent C., Ramos E., Capussotti L., et al. Liver resection of colorectal metastases in elderly patients. *Br J Surg.* (2010);**97**(3):366–76.
29. Papamichael D., Audisio R.A., Glimelius B., de Gramont A., Glynne-Jones R., Haller D., et al. Treatment of colorectal cancer in older patients: International Society of Geriatric

- Oncology (SIOG) consensus recommendations 2013. *Ann Oncol.* (2015);**26**(3):463–76.
30. National Comprehensive Cancer Network. Colon Cancer (version 2.2021) [Internet]. Available from: [https://www.nccn.org/professionals/physician\\_gls/pdf/colon.pdf](https://www.nccn.org/professionals/physician_gls/pdf/colon.pdf)
  31. Nederlandse Vereniging voor Heelkunde. Colorectaal Carcinoom (CRC) (version 2019) [Internet]. Available from: [file:///H:/Downloads/Colorectaal\\_carcinoom\\_CRC\\_\(1\).pdf](file:///H:/Downloads/Colorectaal_carcinoom_CRC_(1).pdf)
  32. Van Cutsem E., Cervantes A., Adam R., Sobrero A., Van Krieken J.H., Aderka D., et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol.* (2016);**27**(8):1386–422.
  33. Casparie M., Tiebosch A.T., Burger G., Blauwgeers H., van de Pol A., van Krieken J.H., et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol.* (2007);**29**(1):19–24.
  34. Reddy S.K., Barbas A.S., Turley R.S., Steel J.L., Tsung A., Marsh J.W., et al. A standard definition of major hepatectomy: resection of four or more liver segments. *HPB.* (2011);**13**(7):494–502.
  35. Mekenkamp L.J., Koopman M., Teerenstra S., van Krieken J.H., Mol L., Nagtegaal I.D., et al. Clinicopathological features and outcome in advanced colorectal cancer patients with synchronous vs metachronous metastases. *Br J Cancer.* (2010);**103**(2):159–64.
  36. Beelen K., Opdam M., Severson T.M., Koornstra R.H.T., Vincent A.D., Wesseling J., et al. PIK3CA mutations, phosphatase and tensin homolog, human epidermal growth factor receptor 2, and insulin-like growth factor 1 receptor and adjuvant tamoxifen resistance in postmenopausal breast cancer patients. *Breast Cancer Res.* (2014);**16**(1):1–10.
  37. Bartlett J.W., Seaman S.R., White I.R., Carpenter J.R. Multiple imputation of covariates by fully conditional specification: Accommodating the substantive model. *Stat Methods Med Res.* (2015);**24**(4):462–87.
  38. Morisot A., Bessaoud F., Landais P., Rebillard X., Tretarre B., Daures J.P. Prostate cancer: net survival and cause-specific survival rates after multiple imputation. *BMC Med Res Methodol.* (2015);**15**:54.
  39. Marshall A., Altman D.G., Holder R.L., Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation : current practice and guidelines. *BMC Med Res Methodol.* (2009);**9**(57):1–9.
  40. Moons K.G., Altman D.G., Reitsma J.B., Ioannidis J.P., Macaskill P., Steyerberg E.W., et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* (2015);**162**(1):W1–73.
  41. Rohatgi A. WebPlotDigitizer version 4.4 [Internet]. Pacifica, California, USA; 2020. Available from: <https://automeris.io/WebPlotDigitizer>
  42. Harrell Jr. F.E., Califf R.M., Pryor D.B., Lee K.L., Rosati R.A. Evaluating the yield of medical tests. *JAMA.* (1982);**247**(18):2543–6.
  43. Pitroda S.P., Khodarev N.N., Huang L., Uppal A., Wightman S.C., Ganai S., et al. Integrated molecular subtyping defines a curable oligometastatic state in colorectal liver metastasis. *Nat Commun.* (2018);**9**(1):1793.
  44. Tsilimigras D.I., Ntanasis-Stathopoulos I., Bagante F., Moris D., Cloyd J., Spartalis E., et al. Clinical significance and prognostic relevance of KRAS, BRAF, PI3K and TP53 genetic mutation analysis for resectable and unresectable colorectal liver metastases: A systematic review of the current evidence. *Surg Oncol.* (2018);**27**(2):280–8.

45. Tosi F, Magni E, Amatu A, Mauri G, Bencardino K, Truini M, et al. Effect of KRAS and BRAF Mutations on Survival of Metastatic Colorectal Cancer After Liver Resection: A Systematic Review and Meta-Analysis. *Clin Color Cancer*. (2017);**16**(3):e153–63.
46. Yamashita S, Chun Y.S., Kopetz S.E., Vauthey J.N. Biomarkers in colorectal liver metastases. *Br J Surg*. (2018);**105**(6):618–27.
47. Rahbari N.N., Reissfelder C., Schulze-Bergkamen H., Jager D., Buchler M.W., Weitz J., et al. Adjuvant therapy after resection of colorectal liver metastases: the predictive value of the MSKCC clinical risk score in the era of modern chemotherapy. *BMC Cancer*. (2014);**14**:174.
48. Ayez N., Lalmahomed Z.S., van der Pool A.E., Vergouwe Y., van Montfort K., de Jonge J., et al. Is the clinical risk score for patients with colorectal liver metastases still useable in the era of effective neoadjuvant chemotherapy? *Ann Surg Oncol*. (2011);**18**(10):2757–63.
49. Hirokawa F, Hayashi M, Miyamoto Y, Asakuma M, Shimizu T, Komeda K, et al. Reconsideration of the indications for adjuvant chemotherapy for liver metastases from colorectal cancer after initial hepatectomy. *Ann Surg Oncol*. (2014);**21**(1):139–46.
50. Nordlinger B., Sorbye H., Glimelius B., Poston G.J., Schlag P.M., Rougier P., et al. Perioperative FOLFOX4 chemotherapy and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC 40983): long-term results of a randomised, controlled, phase 3 trial. *Lancet Oncol*. (2013);**14**(12):1208–15.
51. Kanemitsu Y, Shimizu Y, Mizusawa J, Inaba Y, Hamaguchi T, Shida D, et al. A randomized phase II/III trial comparing hepatectomy followed by mFOLFOX6 with hepatectomy alone for liver metastasis from colorectal cancer: JCOG0603 study. *J Clin Oncol*. (2020);**38**(15\_suppl):4005.
52. Ayez N., van der Stok E.P., de Wilt H., Radema S.A., van Hillegersberg R., Roumen R.M., et al. Neo-adjuvant chemotherapy followed by surgery versus surgery alone in high-risk patients with resectable colorectal liver metastases: the CHARISMA randomized multicenter clinical trial. *BMC Cancer*. (2015);**15**:180.
53. Engstrand J., Nilsson H., Stromberg C., Jonas E., Freedman J. Colorectal cancer liver metastases - a population-based study on incidence, management and survival. *BMC Cancer*. (2018);**18**(1):78.

## SUPPLEMENTARY MATERIAL

**Supplementary Table 1.** Assumptions regarding baseline characteristics, systemic treatment, local treatment and survival outcomes

### Assumptions regarding baseline characteristics:

- RAS and BRAF mutation are considered mutual exclusive, therefore patients with RAS mutations or BRAF mutations, were assumed to have BRAF-wildtype or RAS-wildtype status, respectively.
- Primary tumour nodal status was defined primarily on pathologic N-stadium. When pN stage was missing, cN stage (radiological) was used.
- If number of metastases was not given and code 77 was used (accounting for diffuse metastatic disease in the liver) then number of metastases was scored as 20.

### Assumptions regarding systemic treatment regimens and strategies:

- Systemic treatment includes both chemotherapy and/or targeted therapy.
- A combination regimen is defined as all systemic agents starting within 4 weeks after start of the first agent and started before progression of disease.
- If bevacizumab was started more than 4 weeks after the start of the first agent but before stop of this agent and before progression of disease, we assume bevacizumab was part of this combination regimen.
- If a treatment line continues despite of progression, e.g., in case of reintroduction of the same or an equivalent regimen after a therapy break and detected progression, we regard this as continuation of the same treatment line.
- If oxaliplatin only is registered, we assume this was part of a capecitabine and oxaliplatin (CapOx) regimen of which capecitabine was not registered, so we add capecitabine. We assume this is due to a registration error, in which the administration of capecitabine has not been noticed by the data manager since it is registered differently as oral medication.
- Systemic therapy was considered adjuvant systemic therapy for primary tumour when started < 12 weeks after resection of primary tumour and started before diagnosis of metastases in patients with metachronous disease.
- Capecitabine monotherapy was considered radiosensitizer for primary tumour when started before primary tumour resection and before diagnosis of metastases and with notification to have received chemoradiotherapy.
- Systemic therapy was considered pre-operative therapy (neo-adjuvant or induction) before liver resection when the therapy ended within 120 days before liver resection. Adjuvant therapy after resection of primary tumour or chemotherapy as radiosensitizer was excluded.
- Systemic therapy was considered adjuvant therapy after liver resection when the therapy started within 120 days after liver resection. Chemotherapy as radiosensitizer was excluded.
- Systemic therapy was considered peri-operative therapy of liver resection when the systemic therapy was given < 120 days before and < 120 days after liver resection
- When systemic therapy was given between two liver procedures before progression of disease, the first liver procedure was considered as staging procedure and systemic therapy was considered as pre-operative systemic therapy (neo-adjuvant or induction) for surgery

- A treatment line is defined as systemic therapy (monotherapy or combination regimen) administered at the same time until suspension, regardless of reason for discontinuation.
- Treatment is considered as next line if an agent of a new drug group is started that is not applied in the previous systemic treatment regimen.
- If the same or an equivalent systemic treatment regimen is (re)started, this is considered continuation of the same treatment line, e.g., CapOx to 5-FU/oxaliplatin (FOLFOX).

---

**Assumptions regarding local treatment**

---

- Local treatments are categorized as follows:
  - o 1 stage (1 procedure)
  - o 2-stage (2 procedures < 120 days apart)
- R-status: when two stage procedure and first procedure was R2 resection and second procedure was R1/R0 resection then 2-stage resection considered as R-status of last procedure.
- when 2-stage resection and one procedure was R1 resection and other local treatment was R0 resection then considered as R1 resection.

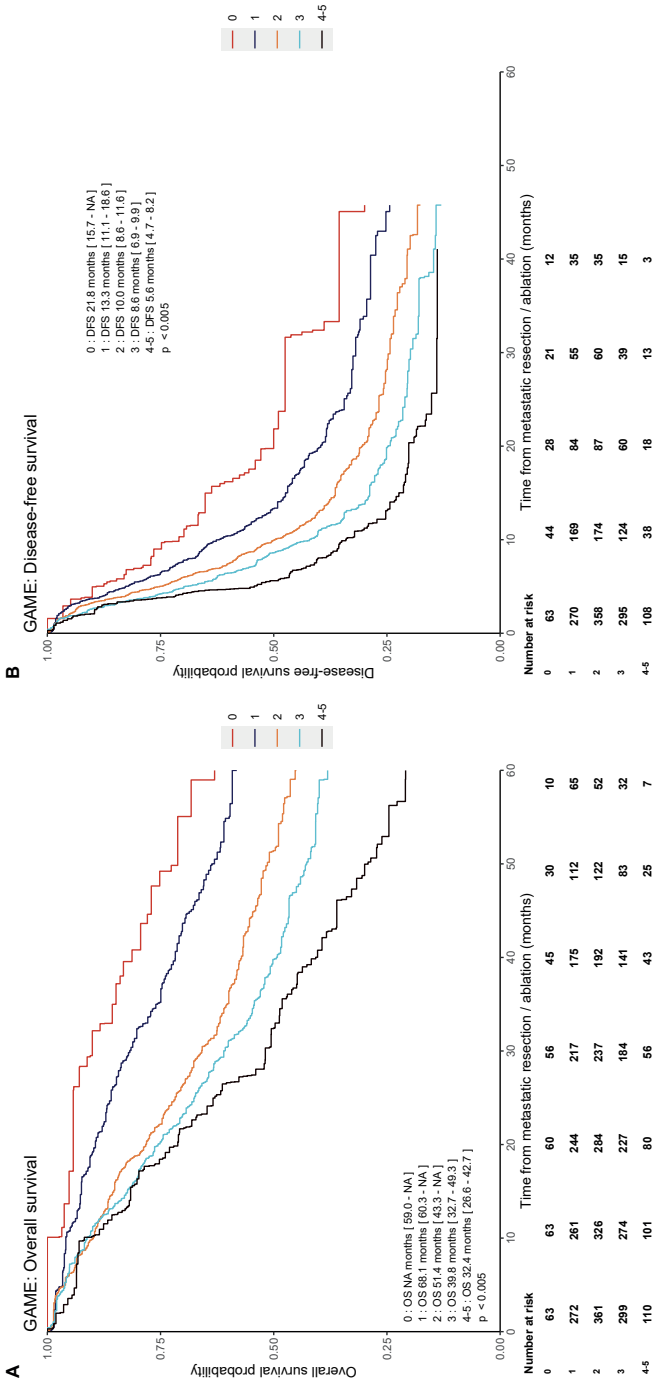
---

**Assumptions regarding progression of disease and survival:**

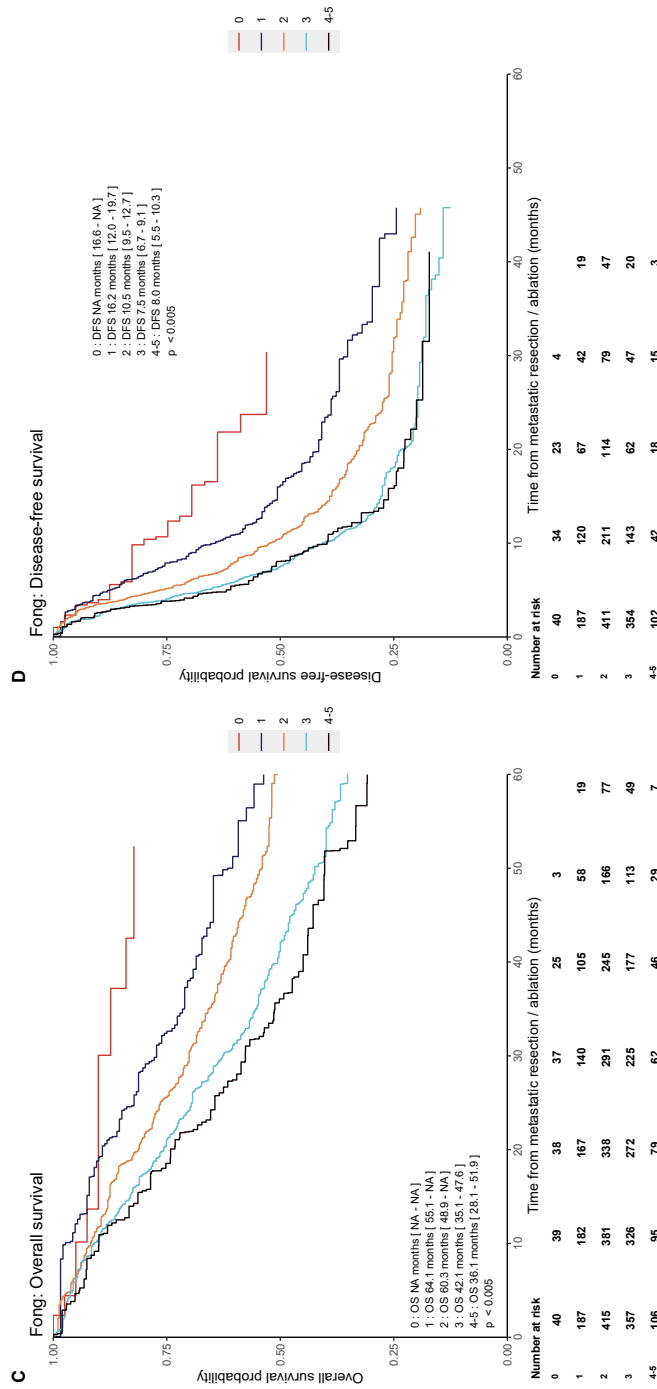
---

- Date of new episode is considered as time of progression.
  - When disease progression is documented < 14 days of liver resection we assume this was part of the liver resection and first new episode is considered as time of progression.
  - Disease-free survival is calculated from date of first liver procedure to date of progression. In case of 2-stage resection, DFS is calculated from last liver procedure.
  - If no recurrence is registered:
    - o If end of follow up is registered and reason end of follow up is: death, then date of death is registered as event of DFS;
    - o If end of follow up is registered and reason end of follow up is other then death then DFS is censored on date of end of follow up;
    - o If no date of end of follow up is registered then DFS is censored on date of last visit;
    - o If none of these dates are registered then DFS is documented as missing.
  - Lymph node metastases registered as abdominal lymph nodes at time of first liver metastases were considered extrahepatic disease and as so classified as not-liver only disease.
  - Overall survival (OS) after resection was defined as date of first resection till date of last documented vital status as documented by the municipal population registry. In case of 2-stage resection, OS is calculated from date of last liver procedure
    - o If the documented date of disease-free survival is after date of documented survival then the date of disease-free survival is date of last survival
  - Patients who did not die are censored on the date last known to be alive in the municipal population registry.
-

**Supplementary Figure 1.** Kaplan-Meier analysis showing overall survival and disease-free survival curves of the total cohort in scores following the GAME clinical risk score (A and B) or the Fong clinical risk score (C and D).

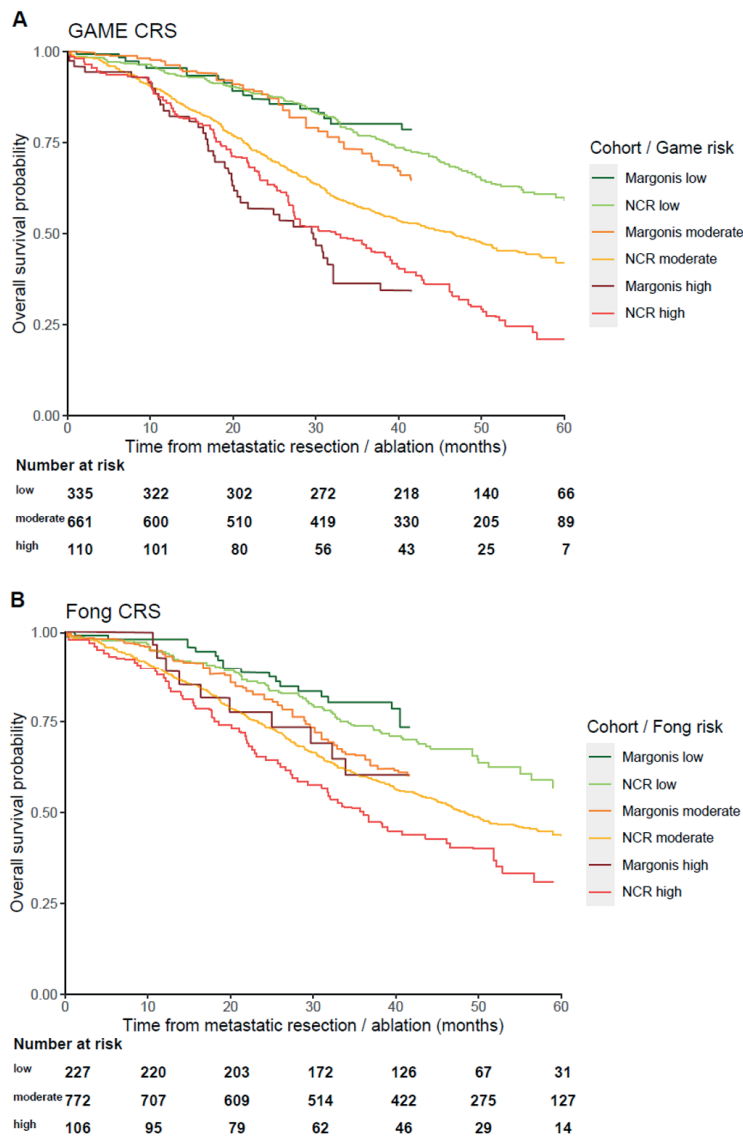


Supplementary Figure 1 (continued).



Suppl. Fig. 1. Abbreviations: DFS (disease-free survival), NA (not applicable), OS (overall survival).

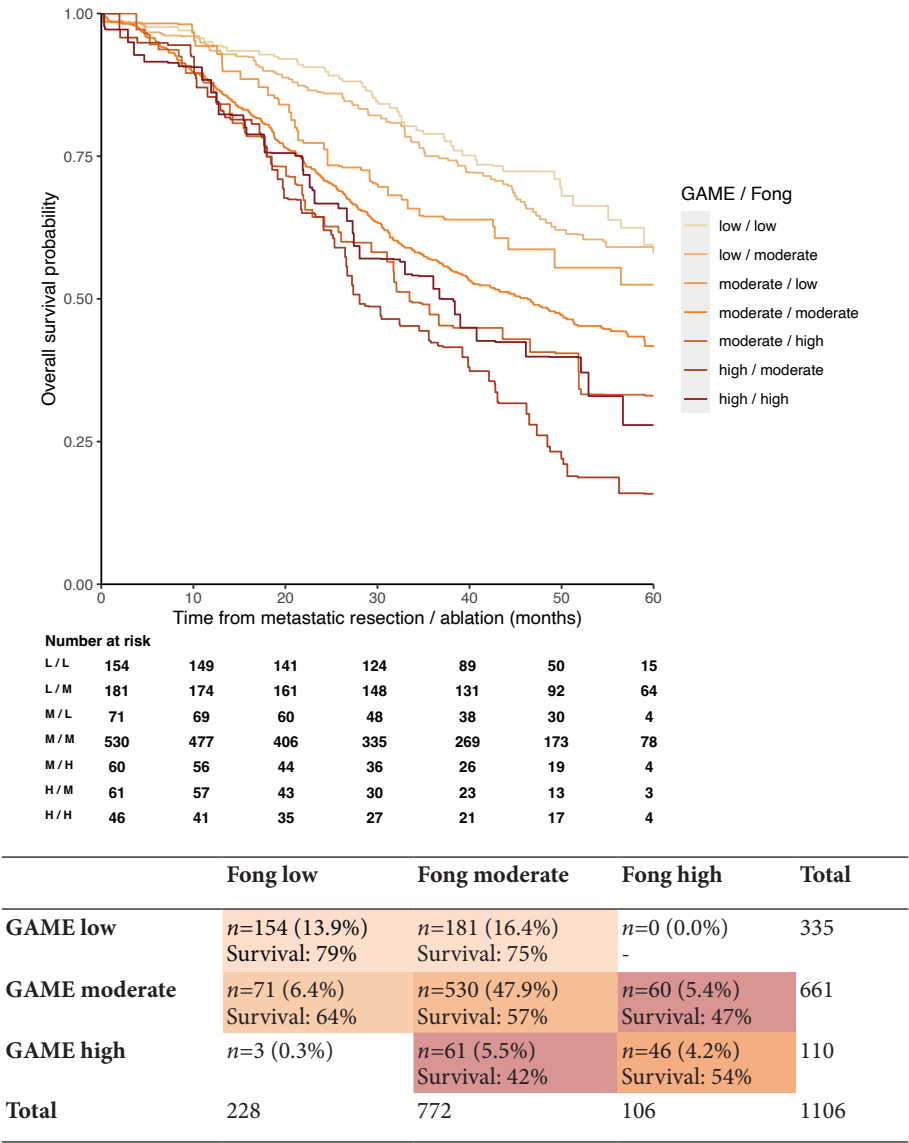
**Supplementary Figure 2.** Calibration of the expected and observed survival outcomes in the NCR cohort of the risk groups (low, moderate and high) according to: A) GAME and B) Fong prediction model.



Suppl. Fig. 2. This calibration is done by digitizing the original survival curves of the two prediction models as published by *Margonis et al.* (expected outcomes, **darker lines**) and these curves are compared with the actual survival curves of the NCR cohort (observed outcomes, **lighter lines**). The risk table displays the number at risk for the NCR cohort.

*Abbreviations:* CRS (clinical risk score), NCR (Netherlands Cancer Registry cohort).

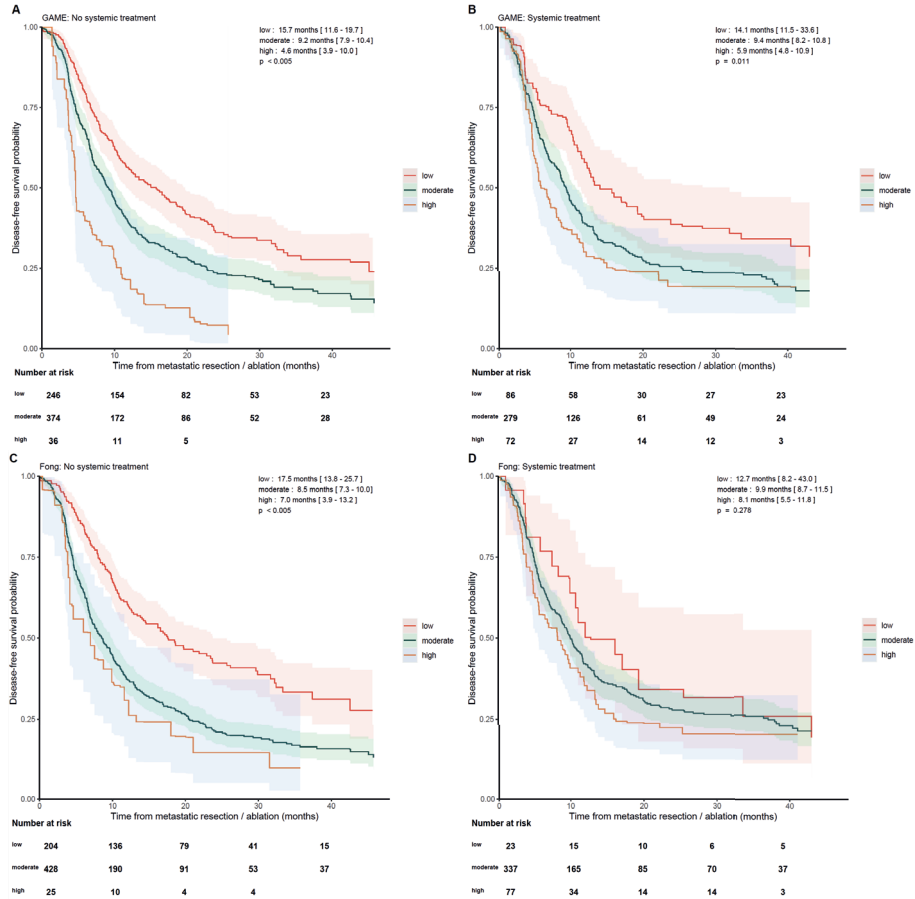
**Supplementary Figure 3.** A combined figure containing on the left, a contingency table showing the frequency distribution of patients among the risk categories (low, moderate, high) following the Fong and GAME prediction model, their corresponding 3-year survival rate estimate, which is also indicated by the heat map for each category



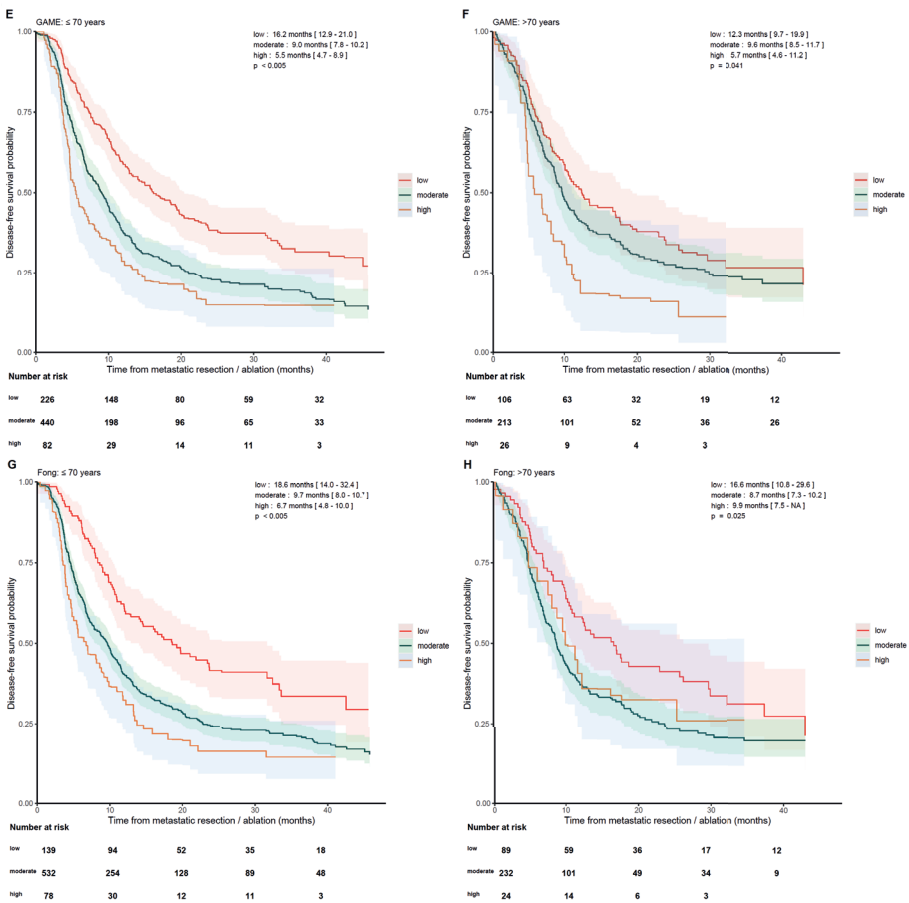
Suppl. Fig. 3. The corresponding survival curves for the groups are displayed in the KM plot on the right. Number of patients does not count up to total number of patients in study due to rounding effects of analyses in the imputed database.

Abbreviations: *n* (count).

**Supplementary Figure 4.** Kaplan-Meier analysis showing disease-free survival (DFS) curves for GAME risk groups for patients without (A) and with (B) perioperative systemic therapy and for age groups  $\leq 70$  years (E) and  $> 70$  years (F). The DFS outcomes of Fong risk groups are shown for patients without (C) and with (D) perioperative systemic therapy and for age groups  $\leq 70$  years (G) and  $> 70$  years (H).



Supplementary Figure 4 (continued).



Suppl. Fig. 4. Red lines represent the low risk groups, green lines the moderate risk and yellow lines the high risk group.

**Supplementary Table 2.** Pooled Harrell's concordance index with 95% confidence intervals for 1- and 3-year overall survival and disease-free survival outcomes for GAME and Fong risk scores in subgroups without and with perioperative systemic therapy

WITHOUT PERIOPERATIVE SYSTEMIC THERAPY									
	GAME score		Survival estimates GAME risk categories		Fong score		Survival estimates Fong risk categories		
	C-index [95% C.I.]	Low (%)	Moderate (%)	High (%)	C-index [95% C.I.]	Low (%)	Moderate (%)	High (%)	
OS									
1-year OS	0.592 [0.518-0.667]	94	89	84	0.596 [0.532-0.661]	95	89	84	
3-year OS	0.610 [0.574-0.647]	78	57	43	0.593 [0.559-0.628]	75	60	48	
5-year OS	0.602 [0.569-0.635]	60	41	26	0.594 [0.563-0.624]	59	42		
DFS									
1-year DFS	0.584 [0.553-0.614]	56	39	21	0.606 [0.578-0.635]	61	37	31	
3-year DFS	0.579 [0.551-0.606]	28	18	9	0.601 [0.575-0.627]	33	16		
WITH PERIOPERATIVE SYSTEMIC THERAPY									
	GAME score		Survival estimates GAME risk categories		Fong score		Survival estimates Fong risk categories		
	C-index [95% C.I.]	Low (%)	Moderate (%)	High (%)	C-index [95% C.I.]	Low (%)	Moderate (%)	High (%)	
OS									
1-year OS	0.588 [0.511-0.664]	96	88	86	0.538 [0.460-0.617]	96	89	88	
3-year OS	0.590 [0.549-0.631]	74	56	49	0.556 [0.515-0.598]	60	60	51	
5-year OS	0.590 [0.554-0.627]	58	43	23	0.557 [0.519-0.594]		46	29	
DFS									
1-year DFS	0.589 [0.551-0.626]	58	39	30	0.563 [0.528-0.598]	50	42	35	
3-year DFS	0.581 [0.547-0.616]	34	23	19	0.559 [0.527-0.591]	26	26	20	

Survival estimates at 1- and 3- years for low, moderate and high risk groups according to GAME and Fong prediction model. When not indicated, the number of patients was too small to calculate the survival estimate.

Abbreviations: C-index (concordance index), DFS (disease-free survival), OS (overall survival).

**Supplementary Table 3.** Pooled Harrell's concordance index with 95% confidence intervals for 1- and 3-year overall survival and disease-free survival outcomes for GAME and Fong risk scores in subgroups of  $\leq 70$  years and  $> 70$  years. Survival estimates at 1- and 3-years for low, moderate and high risk groups according to GAME and Fong prediction model.

AGE ≤ 70 YEARS										
	GAME score		Survival estimates GAME risk categories			Fong score		Survival estimates Fong risk categories		
	C-index [95% C.I.]	Low (%)	Moderate (%)	High (%)	C-index [95% C.I.]	Low (%)	Moderate (%)	High (%)		
OS										
1-year	0.611 [0.540-0.681]	96	90	88	0.583 [0.515-0.650]	96	92	86		
3-year	0.618 [0.585-0.652]	82	59	50	0.588 [0.555-0.621]	76	64	48		
5-year	0.613 [0.585-0.642]	64	45	21	0.584 [0.554-0.613]	56	48	32		
DFS										
1-year	0.601 [0.573-0.628]	59	38	29	0.595 [0.569-0.621]	61	40	32		
3-year	0.594 [0.568-0.619]	32	19	15	0.591 [0.567-0.614]	34	21	14		
AGE > 70 YEARS										
	GAME score		Survival estimates GAME risk categories			Fong score		Survival estimates Fong risk categories		
	C-index [95% C.I.]	Low (%)	Moderate (%)	High (%)	C-index [95% C.I.]	Low (%)	Moderate (%)	High (%)		
OS										
1-year	0.572 [0.491-0.654]	90	85	79	0.601 [0.527-0.675]	93	83	89		
3-year	0.580 [0.533-0.627]	67	52	37	0.590 [0.547-0.634]	70	50	58		
5-year	0.575 [0.531-0.618]	49	36	37	0.589 [0.548-0.630]	58	34			
DFS										
1-year	0.554 [0.511-0.597]	51	41	22	0.584 [0.542-0.625]	58	37	41		
3-year	0.547 [0.507-0.587]	26	23	22	0.575 [0.536-0.613]	31	20			

When not indicated, the number of patients was too small to calculate the survival estimate.  
Abbreviations: C-index (concordance index), DFS (disease-free survival), OS (overall survival).





# PREDICTING EARLY EXTRAHEPATIC RECURRENCE AFTER LOCAL TREATMENT OF COLORECTAL LIVER METASTASES

Submitted to *Annals of Surgery*.

Emerens Wensink<sup>†</sup>  
Karen Bolhuis<sup>†</sup>  
Marloes Elferink  
Remond Fijneman  
Onno Kranenburg  
Inne Borel Rinkes  
Miriam Koopman  
Rutger-Jan Swijnenburg  
Geraldine Vink  
Jeroen Hagendoorn  
Cornelis Punt  
Jeanine Roodhart<sup>‡</sup>  
Sjoerd Elias<sup>‡</sup>

<sup>†</sup> Both authors share equal contribution.

<sup>‡</sup> Both authors share last authorship.

## ABSTRACT

### *Objective*

We aimed to develop and internally validate a prediction model for early extrahepatic recurrence (EHR) after local treatment of colorectal liver metastases (CRLM).

### *Summary Background Data*

Patients who develop early EHR ( $\leq 6$  months) may not benefit from local treatment of CRLM. However, prediction models for early EHR are not available.

### *Methods*

We used a population-based cohort of 1077 patients locally treated for CRLM to develop and internally validate a prediction model for EHR using Cox regression. Performance assessment included calibration, discrimination, net benefit, and generalizability by internal-external cross-validation. The prognostic relevance of 6-month EHR was evaluated by landmark analysis.

### *Results*

During a median follow-up of 35 months, 557 patients developed EHR, and 249 died. The median OS for patients with EHR within six months after CRLM treatment was 19.5 months (95% C.I. 15.6-23.0) versus not reached (45.3-not reached). The EHR prediction model included sidedness of primary tumour, T-stage and N-stage of primary tumour, *RAS/BRAF*<sup>V600E</sup> mutational status, and number and size of CRLM. The model was well-calibrated, yielded overoptimism-corrected 6-month EHR risks between 5.9-56.0% (interquartile interval 12.9-22.0%). Harrell's C-index through 6 and 12 months was 0.663 (0.624-0.702) and 0.661 (0.632-0.689), respectively. Patients in the highest risk quartile had an observed 6-month EHR risk of 32% versus 6% in the lowest quartile. Expected generalizability was good. EHR risk-informed CRLM treatment decisions yielded net benefit at 6-month EHR thresholds of 0-40%.

### *Conclusions*

Early EHR after local treatment of CRLM has a major impact on prognosis and can be predicted with routine clinical information.

## INTRODUCTION

Colorectal cancer liver metastases (CRLM) are the major cause of colorectal cancer (CRC)-related death<sup>1</sup>. Local treatment of CRLM without extrahepatic metastatic involvement, such as liver resection, offers the only chance for cure or long-term survival, with 5-year survival rates of up to 55%<sup>2-5</sup>. Improved surgical techniques, optimization of systemic neoadjuvant induction treatment with high response rates and more lenient eligibility criteria have increased the number of patients assessed as technically resectable and undergoing CRLM resection<sup>4,6</sup>.

However, relapse after local CRLM treatment occurs in up to 75% of patients, often with unresectable recurrences and subsequent decreased survival<sup>5,7,8</sup>. Numerous prediction models for (recurrence-free) survival after local treatment of CRLM exist<sup>9-15</sup>, but these are not widely used to guide decision-making due to their inability to identify patients with a sufficiently short survival to render local treatment unjustified. Aspects which might contribute to this include suboptimal incorporation of prognostic factors and the use of (recurrence-free) survival as an endpoint.

A limitation of recurrence-free survival (RFS) as endpoint is its inability to discriminate between intra- and extrahepatic recurrences. Patients with liver-limited recurrences may be eligible for repeat resections, resulting in long-term survival and thus should not be excluded from local treatment<sup>8,16-18</sup>. In contrast, patients with extrahepatic recurrence (EHR) rarely undergo repeated local treatment<sup>18-20</sup>. An early recurrence, usually defined as occurring within six months<sup>21,22</sup>, and EHR are independently associated with poor overall survival (OS) in patients receiving local treatment for CRLM<sup>21-23</sup>. Therefore local treatment of CRLM may not be justified in patients who develop an early EHR. Being able to predict early EHR may spare patients an invasive treatment, treatment-related morbidity and avoid delay in starting systemic treatment which may effectively treat the systemic disease present. Moreover, early EHR estimates may stratify patients for perioperative systemic therapy, since patients receiving local CRLM treatment have no OS benefit from perioperative systemic therapy<sup>24,25</sup>.

Although early EHR after local treatment of CRLM is of major clinical importance, to our knowledge, no prediction models for early EHR exist. Also, novel prognostic factors, such as primary tumour localization and *RAS*/*BRAF*<sup>V600E</sup> tumour mutational status, may aid in better identifying patients at high risk for early EHR. Patients with right-sided primary tumours have a worse prognosis after local treatment of CRLM,

more recurrences at multiple sites and less repeated local treatment as compared to patients with left-sided primaries<sup>26,27</sup>. The presence of *RAS* and *BRAF*<sup>V600E</sup> mutations is associated with a higher recurrence rate of up to 94%, with EHR not amenable to local therapy and a shorter EHR-free survival (EHRFS)<sup>8,28–30</sup>.

We aimed to develop and internally validate a prediction model, which incorporates primary tumour location and *RAS*/*BRAF*<sup>V600E</sup> mutational status as novel prognostic factors, for early EHR following local treatment of CRLM using a population-based cohort.

## METHODS

### Patient cohort

All patients initially diagnosed with CRC between January 1st 2015 and December 31st 2016, who underwent local treatment (resection and/or local ablation) for CRLM, were identified in the Netherlands Cancer Registry (NCR), a population-based registry of all newly diagnosed cancer patients in the Netherlands<sup>31</sup>. Patients with extrahepatic metastases before resection, R2 liver resections, appendix carcinoma, concomitant local liver treatment other than resection or ablation, and without any follow-up information were excluded. The scientific committee of the Netherlands Comprehensive Cancer Organisation (IKNL) approved the research protocol and use of this data, and the requirement of written informed consent was waived for this study. The study was performed in accordance with the Declaration of Helsinki and reported according to the TRIPOD guidelines<sup>32</sup>.

### Candidate predictor variables

Data was extracted from the NCR including but not limited to: age, sex, American Joint Committee on Cancer (AJCC) tumour stage (T-stage), nodal stage (N-stage) of the primary tumour, location of the primary tumour (right-sided (coecum-transverse colon), left-sided (splenic flexure-rectosigmoid) and rectum), disease-free interval between detection of primary tumour and metastases (DFI), size (mm) and number of liver metastases, serum CEA level (ug/L) prior to local treatment of CRLM, type of local treatment, resection margin status (R0 versus R1) and administered perioperative systemic treatment. A major resection was defined as resection of  $\geq 4$  liver segments<sup>33</sup>, synchronous disease as DFI  $\leq 6$  months<sup>34</sup> and perioperative systemic therapy as therapy administered  $\leq 100$  days before and/or after local CRLM treatment and initiated prior

to progression of disease. Intent of systemic treatment was not registered, precluding a distinction between neo-adjuvant or induction systemic therapy. However, as Dutch guidelines for CRC<sup>35</sup> recommend not to administer perioperative systemic therapy in initially resectable CRLM, we assume that preoperative systemic treatment was given as induction treatment to achieve CRLM resectability. Further assumptions regarding systemic treatment are described in Supplementary Table 1.

*RAS/BRAF*<sup>V600E</sup> mutational status was retrieved from the NCR and the national automated pathological archive (PALGA<sup>36</sup>), determined in daily practice on primary tumour or metastases at any time during the disease course. Missing *KRAS*, *NRAS* and *BRAF*<sup>V600E</sup> mutation status was complemented by an additional Sequenom Massarray mutation analysis of tumour tissue of 250 patients. These 250 additional samples were selected in such a way to maximize mutation status information for patient subgroups in which that information was otherwise underrepresented, increasing the likelihood of successful multiple imputation<sup>37</sup>.

## Patient outcomes

Follow-up data for recurrences was collected from medical records until May 2020 and survival was obtained by linkage with the municipal population registry on January 31<sup>st</sup>, 2021. OS was defined as date of first local treatment for CRLM till date of death or last follow-up. RFS and EHRFS was defined as date of first local treatment of CRLM till date of a RFS or EHRFS event, which was defined as first recurrence of disease or first EHR or death, whichever occurred first, or censored on last date of RFS or EHRFS without event, respectively. In two-stage resections, OS, RFS and EHRFS was calculated since the date of the last procedure. If follow-up for recurrences was shorter than follow-up for survival, all survival follow-up beyond the last follow-up for recurrences was discarded for assessment of RFS, EHRFS or OS. In all patients a minimum of one year RFS and two year OS follow-up was ensured. All assumptions regarding OS, RFS and EHRFS are included in Supplementary Table 1.

## Statistical analysis

We used standard descriptive statistics to describe the study population, including medians and interquartile intervals (IQI) for continuous variables, and frequency and percentages for categorical variables. Follow-up data and patient outcomes were described using (reverse) Kaplan Meier approaches.

## Prediction model development and performance assessment

Early EHR ( $\leq 6$  months<sup>21,22</sup>) was defined as the clinically relevant primary endpoint of the model, due to the poor prognosis in patients with early EHR and lower chance of repeat local treatment, in contrast to patients with liver-only recurrences. We assessed the prognostic impact of our primary endpoint using landmark analysis at six months after CRLM treatment.

Based on published recommendations<sup>38</sup>, we had sufficient data to model 17 coefficients. Nine candidate predictors were selected for model development by assessment of a multidisciplinary team based on literature of previous prediction models and novel prognostic factors<sup>9–12,26,39</sup>. The predictors, including four continuous variables modeled non-linearly, were: neoadjuvant systemic treatment, primary tumour location, T-stage, N-stage, *RAS/BRAF*<sup>V600E</sup> mutational status, number of liver metastases, size of largest liver metastasis, pre-operative CEA and DFI. We used multiple imputation using multivariate imputation by chained equations (MICE)<sup>40</sup> to account for missing data, generating 53 imputed datasets (based on the percentage of patients with any missing data in the candidate predictors).

A prediction model for EHRFS after local treatment of CRLM was developed using Cox regression, with a time-horizon of 12 months to improve the effective sample size, but with a primary evaluation of the model's performance for EHR  $\leq 6$  months. The prediction model was developed in the whole cohort, using Akaike Information Criterion (AIC)-based backward selection in each imputed dataset leading to a primary model only including predictors selected in  $\geq 50\%$  of imputed datasets, which was then refitted in each imputed dataset to obtain a pooled selection model using Rubin's rules ('EHR model'). Adjuvant systemic therapy was included in all models using an offset for expected therapeutic efficacy based on the pooled adjuvant systemic treatment effect from published randomized controlled trials<sup>24,25</sup>.

Model performance was assessed using calibration plots for 6- and 12-month EHR risk, discrimination (Harrell's C-index, Uno's C-index through 6-months and 12-months), time-dependent receiver operator characteristic (ROC) curve, Nagelkerke's  $R^2$  and decision curve analysis. Each measure was determined for each imputed dataset separately and pooled using Rubin's rules, incorporating appropriate data-transformation steps. Decision curve analysis was used to assess the net benefit associated with CRLM treatment decisions based on a given threshold value for

6-month or 12-month EHRFS probability<sup>41</sup>. To visualize the model's potential relevance, we used Kaplan-Meier curves for EHRFS, RFS and OS, categorizing patients based on quartiles of predicted EHR risk. The reported performance measures were based on predicted EHR-risks which included the effect of adjuvant treatment (if given).

We used internal validation by 500-fold bootstrap resampling, repeating all model-development steps in each bootstrap sample, to obtain an overoptimism-corrected model (using uniform shrinkage) and C-index. We used internal-external cross-validation, including all modeling steps, to evaluate the generalizability of the model based on three geographic regions.

In an exploratory analysis, we tested whether the prognostic value of *RAS* mutation for EHRFS depended on the administration of preoperative systemic treatment as reported by others<sup>28,29</sup>, using a multivariable model with a *RAS*\*preoperative systemic treatment interaction term.

A more detailed description of the methods is described in the Supplementary Methods. Analyses were performed using SPSS software version 25 (IBM, Armonk, New York, USA) and R version 4.0.3 (2020-10-10) with the following libraries: rms (V6.2-0), pec (V2022.03.0), survival (V3.3-1), mice (V3.14.0), survivalROC (V1.0.3).

## RESULTS

### Patient cohort

All 1105 patients who underwent local treatment (resection and/or ablation) for CRLM were selected from the NCR for analysis, for which the primary endpoint regarding early EHR was available in 1077 patients. In 11 of the 1105 (<1%) patients, no follow-up data was available.

The patient characteristics of the cohort are displayed in Table 1. Overall, the median age was 66 years, 403 (37%) were females, 797 (74%) presented with synchronous disease and 256 (24%) with a right-sided primary tumour. A total of 427 (40%) patients received systemic treatment, and in 173 (16%) patients a major resection (hemihepatectomy with/without local ablation and/or wedge resection) was performed. The *RAS*/*BRAF* mutation status was available in 701 (65%) patients, of whom 352 (50%) harbored a *RAS* mutation and 19 (3%) a *BRAF*<sup>V600E</sup> mutation.

**Table 1.** Characteristics of 1105 Dutch CRC patients diagnosed in 2015-2016 who received local treatment for CRLM

		<b>Overall</b> <i>n=1077</i>
Age, years	Median (IQI)	66 [59-72]
Sex	Female, n (%)	403 (37.4)
Side primary tumour	Right, n (%)	256 (23.8)
	Left, n (%)	460 (42.7)
	Rectum, n (%)	361 (33.5)
Chemoradiotherapy primary tumour	Yes, n (%)	127 (11.8)
T-stage	1, n (%)	27 (2.5)
	2, n (%)	126 (11.8)
	3, n (%)	740 (69.1)
	4, n (%)	178 (16.6)
	Missing	6
N-stage	N0, n (%)	398 (37.0)
	N1, n (%)	380 (35.3)
	N2, n (%)	297 (27.6)
	Missing	2
Stage of disease at diagnosis	I, n (%)	25 (2.3)
	II, n (%)	102 (9.5)
	III, n (%)	185 (17.2)
	IV, n (%)	765 (71.0)
Differentiation grade of CRC	Low, n (%)	15 (1.5)
	Intermediate, n (%)	916 (92.0)
	High, n (%)	65 (6.5)
	Missing	81
Time to metastases	Synchronous, n (%)	797 (74.0)
Number of liver metastases	Median (IQI)	2 [1- 4]
	Missing	41
Size largest liver metastasis, mm	Median (IQI)	24 [16-36]
	Missing, n (%)	83
CEA level, ug/L	Median (IQI)	9.00 [3.30, 36.00]
	Missing	225
Type of surgery	Local ablative therapy only, n (%)	97 (9.0)
	Wedge / segment resection only, n (%)	596 (55.3)
	Minor resection and local ablative therapy, n (%)	211 (19.6)
	Hemihepatectomy with/without ablation/wedge, n (%)	173 (16.1)

		<b>Overall n=1077</b>
R-status	R0, n (%)	841 (78.1)
	R1, n (%)	141 (13.1)
	Unknown due to ablation	95 (8.8)
Perioperative systemic therapy	Neoadjuvant only, n (%)	322 (29.9)
	Adjuvant only, n (%)	51 (4.7)
	Peri-operative, n (%)	54 (5.0)
	None, n (%)	650 (60.4)
Tumour mutational status	<i>RAS/BRAF<sup>V600E</sup></i> wildtype, n (%)	330 (47.1)
	<i>BRAF<sup>V600E</sup></i> mutation, n (%)	19 (2.7)
	<i>RAS</i> mutation, n (%)	352 (50.2)
	Missing ( <i>RAS</i> and/or <i>BRAF</i> status)	376
MMR-status	MMR-deficient, n (%)	15 (2.3)
	MMR-proficient, n (%)	632 (97.7)
	Missing	430

Characteristics of 1105 Dutch patients diagnosed in 2015-2016 with CRC who received local treatment for CRLM are shown for the cohort. The count (%) for categorical variables and median (IQI) for continuous variables is shown.

**Abbreviations:** CEA (carcinoembryonic antigen), CRC (colorectal cancer), CRLM (colorectal liver metastases), IQI (interquartile interval), mm (millimetres), MMR (mismatch repair), N-stage (nodal stage of primary tumour), R-status (resection margin status), T-stage (tumour stage of primary tumour), ug/L (microgram per liter).

## Patient outcomes after local treatment of CRLM

In the cohort, 807 (73%) recurrences were observed during a median follow-up of 35 months. The median OS was 51.3 months (95% C.I. 49.3-not reached (NR)) and RFS was 10.1 months (95% C.I. 9.5-10.9; Supplementary Figure 1). The site of first recurrence was liver-only in 332 (43.3%) patients and extrahepatic (with/without intrahepatic metastases) in 399 (52.2%) patients (Supplementary Table 2). The median EHRFS was 20.4 months (95% C.I. 18.8-23.4). In the cohort, there were 557 (51.7%) EHR events, of which 194 (18.0%)  $\leq 6$  months and 363 (33.7%)  $\leq 12$  months. First EHR concerned a multisite EHR in 127 (26.6%) patients. The site of first EHR was most frequently the lungs and lymph nodes in 213 (44.6%) and 55 (11.5%) patients, respectively, whereas the brain was affected in 6 (1.3%) patients. The site of first EHR was significantly correlated with post-recurrence survival,  $p < 0.001$ , with the shortest post-recurrence survival in patients with brain metastases (1.9 months) and the longest in patients with EHR in the lymph nodes, intra-abdomen or lungs (23.4, 27.1 and 29.3 months, respectively; Supplementary Figure 2). The time to recurrence was similar for all extrahepatic sites,  $p = 0.55$ . Notably, 40 (20.6%) patients with early EHR had a hemihepatectomy and 16 (8.2%) patients a two-stage resection.

## Prognostic relevance of 6-month EHR

Patients who survived until the prespecified landmark time (6 months after local treatment of CRLM,  $n=982$ ) were included in the landmark analysis to compare survival outcomes according to type of recurrence. Within the first six months after local treatment of CRLM, 726 (73.9%) patients had no recurrence, 123 (12.5%) suffered a liver-only recurrence and 133 (13.5%) an EHR (including 100 patients with an extra- and intrahepatic recurrence). The median OS from the landmark time for patients with 6-month EHR after CRLM treatment was 19.5 months (95% C.I. 15.6-23.0; Supplementary Figure 3). The median OS from the landmark time in patients with liver-only recurrence was 30.7 months (95% C.I. 29.0-NR) and not reached (95% C.I. 45.3-NR) in patients without a recurrence within the first six months.

## Prognostic value of tumour mutational status and sidedness of primary tumour

The prognostic value of the novel candidate predictors, tumour mutational status and sidedness of the primary tumour, was first explored using univariable analysis for OS, EHRFS and RFS (Supplementary Figure 4 and Table 2). The median EHRFS for patients with *BRAF*<sup>V600E</sup> mutated, *RAS* mutated and *RAS*/*BRAF*<sup>V600E</sup> wildtype tumours was 11.4 months (95% C.I. 5.8-NR), 18.5 months (95% C.I. 14.3-20.9) and 28.2 months (95% C.I. 22.2-33.9,  $p<0.005$ ), respectively. The EHRFS for patients with right-sided, rectum or left-sided localized tumours was 18.5 (95% C.I. 13.3-32.0), 18.6 (95% C.I. 14.5-23.8) and 23.0 (95% C.I. 20.4-32.3,  $p<0.05$ ) months, respectively.

**Table 2.** Specifications of prediction model for extrahepatic recurrence-free survival

Variable	Level	n	Univariable			Full multivariable			Selection model		
			HR	p-value	HR	p-value	HR	p-value	HR	p-value	Shrunk HR
Neoadjuvant treatment	Not received	755	REF		REF		REF	0.371	-		
	Received	322	1.30 [1.05 - 1.62]	0.02	0.89 [0.69 - 1.15]		-		-		
Sidedness	Right	256	REF		REF		REF	0.074	REF	0.080	
	Left	460	0.73 [0.56 - 0.95]	0.02	0.94 [0.71 - 1.25]		0.94 [0.71 - 1.25]		0.94 [0.71 - 1.25]		0.95
T-stage	Rectum	361	0.96 [0.74 - 1.26]	0.79	1.24 [0.94 - 1.65]		1.24 [0.93 - 1.64]		1.24 [0.93 - 1.64]		1.20
	T1 - T2	153	REF		REF		REF	0.007	REF	0.010	
N-stage	T3	744	1.29 [0.92 - 1.80]	0.15	1.23 [0.87 - 1.74]		1.22 [0.87 - 1.72]		1.22 [0.87 - 1.72]		1.19
	T4	180	1.96 [1.33 - 2.88]	<0.05	1.77 [1.18 - 2.65]		1.72 [1.16 - 2.56]		1.72 [1.16 - 2.56]		1.60
Mutational status	N0	399	REF		REF		REF	<0.005	REF	<0.005	
	N1	380	1.14 [0.88 - 1.48]	0.31	1.21 [0.93 - 1.58]		1.22 [0.94 - 1.59]		1.22 [0.94 - 1.59]		1.19
RAS/BRAF-wildtype	N2	298	1.77 [1.37 - 2.28]	<0.05	1.64 [1.26 - 2.13]		1.66 [1.28 - 2.15]		1.66 [1.28 - 2.15]		1.55
	RAS/BRAF-wildtype	505	REF		REF		REF	<0.005	REF	<0.005	
Number of liver metastases	BRAF-mt	44	2.08 [1.20 - 3.59]	<0.05	2.16 [1.22 - 3.82]		2.13 [1.21 - 3.76]		2.13 [1.21 - 3.76]		1.92
	RAS-mt	528	1.48 [1.16 - 1.88]	<0.05	1.64 [1.26 - 2.13]		1.67 [1.28 - 2.16]		1.67 [1.28 - 2.16]		1.55
Size of largest liver metastasis	Non-linearly		Non-linearly		Non-linearly		Non-linearly	<0.005	Non-linearly	<0.005	
	Non-linearly		Non-linearly		Non-linearly		Non-linearly	<0.005	Non-linearly	<0.005	
Pre-operative CEA	Non-linearly		Non-linearly		Non-linearly		Non-linearly	0.264	-		
	Non-linearly		Non-linearly		Non-linearly		Non-linearly	0.367	-		
Disease-free interval											

Specifications for univariable and multivariable cox regression analyses for all candidate predictors for EHRFS within 12 months after local treatment of CRLM are shown, including the full multivariable models, the pooled selection model ('EHR model'). For the pooled selection model, the apparent model hazard ratios and the overfitting-adjusted hazard ratios (shrunk) are shown ( $\beta_{\text{adjusted}} = \beta_{\text{unadjusted}} \times \text{shrinkage factor}$  obtained via bootstrapping during internal validation). For the full multivariable and selection models, multivariable Wald DJ  $p$ -values are shown. Continuous variables were modeled non-linearly using restricted cubic splines, for which the hazard ratios are shown in Supplementary Figure 5. The number of patients per category ( $n$ ) indicated are the pooled number of patients for each level over the imputed datasets.

**Abbreviations:**  $\beta$  (regression coefficient), CEA (carcinoembryonic antigen), CRLM (colorectal liver metastases), EHRFS (extrahepatic recurrence-free survival), HR (hazard ratio), N-stage (nodal stage of the primary tumour), REF (reference), T-stage (tumour stage of the primary tumour).

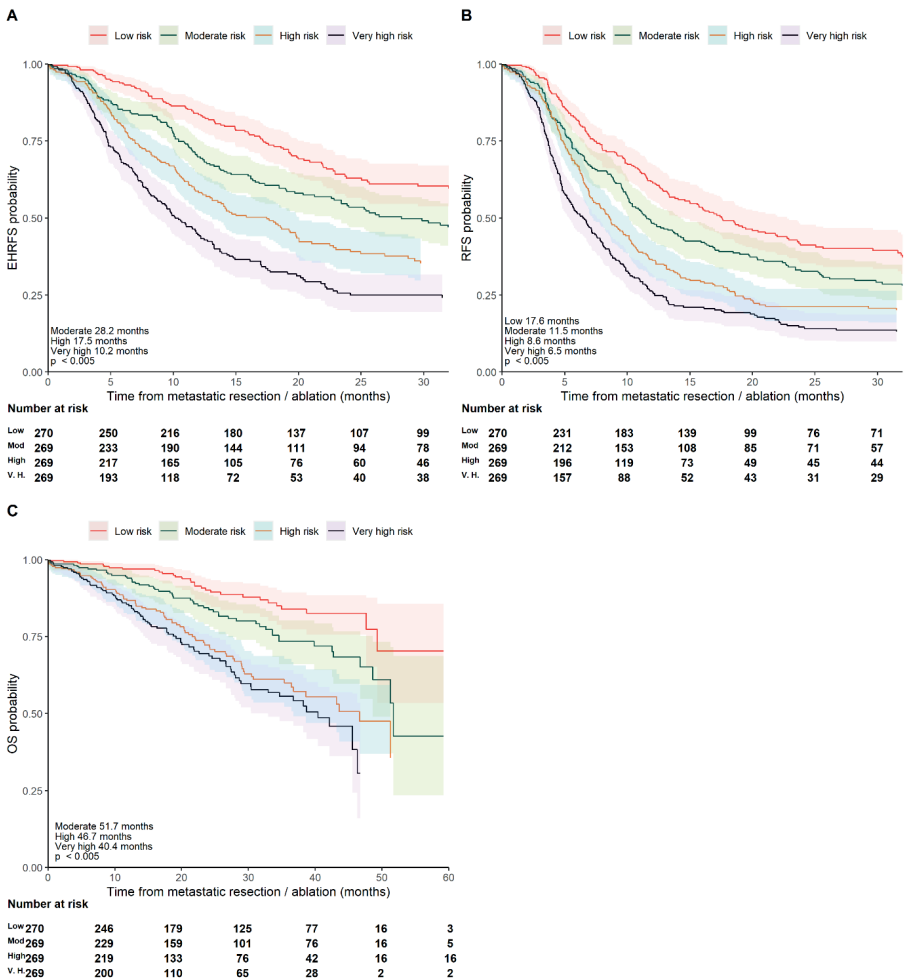
## EHR prediction model

Following AIC-informed backward selection, the model included 6/9 candidate predictor variables: sidedness of the primary tumour, T-stage, N-stage, *RAS/BRAF<sup>V600E</sup>* mutational status, and number and size of liver metastases (preoperative systemic treatment, preoperative CEA and DFI were excluded). The model's HR are shown in Table 2 (non-linear HR plots for continuous variables in Supplementary Figure 5). In an exploratory analysis, including an interaction term between *RAS* mutational status and preoperative systemic treatment did not significantly improve the model fit (Wald's D1 test,  $p=0.194$ ).

## Performance and validation of model

The model was able to discriminate for EHRFS, RFS and OS, based on quartiles of predicted EHR risk (Figure 1). The 6-month EHR rates in the low, intermediate, high and very high risk patient groups were 6% (95% C.I. 4-10), 15% (95% C.I. 11-20), 20% (95% C.I. 16-25) and 32% (95% C.I. 26-38), respectively. Likewise, the model showed good discrimination for RFS and OS, with significant differences in survival among the risk groups.

**Figure 1.** Kaplan Meier plots for EHRFS, RFS and OS according to quartiles of predicted EHR risk groups (low, moderate, high and very high)



**Fig 1.** Kaplan Meier curves for quartiles of predicted EHR risk (low, moderate, high and very high risk) patients are shown for three survival endpoints (EHRFS, RFS and OS). Predicted EHR risk includes EHR or death as an event for EHRFS. The values in the plots summarize the median survival and 95% confidence intervals and the p-value for the log-rank test. The survival probabilities were pooled over the imputed datasets after complementary log-log transformation. If for a given group the median survival was not reached, it is not reported.

**Abbreviations:** EHR (extrahepatic recurrence), EHRFS (extrahepatic recurrence-free survival), OS (overall survival), RFS (recurrence-free survival), V.H. (very high).

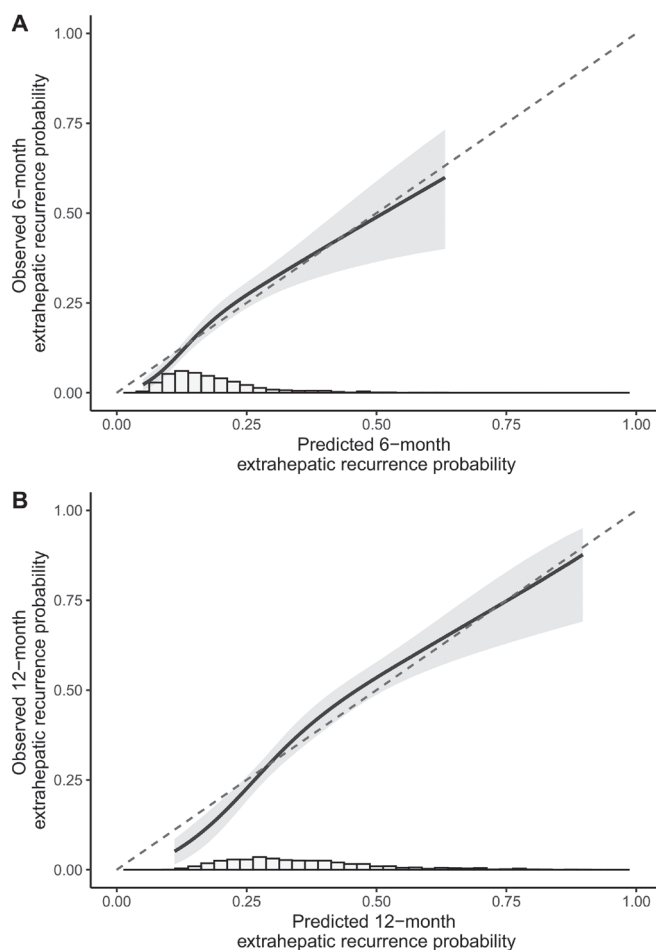
**Figure 2.** Calibration plot for extrahepatic recurrence

Fig 2. Calibration plots for the predicted 6-month (A) and 12-month (B) extrahepatic recurrence (EHR, which includes EHR or death as events for EHRFS) apparent probabilities versus the observed EHR probabilities are shown. The dashed line indicates perfect calibration, the solid line the observed EHR probabilities with in light gray the 95% confidence interval band. The histogram shows the distribution of the predicted EHR probabilities. The integrated calibration index is 0.015 (6-month EHR) and 0.028 (12-month EHR). The median absolute difference was 0.017 (6-month) and 0.030 (12-month), with a maximum absolute difference of 0.03 (6-month) and 0.06 (12-month).

*Abbreviations:* EHR (extrahepatic recurrence), EHRFS (extrahepatic recurrence-free survival).

The performance of the prediction model was assessed by calibration and discrimination. The estimated and observed risks for EHR or death were well-calibrated (Figure 2).

The observed/expected ratio was 1.015 (95% C.I. 0.911-1.120), which is close to 1. For discrimination, Harrell's C-index through 6 and 12-months was 0.663 (95% C.I. 0.624-0.702) and 0.661 (95% C.I. 0.632-0.689), respectively, and similar for Uno's C-index. The 6 and 12-month area under the time-dependent ROC curves was 0.668 (95% C.I. 0.626-0.709) and 0.671 (95% C.I. 0.636-0.707), respectively (Supplementary Figure 6). Nagelkerke's  $R^2$  was 0.094. The shrinkage factor obtained through internal validation was 0.86, with the shrunken HR in Table 2. The shrunken model yielded overoptimism-corrected 6-month risks for EHR or death between 5.9-56.0% (interquartile interval 12.9-22.0%). The optimism-adjusted Harrell's C-index through 6- and 12-months was 0.643 (95% C.I. 0.605-0.682) and 0.641 (95% C.I. 0.612-0.669). The full model specifications are shown in Appendix 1.

The model was further validated for generalizability by internal-external cross-validation using three geographical regions, which indicated that the models developed on the other regions showed adequate performance in each left-out geographical region, with some variation in the calibration slopes and observed/expected ratio, but less in the C-index (Supplementary Figure 7).

## Decision curve analysis analyzing net benefit when using model-guided CRLM treatment decisions

We examined the potential net benefit of the model for clinical decision-making regarding local treatment of CRLM through decision curve analysis. EHR model-guided treatment of CRLM (compared to non-informed decision-making by treating all or no patients) results in net benefit for patients for 6-month EHR risk thresholds of 0-40% and for 12-month EHR risk thresholds of 0-60% (Figure 3).

**Figure 3.** Decision curve analysis for informed decision-making by selecting patients for local CRLM treatment according to the model's predicted EHR probability

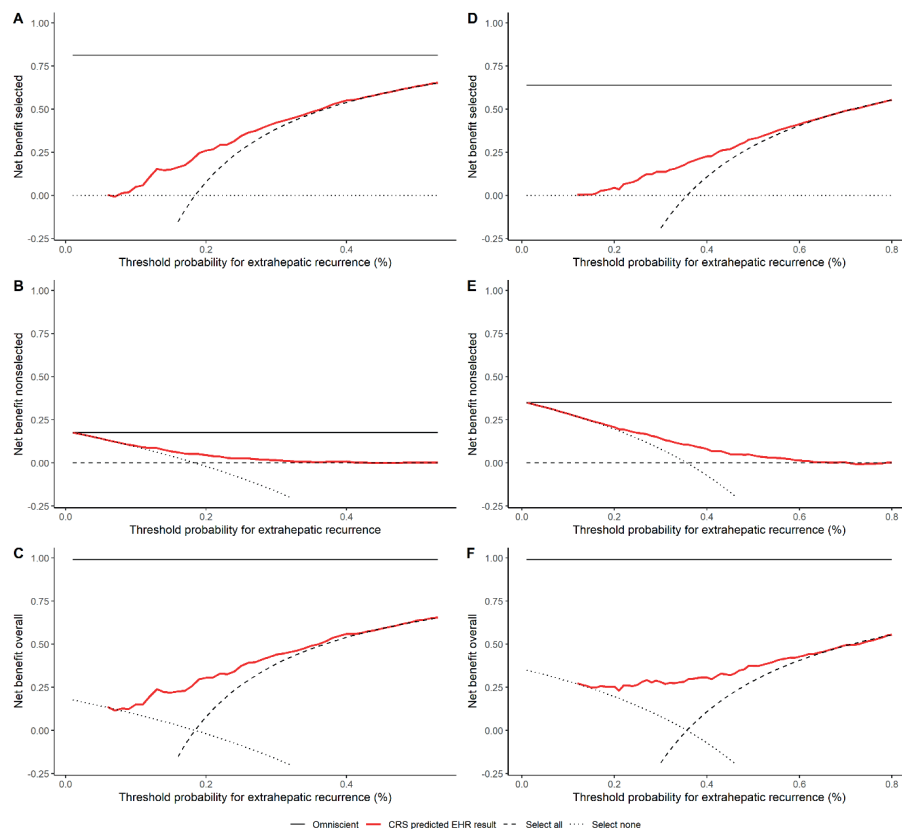


Fig 3. Decision curve analysis plots are shown indicating the net benefit obtained for a given threshold value for 6-month extrahepatic recurrence probability (A – C) and 12-month extrahepatic recurrence probability (D – F), which includes EHR or death as an EHRFS event. The net benefit was compared across 3 situations: non-informed decision-making (selecting all patients or selecting no patients (dashed and dotted lines, respectively)) and for informed decision-making by selecting patients for local CRLM treatment according to the model's predicted EHRFS probability (red line). As comparison, the black line represents an omniscient model (all knowing model). **A & D.** The net benefit of treating patients with local treatment for CRLM ('selected patients') is determined using the true positives (patients with predicted EHRFS probability ( $p_{\text{EHRFS}}$ ) above the threshold value and not having had an EHR) versus false positives ( $p_{\text{EHRFS}} > \text{threshold}$  and the patient did have an EHR) for a range of threshold values (0-1), with the benefit of false positives weighted relative to the threshold value. For consistency, the net benefit is shown for a range of thresholds for extrahepatic recurrence (extrahepatic recurrence probability =  $1 - \text{EHRFS probability}$ ). **B & E.** The net benefit of not treating patients with local treatment for CRLM ('nonselected patients') is determined using the true negatives (patients with  $p_{\text{EHRFS}} < \text{threshold}$  and having an EHR) versus false negatives (patients with  $p_{\text{EHRFS}} < \text{threshold}$  and not having had an EHR) for a range of threshold values (0-1), with the benefit of false negatives weighted relative to the threshold value. **C & F.** The overall net benefit is the sum of the net benefit of the selected and nonselected patients.

*Abbreviations:* EHR (extrahepatic recurrence), EHRFS (extrahepatic recurrence-free survival).

## DISCUSSION

We developed a prediction model for early EHR in a nationwide, population-based cohort of patients with local treatment of CRLM. The model incorporated tumour *RAS*/*BRAF*<sup>V600E</sup> mutational status and sidedness of primary tumour alongside traditional prognostic factors. Early EHR after local CRLM treatment is of major clinical importance and can be meaningfully predicted with routine clinical information. Our EHR prediction model discriminates between patients based on EHR rates, reflected in differing EHRFS, RFS and OS. The EHR prediction model's expected generalizability is good.

Prediction models are increasingly used and can facilitate shared risk-informed decision-making for interventions, manage patient expectations or select patients for inclusion in trials. However, clinical application of available prediction models for local CRLM treatment is hampered by lack of generalizability, loss of predictive performance by simplification of models (e.g. categorizing continuous variables) and low clinical utility<sup>37</sup>. Published models were developed to predict RFS and OS. With increasing possibilities for repeated CRLM resections of recurrences with favorable survival outcomes<sup>16,17</sup>, site-agnostic RFS and OS become a less relevant outcome for prediction models. Our study confirmed that about half of patients have a liver-limited first recurrence and experience long-term survival. Although RFS and OS are meaningful outcomes to manage expectations, EHRFS as outcome may guide clinical decisions for patients with CRLM.

To our knowledge, our model is the first to predict early EHR in patients after local treatment of CRLM. Local CRLM treatment should ideally be avoided in patients who experience an early EHR (18% of patients). These patients evidently have systemic disease, have a poor prognosis and are often not eligible for repeated local treatment<sup>18–23</sup>. The poor OS we demonstrated in patients with early EHR is comparable to the expected OS of mCRC patients undergoing palliative systemic treatment<sup>42</sup>. Patients at high risk for early EHR are therefore unnecessarily exposed to potential perioperative risks and may be harmed by delaying palliative systemic treatment. The EHR prediction model can be used to confirm that local treatment should be pursued in low-risk patients. However, it is currently difficult for the EHR prediction model to identify patients with a sufficiently high predicted risk which would justify avoiding local CRLM treatment. The EHR prediction model may also aid clinical decision-

making by identifying moderate/high-risk patients for early EHR who may benefit from perioperative systemic treatment. A treatment strategy for these patients may be to initiate long-lasting systemic treatment, and upon sustained response, perform local treatment of CRLM. Once externally validated, the EHR model will lend well for studies examining the optimal treatment by stratifying patients who are at moderate/high risk for early EHR.

The strength of our study is that our EHR prediction model was developed in a nationwide cohort of patients encompassing 39 academic, teaching and regional hospitals, and is thus representative for a general CRLM population undergoing local CRLM treatment. The cohort had minimal (<1%) loss to follow-up, likely not affecting its generalizability. Furthermore, the EHR prediction model included *RAS* and *BRAF*<sup>V600E</sup> mutational status, novel prognostic factors for outcomes after local treatment of CRLM. As not routinely available for all patients, we performed additional mutation analysis (resulting in 65% available) and handled remaining missing data by multiple imputation. *RAS* and *BRAF*<sup>V600E</sup> mutations are associated with an increased incidence of EHR<sup>29,30</sup>. Patients with *BRAF*<sup>V600E</sup> mutations and early unsalvageable recurrences have a poor survival after local treatment of CRLM<sup>13,29,43</sup>. Effective palliative systemic treatments are available for *BRAF*<sup>V600E</sup> mutated mCRC, further emphasizing the need to include *BRAF* mutational status in prediction models for local CRLM treatment. Only three prediction models included *RAS* and *BRAF* mutation status<sup>13–15</sup>, potentially due to the low prevalence of *BRAF* mutations in patients with local treatment of CRLM (approximately 2%)<sup>13</sup>. In contrast to previous studies<sup>28,44</sup>, there was no interaction between neoadjuvant treatment status and *RAS* mutational status in our cohort.

However, our study has some limitations. Firstly, patients in our cohort were selected based on primary tumour diagnosis in 2015 and 2016 with subsequent local treatments of CRLM until January 2019. Thus, our cohort does not include metachronous disease with a DFI beyond four years. Secondly, our prediction model does not and could not robustly specify site of recurrence in patients, which may be relevant since patients with lung-only recurrences might be able to undergo local treatment and experience long-term survival, although this practice is based on retrospective highly selected patient series with small numbers<sup>45,46</sup>. Although we were unable to externally validate our prediction model beyond the internal-external cross-validation, the full EHR prediction model specifications have been provided to facilitate external validation in other patient cohorts.

The performance of our model could be further improved by including additional promising histopathologic or tumour genetic features which may better identify high-risk patients<sup>15</sup>. Examples include distinct histopathological growth patterns, the Immunoscore (based on T-cell infiltration), a six-gene panel and liquid biopsies (detecting circulating tumour DNA)<sup>47-50</sup>. Incorporating these additional features in an updated prediction model for local CRLM treatment may help identify patients at sufficiently high-risk for early EHR to optimize the treatment strategy for these patients.

## CONCLUSION

We analysed population-based outcomes of patients after local treatment (resection and/or ablation) of CRLM. Early EHR is a valuable and informative alternative endpoint for prediction models. The EHR prediction model, including *RAS/BRAF<sup>V600E</sup>* mutation status and sidedness, showed robust performance in discriminating between patients based on EHR probability, reflected in differing EHRFS, RFS and OS. After further external validation, the EHR prediction model might offer guidance in clinical decision-making in patients with resectable CRLM.

## ACKNOWLEDGEMENTS

The authors thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry, as well as IKNL staff for scientific advice. We would like to acknowledge the NKI-AVL Core Facility Molecular Pathology & Biobanking (CFMPB) for supplying lab support and thank Willem Hoefakker and Pien van Diemen for mutational analysis and their contribution to the study. We would like to acknowledge the Sacha Swarttouw-Hijmans fund for their supporting grant.

## FUNDING

This research was supported by a grant by the Sacha Swarttouw-Hijmans Fund.

## REFERENCES

1. Zhou H., Liu Z., Wang Y., Wen X., Amador E.H., Yuan L., et al. Colorectal liver metastasis: molecular mechanism and interventional therapy. *Signal Transduct Target Ther.* (2022);7(70):1–25.
2. Kanas G.P., Taylor A., Primrose J.N., Langeberg W.J., Kelsh M.A., Mowat F.S., et al. Survival after liver resection in metastatic colorectal cancer: review and meta-analysis of prognostic factors. *Clin Epidemiol.* (2012);4:283–301.
3. Andreou A., Aloia T.A., Brouquet A., Dickson P. V, Zimmitti G., Maru D.M., et al. Margin status remains an important determinant of survival after surgical resection of colorectal liver metastases in the era of modern chemotherapy. *Ann Surg.* (2013);257(6):1079–88.
4. Bolhuis K., Kos M., van Oijen M.G.H., Swijnenburg R.J., Punt C.J.A. Conversion strategies with chemotherapy plus targeted agents for colorectal cancer liver-only metastases: A systematic review. *Eur J Cancer.* (2020);141:225–38.
5. de Jong M.C., Pulitano C., Ribero D., Strub J., Mentha G., Schulick R.D., et al. Rates and patterns of recurrence following curative intent surgery for colorectal liver metastasis: an international multi-institutional analysis of 1669 patients. *Ann Surg.* (2009);250(3):440–8.
6. Kopetz S., Chang G.J., Overman M.J., Eng C., Sargent D.J., Larson D.W., et al. Improved survival in metastatic colorectal cancer is associated with adoption of hepatic resection and improved chemotherapy. *J Clin Oncol.* (2009);27(22):3677–83.
7. Imai K., Allard M.A., Benitez C.C., Vibert E., Sa Cunha A., Cherqui D., et al. Early Recurrence After Hepatectomy for Colorectal Liver Metastases: What Optimal Definition and What Predictive Factors? *Oncologist.* (2016);21(7):887–94.
8. Okuno M., Goumard C., Kopetz S., Vega E.A., Joechle K., Mizuno T., et al. RAS Mutation is Associated with Unsalvageable Recurrence Following Hepatectomy for Colorectal Cancer Liver Metastases. *Ann Surg Oncol.* (2018);25(8):2457–66.
9. Fong Y., Fortner J., Sun R.L., Brennan M.F., Blumgart L.H. Clinical Score for Predicting Recurrence After Hepatic Resection for Metastatic Colorectal Cancer. *Ann Surg.* (1999);230(3):309.
10. Margonis G.A., Sasaki K., Gholami S., Kim Y., Andreatos N., Rezaee N., et al. Genetic And Morphological Evaluation (GAME) score for patients with colorectal liver metastases. *Br J Surg.* (2018);105(9):1210–20.
11. Paredes A.Z., Hyer J.M., Tsilimigras D.I., Moro A., Bagante F., Guglielmi A., et al. A Novel Machine-Learning Approach to Predict Recurrence After Resection of Colorectal Liver Metastases. *Ann Surg Oncol.* (2020);27(13):5139–47.
12. Brudvik K.W., Jones R.P., Giulianti F., Shindoh J., Passot G., Chung M.H., et al. RAS Mutation Clinical Risk Score to Predict Survival After Resection of Colorectal Liver Metastases. *Ann Surg.* (2019);269(1):120–6.
13. Buisman F.E., Giardiello D., Kemeny N.E., Steyerberg E.W., Höppener D.J., Galjart B., et al. Predicting 10-year survival after resection of colorectal liver metastases; an international study including biomarkers and perioperative treatment. *Eur J Cancer.* (2022);168:25–33.
14. Chen Y., Chang W., Ren L., Chen J., Tang W., Liu T., et al. Comprehensive Evaluation of Relapse Risk (CERR) Score for Colorectal Liver Metastases: Development and Validation. *Oncologist.* (2020);25(7):e1031–41.

15. Lang H., Baumgart J., Heinrich S., Tripke V., Passalacqua M., Maderer A., et al. Extended Molecular Profiling Improves Stratification and Prediction of Survival After Resection of Colorectal Liver Metastases. *Ann Surg.* (2019);**270**(5):799–805.
16. Nanji S., Tsang M.E., Wei X., Booth C.M. Outcomes after repeat hepatic resection for recurrent metastatic colorectal cancer: A population-based study. *Am J Surg.* (2017);**213**(6):1053–9.
17. Matsuoka H., Morise Z., Tanaka C., Hayashi T., Ikeda Y., Maeda K., et al. Repeat hepatectomy with systemic chemotherapy might improve survival of recurrent liver metastasis from colorectal cancer—a retrospective observational study. *World J Surg Oncol.* (2019);**17**(1):33.
18. Isoniemi H., Uutela A., Nordin A., Lantto E., Kellokumpu I., Ovissi A., et al. Centralized repeated resectability assessment of patients with colorectal liver metastases during first-line treatment: prospective study. *Br J Surg.* (2021);**108**(7):817–25.
19. Battula N., Tsapralis D., Mayer D., Isaac J., Muiesan P., Sutcliffe R.P., et al. Repeat liver resection for recurrent colorectal metastases: A single-centre, 13-year experience. *Hpb.* (2014);**16**(2):157–63.
20. Butte J.M., Gönen M., Allen P.J., Peter Kingham T., Sofocleous C.T., DeMatteo R.P., et al. Recurrence After Partial Hepatectomy for Metastatic Colorectal Cancer: Potentially Curative Role of Salvage Repeat Resection. *Ann Surg Oncol.* (2015);**22**(8):2761–71.
21. Malik H.Z., Gomez D., Wong V., Al-Mukthar A., Toogood G.J., Lodge J.P.A., et al. Predictors of early disease recurrence following hepatic resection for colorectal cancer metastasis. *Eur J Surg Oncol.* (2007);**33**(8):1003–9.
22. Viganò L., Capussotti L., Lapointe R., Barroso E., Hubert C., Giuliani F., et al. Early recurrence after liver resection for colorectal metastases: risk factors, prognosis, and treatment. A LiverMetSurvey-based study of 6,025 patients. *Ann Surg Oncol.* (2014);**21**(4):1276–86.
23. Stewart C.L., Warner S., Ito K., Raoof M., Wu G.X., Kessler J., et al. Cytoreduction for colorectal metastases: liver, lung, peritoneum, lymph nodes, bone, brain. When does it palliate, prolong survival, and potentially cure? *Curr Probl Surg.* (2018);**55**(9):330–79.
24. Nordlinger B., Sorbye H., Glimelius B., Poston G.J., Schlag P.M., Rougier P., et al. Perioperative FOLFOX4 chemotherapy and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC 40983): long-term results of a randomised, controlled, phase 3 trial. *Lancet Oncol.* (2013);**14**(12):1208–15.
25. Kanemitsu Y., Shimizu Y., Mizusawa J., Inaba Y., Hamaguchi T., Shida D., et al. Hepatectomy Followed by mFOLFOX6 Versus Hepatectomy Alone for Liver-Only Metastatic Colorectal Cancer (JCOG0603): A Phase II or III Randomized Controlled Trial. *J Clin Oncol.* (2021);**39**(34):3789–99.
26. Buisman F.E., Galjart B., Buettner S., Groot Koerkamp B., Grunhagen D.J., Verhoef C. Primary tumor location and the prognosis of patients after local treatment of colorectal liver metastases: a systematic review and meta-analysis. *HPB.* (2020);**22**(3):351–7.
27. Russolillo N., Sperti E., Langella S., Menonna F., Allietta A., Di Maio M., et al. Impact of primary tumor location on patterns of recurrence and survival of patients undergoing resection of liver metastases from colon cancer. *Hpb.* (2020);**22**(1):116–23.
28. Takeda Y., Mise Y., Takahashi Y., Ito H., Inoue Y., Yoshioka R., et al. Limited Prognostic Value of KRAS in Patients Undergoing Hepatectomy for Colorectal Liver Metastases. *Ann Surg Oncol.* (2022);**29**(4):2383–91.

29. Kobayashi S., Takahashi S., Takahashi N., Masuishi T., Shoji H., Shinozaki E., et al. Survival Outcomes of Resected BRAF V600E Mutant Colorectal Liver Metastases: A Multicenter Retrospective Cohort Study in Japan. *Ann Surg Oncol.* (2020);**27**(9):3307–15.
30. Schirripa M., Bergamo F., Cremolini C., Casagrande M., Lonardi S., Aprile G., et al. BRAF and RAS mutations as prognostic factors in metastatic colorectal cancer patients undergoing liver resection. *Br J Cancer.* (2015);**112**(12):1921–8.
31. (IKNL) N.C.C.O. Netherlands Cancer Registry (NCR) [Internet]. Available from: <https://iknl.nl/en/ncr>
32. Moons K.G.M., Altman D.G., Reitsma J.B., Ioannidis J.P.A., Macaskill P., Steyerberg E.W., et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med.* (2015);**162**(1):W1–73.
33. Reddy S.K., Barbas A.S., Turley R.S., Steel J.L., Tsung A., Marsh J.W., et al. A standard definition of major hepatectomy: resection of four or more liver segments. *HPB.* (2011);**13**(7):494–502.
34. Mekenkamp L.J., Koopman M., Teerenstra S., van Krieken J.H., Mol L., Nagtegaal I.D., et al. Clinicopathological features and outcome in advanced colorectal cancer patients with synchronous vs metachronous metastases. *Br J Cancer.* (2010);**103**(2):159–64.
35. Nederlandse Vereniging voor Heelkunde. Colorectaal Carcinoom (CRC) (version 2019) [Internet]. Available from: [file:///H:/Downloads/Colorectaal\\_carcinoom\\_CRC\\_\(1\).pdf](file:///H:/Downloads/Colorectaal_carcinoom_CRC_(1).pdf)
36. Casparie M., Tiebosch A.T., Burger G., Blauwgeers H., van de Pol A., van Krieken J.H., et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol.* (2007);**29**(1):19–24.
37. Bolhuis K., Wensink G.E., Elferink M.A.G., Bond M.J.G., Dijksterhuis W.P.M., Fijneman R.J.A., et al. External Validation of Two Established Clinical Risk Scores Predicting Outcome after Local Treatment of Colorectal Liver Metastases in a Nationwide Cohort. *Cancers (Basel).* (2022);**14**(10):2356.
38. Riley R.D., Snell K.I.E., Ensor J., Burke D.L., Harrell F.E., Moons K.G.M., et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med.* (2019);**38**(7):1276–96.
39. Yamashita S., Chun Y.S., Kopetz S.E., Vauthey J.N. Biomarkers in colorectal liver metastases. *Br J Surg.* (2018);**105**(6):618–27.
40. van Buuren S., Groothuis-Oudshoorn K., Vink G., Doove L., Jolani S., Schouten R., et al. mice: Multivariate Imputation by Chained Equations in R. R package. *J Stat Softw.* (2011);**45**(3):1–67.
41. Rousson V., Zumbo T. Decision curve analysis revisited: Overall net benefit, relationships to ROC curve analysis, and application to case-control studies. *BMC Med Inform Decis Mak.* (2011);**11**(45):1–9.
42. Cremolini C., Antoniotti C., Rossini D., Lonardi S., Loupakakis F., Pietrantonio F., et al. Upfront FOLFOXIRI plus bevacizumab and reintroduction after progression versus mFOLFOX6 plus bevacizumab followed by FOLFIRI plus bevacizumab in the treatment of patients with metastatic colorectal cancer (TRIBE2): a multicentre, open-label, phase 3, randomised, controlled trial. *Lancet Oncol.* (2020);**21**(4):497–507.

43. Uutela A., Osterlund E., Halonen P., Kallio R., Ålgars A., Salminen T., et al. Resectability, conversion, metastasectomy and outcome according to RAS and BRAF status for metastatic colorectal cancer in the prospective RAXO study. *Br J Cancer*. (2022);
44. Margonis G.A., Kim Y., Sasaki K., Samaha M., Buettner S., Amini N., et al. Activating KRAS mutation is prognostic only among patients who receive preoperative chemotherapy before resection of colorectal liver metastases. *J Surg Oncol*. (2016);**114**(3):361–7.
45. Shah S.A., Haddad R., Al-Sukhni W., Kim R.D., Greig P.D., Grant D.R., et al. Surgical resection of hepatic and pulmonary metastases from colorectal carcinoma. *J Am Coll Surg*. (2006);**202**(3):468–75.
46. Neeff H., Hörth W., Makowiec F., Fischer E., Imdahl A., Hopt U.T., et al. Outcome after resection of hepatic and pulmonary metastases of colorectal cancer. *J Gastrointest Surg*. (2009);**13**(10):1813–20.
47. Vigano L., Banciforte B., Laurenti V., Costa G., Procopio F., Cimino M., et al. The Histopathological Growth Pattern of Colorectal Liver Metastases Impacts Local Recurrence Risk and the Adequate Width of the Surgical Margin. *Ann Surg Oncol*. (2022);S709–10.
48. Wang R., Dai W., Gong J., Huang M., Hu T., Li H., et al. Development of a novel combined nomogram model integrating deep learning-pathomics, radiomics and immunoscore to predict postoperative outcome of colorectal cancer lung metastasis patients. *J Hematol Oncol*. (2022);**15**(1):4–9.
49. Wada Y., Shimada M., Morine Y., Ikemoto T., Saito Y., Baba H., et al. A transcriptomic signature that predicts cancer recurrence after hepatectomy in patients with colorectal liver metastases. *Eur J Cancer*. (2022);**163**:66–76.
50. Bidard F.C., Kiavue N., Ychou M., Cabel L., Stern M.H., Madic J., et al. Circulating tumor cells and circulating tumor DNA detection in potentially resectable metastatic Colorectal cancer: A prospective ancillary study to the Unicancer Prodige-14 Trial. *Cells*. (2019);**8**(516):1–13.

## SUPPLEMENTARY MATERIAL

**Supplementary Table 1.** Assumptions regarding systemic treatment and survival outcomes

<b>Assumptions regarding progression of disease and survival:</b>
Date of new episode is considered as time of progression.
When disease progression is documented < 14 days of liver resection we assume this was part of the liver resection and first new episode is considered as time of progression.
Recurrence-free survival (RFS) is calculated from date of first liver procedure to date of progression. In case of 2-stage resection, RFS is calculated from last liver procedure.
RFS is calculated from date of first liver procedure to date of progression. In case of 2-stage resection, RFS is calculated from last liver procedure.
If no recurrence is registered:
If end of follow-up is registered and reason end of follow-up is: death, then date of death is registered as event of RFS;
If end of follow-up is registered and reason end of follow-up is other than death then RFS is censored on date of end of follow up;
If no date of end of follow-up is registered then RFS is censored on date of last visit;
If none of these dates are registered then RFS is documented as missing.
Extrahepatic recurrence-free survival (EHRFS) is calculated from date of first liver procedure to date of progression. In case of 2-stage resection, EHRFS is calculated from last liver procedure.
If no extrahepatic recurrence is registered:
If end of follow-up is registered and reason end of follow up is: death, then date of death is registered as event of EHRFS;
If end of follow-up is registered and reason end of follow up is other than death then EHRFS is censored on date of end of follow up;
If no date of end of follow-up is registered then EHRFS is censored on date of last visit;
If no last visit is registered but event for RFS is registered then EHRFS is censored on date of RFS;
If none of these dates are registered then EHRFS is documented as missing.
Extrahepatic disease was defined as presence of disease outside the liver or metastasectomy outside the liver.
Lymph node metastases registered as abdominal lymph nodes at time of first liver metastases were considered extrahepatic disease and as so classified as not-liver only disease.
Overall survival (OS) is calculated from date of diagnosis of metastatic disease.
Patients who did not die are censored on the date last known to be alive in the GBA (the municipal population registry).
OS after resection is calculated from date of first liver procedure. In case of 2-stage resection, OS is calculated from date of last liver procedure.
Primary tumour nodal stage was defined primarily on pathologic N-stage. When pN stage was missing, cN stage (radiological) was used.
If number of metastases was not given and code 77 was used (accounting for diffuse metastatic disease in the liver) then number of metastases was scored as 20.
If performance status was missing, this was scored as 0-1 because patients were considered physically good enough for resection.
<b>Assumptions regarding systemic treatment regimens and strategies:</b>
Systemic treatment includes both chemotherapy and/or targeted therapy.
A combination regimen is defined as all systemic agents starting within 4 weeks after start of the first agent and started before progression of disease.

If bevacizumab was started more than 4 weeks after the start of the first agent but before stop of this agent and before progression of disease, we assume bevacizumab was part of this combination regimen.

If a treatment line continues despite of progression, e.g., in case of reintroduction of the same or an equivalent regimen after a therapy break and detected progression, we regard this as continuation of the same treatment line.

If oxaliplatin only is registered, we assume this was part of a capecitabine and oxaliplatin (CAPOX) regimen of which capecitabine was not registered, so we add capecitabine. We assume this is due to a registration error, in which the administration of capecitabine has not been noticed by the data manager.

Systemic therapy was considered adjuvant systemic therapy for primary tumour when started < 12 weeks after resection of primary tumour and started before diagnosis of metastases in patients with metachronous disease.

Capecitabine monotherapy was considered radiosensitizer for primary tumour when started before primary tumour resection and before diagnosis of metastases and with notification to have received chemoradiotherapy.

Systemic therapy was considered pre-operative therapy (neo-adjuvant or induction) before liver resection when the therapy ended within 120 days before liver resection. Adjuvant therapy after resection of primary tumour or chemotherapy as radiosensitizer was excluded.

Systemic therapy was considered adjuvant therapy after liver resection when the therapy started within 120 days after liver resection. Chemotherapy as radiosensitizer was excluded.

Systemic therapy was considered peri-operative therapy of liver resection when the systemic therapy was given <120 days before and < 120 days after liver resection

When systemic therapy was given between two liver procedures before progression of disease, the first liver procedure was considered as staging procedure and systemic therapy was considered as pre-operative systemic therapy (neo-adjuvant or induction) for surgery 2

Treatment strategies are categorized as follows:

Treatment regimens containing chemotherapy, without targeted therapy, subdivided in: monotherapy (1 chemotherapy agent), doublets (2 chemotherapy agents) and triplets (3 chemotherapy agents);

Treatment regimens containing targeted therapy with or without chemotherapy, subdivided in: bevacizumab-containing regimens, and anti-EGFR targeted therapy-containing regimens.

Systemic therapy regimens are categorized as follows:

Fluoropyrimidine monotherapy (e.g. 5-fluorouracil [5-FU], capecitabine);

Oxaliplatin-based doublet therapy (e.g. capecitabine + oxaliplatin (CAPOX), 5-FU/oxaliplatin [FOLFOX]);

Irinotecan-based doublet therapy (e.g. capecitabine + irinotecan (CAPIRI), 5-FU/irinotecan [FOLFIRI])

Triplet systemic therapy (5-fluorouracil [5-FU], oxaliplatin and irinotecan)

Targeted therapy (anti-EGFR therapy; cetuximab or panitumumab, and bevacizumab)

A treatment line is defined as systemic therapy (monotherapy or combination regimen) administered at the same time until suspension, regardless of reason for discontinuation.

Treatment is considered as next line if an agent of a new drug group is started that is not applied in the previous systemic treatment regimen.

If the same or an equivalent systemic treatment regimen is (re)started, this is considered continuation of the same treatment line, e.g., CAPOX to FOLFOX.

#### **Assumptions regarding mutational status:**

*RAS* and *BRAF* mutation are considered mutual exclusive, therefore patients with *RAS* mutations or *BRAF* mutations, were assumed to have *BRAF*-wildtype or *RAS*-wildtype status, retrospectively.

---

**Assumptions regarding systemic treatment lines:**

---

A treatment line is defined as systemic therapy (monotherapy or combination regimen) administered at the same time until suspension, regardless of reason for discontinuation.

Treatment is considered as next line if an agent of a new drug group is started that is not applied in the previous systemic treatment regimen.

If the same or an equivalent systemic treatment regimen is (re)started, this is considered continuation of the same treatment line, e.g. CAPOX to FOLFOX.

---

**Assumptions regarding local treatment**

---

Local treatments are categorized as follows: 1 stage (1 procedure); 2-stage (2 procedures < 120 days apart)

R-status:

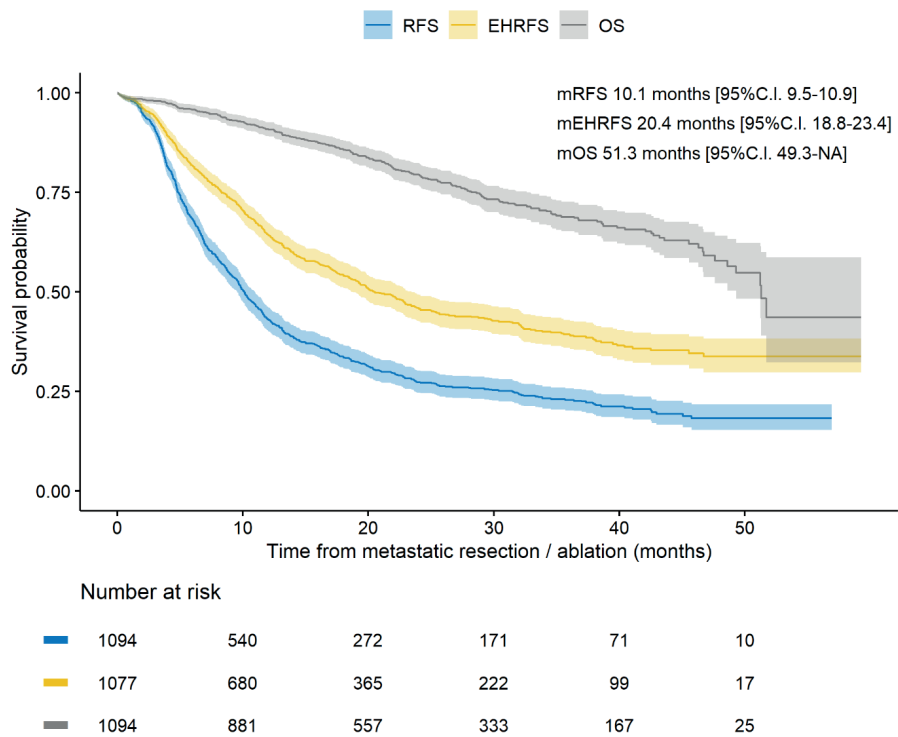
when two stage procedure and first procedure was R2 resection and second procedure was R1/R0 resection then 2-stage resection considered as R-status of last procedure.

when 2-stage resection and one procedure was R1 resection and other local treatment was R0 resection then considered as R1 resection.

---

*Abbreviations:* 5-FU (5-fluorouracil), CAPIRI (capecitabine + irinotecan), CAPOX (capecitabine + oxaliplatin), cN (radiological nodal-stage), EGFR (epidermal growth factor receptor), EHRFS (extrahepatic recurrence-free survival), FOLFIRI (5-FU + irinotecan), FOLFOX (5-FU + oxaliplatin), GBA (the municipal population registry), OS (overall survival), pN (pathological nodal-stage), RFS (recurrence-free survival).

**Supplementary Figure 1.** Kaplan-Meier analysis showing RFS, EHRFS and OS curves of the total cohort.



Suppl. Fig. 1. Kaplan Meier survival curves with 95% confidence intervals are shown for RFS, EHRFS and OS for the whole cohort. The number of patients at risk is indicated in the risk table.

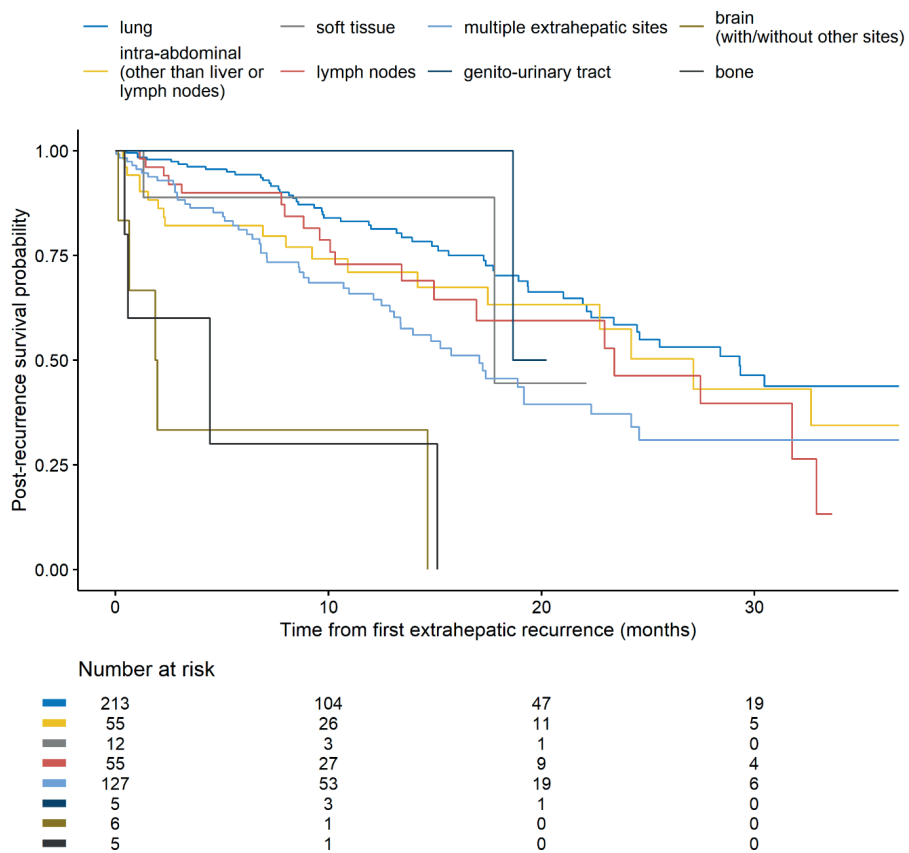
*Abbreviations:* EHRFS (extrahepatic recurrence-free survival), mEHRFS (median EHRFS), mOS (median OS), mRFS (median RFS), OS (overall survival), RFS (recurrence-free survival).

**Supplementary Table 2.** Detailed information about first recurrence and first extrahepatic recurrence

	<b>Total cohort</b> <b><i>n</i> = 1105</b> <b><i>n</i> (%)</b>
<b>RFS data</b>	
<b>Event</b>	
No	287 (26.0)
Yes	807 (73.0)
Recurrence	765
Death	42
Missing	11 (1.0)
<b>Site of first recurrence, <i>n</i> = 765</b>	
Liver-only	332 (43.3)
Extrahepatic	399 (52.2)
Missing site of recurrence	34 (4.4)
<b>EHRFS data</b>	
<b>Extrahepatic event during follow up</b>	
No	520 (48.3)
Yes	557 (51.7)
Recurrence	478
Death	79
Missing	28 (2.5)
<b>Site of first extrahepatic recurrence, <i>n</i> = 478</b>	
Lung	213 (44.6)
Intra-abdominal	55 (11.5)
Lymph nodes	55 (11.5)
Bone	5 (1.0)
Genito-urinary tract	5 (1.0)
Soft tissue	12 (2.5)
Brain with/without other sites	6 (1.3)
Multiple extrahepatic sites	127 (26.6)

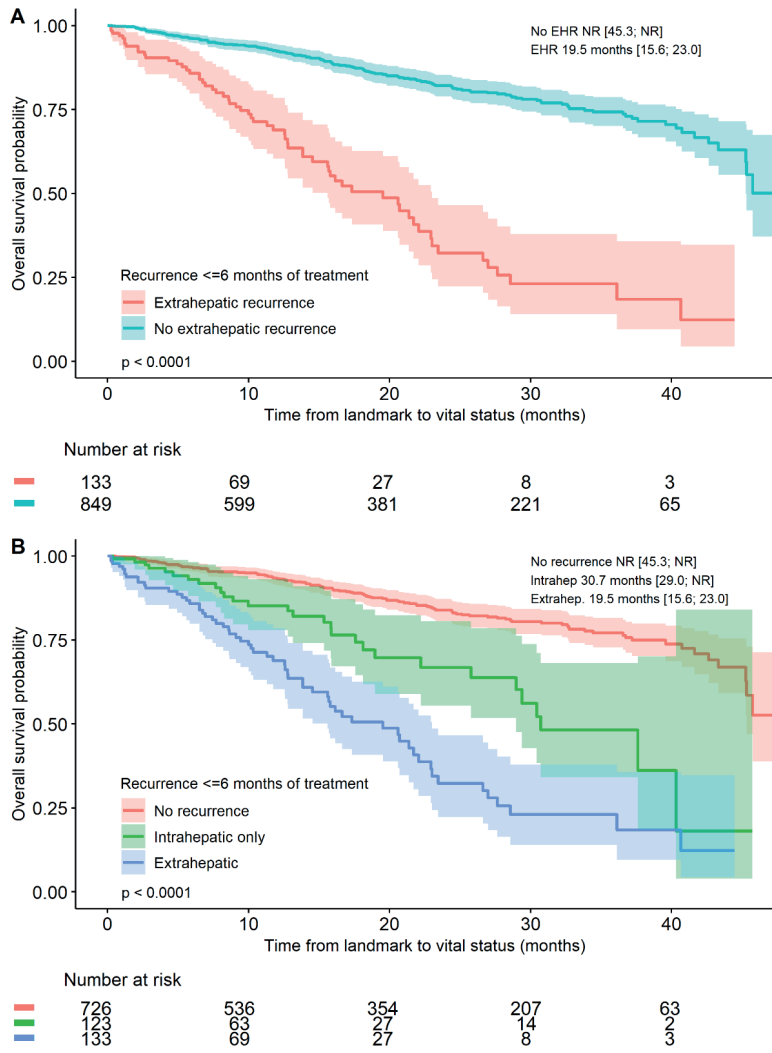
*Abbreviations:* *n* (count), RFS (recurrence-free survival), EHRFS (extrahepatic recurrence-free survival).

**Supplementary Figure 2.** Post-recurrence overall survival of patients according to site of extrahepatic recurrence



Suppl. Fig. 2. A KM plot of post-recurrence survival in patients according to the first site of extrahepatic recurrence. Categories indicate the site of extrahepatic metastasis, but may also include hepatic localization. A log-rank test for the post-recurrence survival probability per site of extrahepatic recurrence was performed ( $p<0.0001$ ). A log-rank test for the time to recurrence per site of extrahepatic recurrence was performed ( $p=0.55$ ).

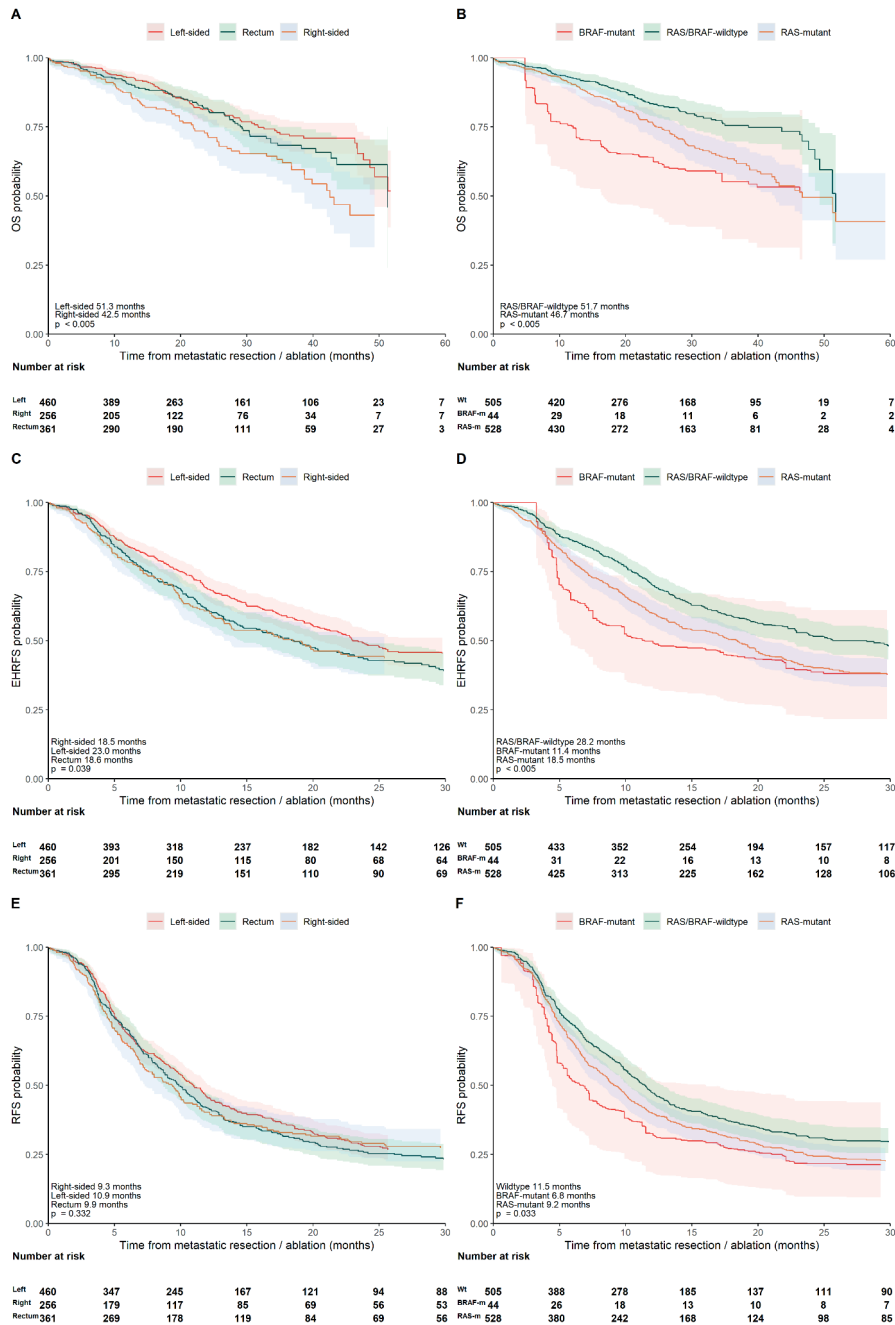
**Supplementary Figure 3.** Landmark analysis at six months showing Kaplan-Meier curves of patients with extrahepatic, intrahepatic-only and no recurrences within six months



Suppl. Fig. 3. A landmark analysis Kaplan Meier survival plot is shown, indicating the overall survival of patients after the landmark point (6 months after local treatment of CRLM), according to the site of recurrence which patients had experienced within 6 months of local treatment of CRLM. A. The 2 groups are extrahepatic recurrence (which includes death as an event) versus no extrahepatic recurrence. B. The 3 recurrence site groups are: no recurrence, intrahepatic recurrence only, extrahepatic recurrence (including  $n=100$  intra + extrahepatic recurrence). The log-rank  $p$ -value is indicated in the plot along with the observed median survival and 95% confidence intervals.

*Abbreviations:* EHR (extrahepatic recurrence), n (count), NR (not reached).

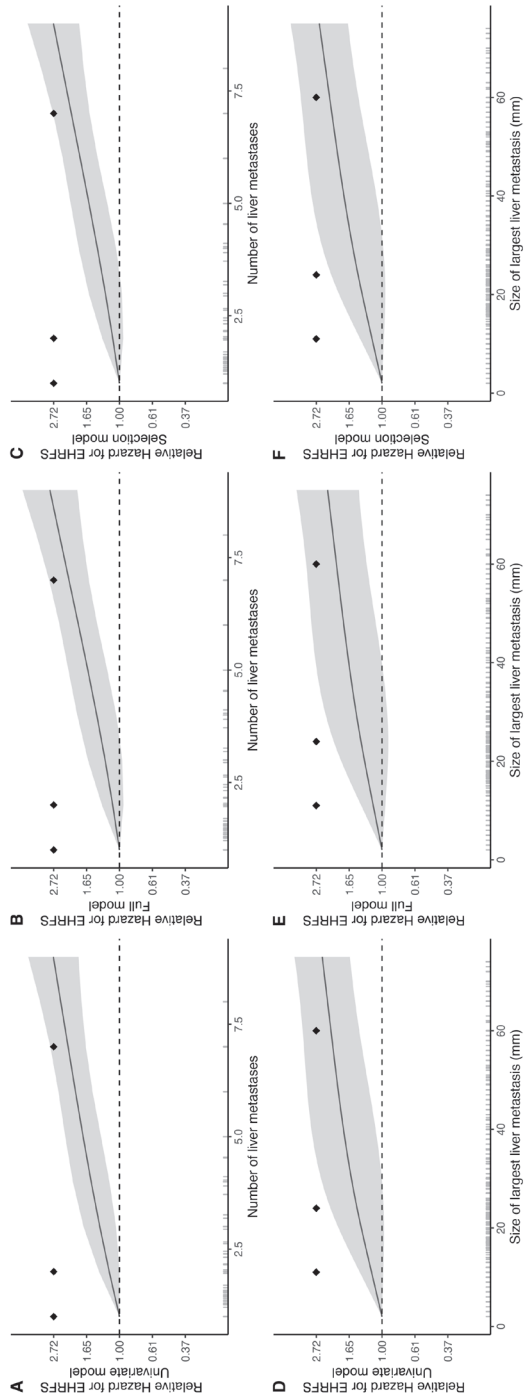
**Supplementary Figure 4.** Kaplan-Meier curves describing post-resection overall survival, recurrence-free survival and extrahepatic recurrence-free survival in the total cohort according to location of primary tumour and RAS/BRAF mutational status.



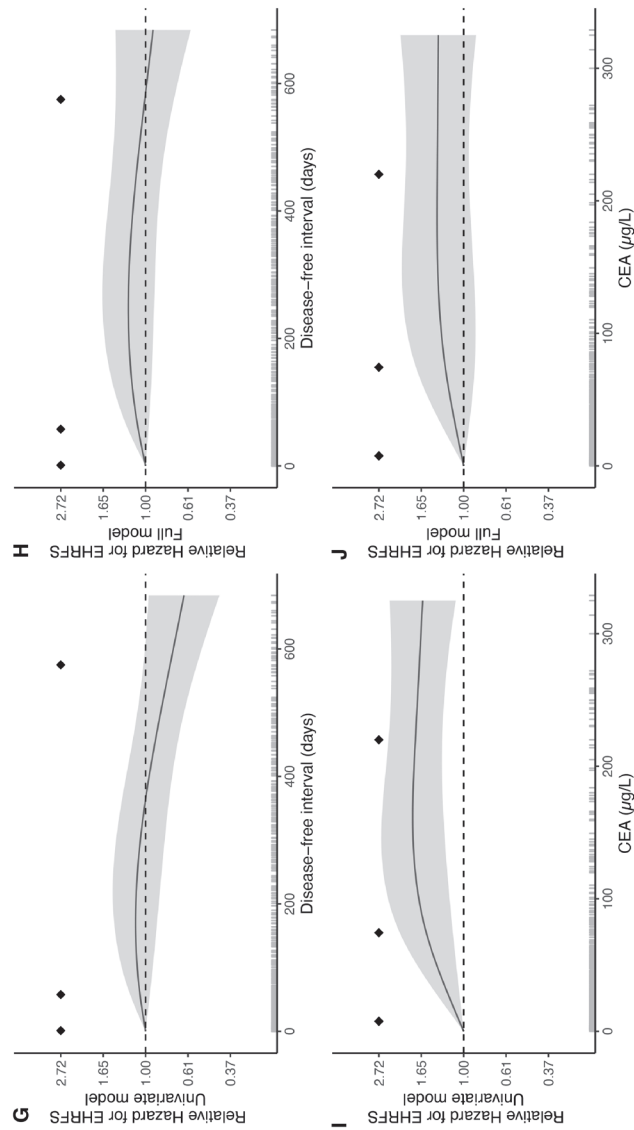
► Suppl. Fig. 4. Kaplan-Meier curves describing post-resection overall survival (A and B), and recurrence-free survival (C and D) and extrahepatic recurrence-free survival (E and F) in the total cohort according to location of primary tumour (A, C, D) and *RAS/BRAF* mutational status (B, D, F) using the imputed dataset. The observed median survival and 95% confidence intervals are indicated in the plot and not indicated if not reached within the follow-up period.

*Abbreviations:* *BRAF*-m (*BRAF*-mutant), EHRFS (extrahepatic recurrence-free survival), mEHRFS (median EHRFS), mOS (median OS), mRFS (median RFS), OS (overall survival), *RAS*-m (*RAS*-mutant), RFS (recurrence-free survival), Wt (*RAS/BRAF*-wildtype).

Supplementary Figure 5. Hazard ratio for EHRFS for continuous variables modeled non-linearly using restricted cubic splines



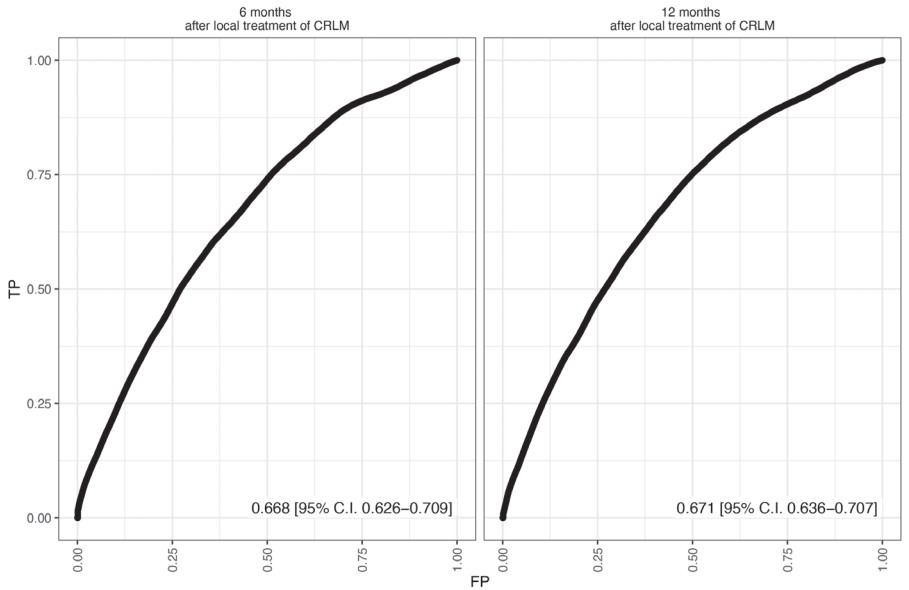
Supplementary Figure 5 (continued).



Suppl. Fig. 5. The pooled relative hazard and 95% confidence interval ribbon for EHRFS within 12 months from local treatment of CRLM according to one of three models (univariable, full multivariable and selection multivariable) are shown in the y-axis relative to the observed value of the continuous variable plotted on the x-axis. Each column indicates a model type with the univariable results in plots A, D, G, I; the full multivariable results in plots B, E, H, J; and the selection multivariable results in plots C and F (disease-free interval and pre-operative CEA were not included in the selection model). Each row shows a continuous variable, with number of liver metastases (A – C), size of largest liver metastasis (D – F), disease-free interval (G–H) and pre-operative CEA values (I–J). All 4 continuous variables were modeled using restricted cubic splines analysis with 3 knot positions (positions are indicated by the dots in the plot). The frequency of the observed values are indicated along the x-axis.

Abbreviations: CEA (carcinoembryonic antigen), CRLM (colorectal liver metastasis), EHRFS (extrahepatic recurrence-free survival), mm (millimeter), µg/L (microgram per liter).

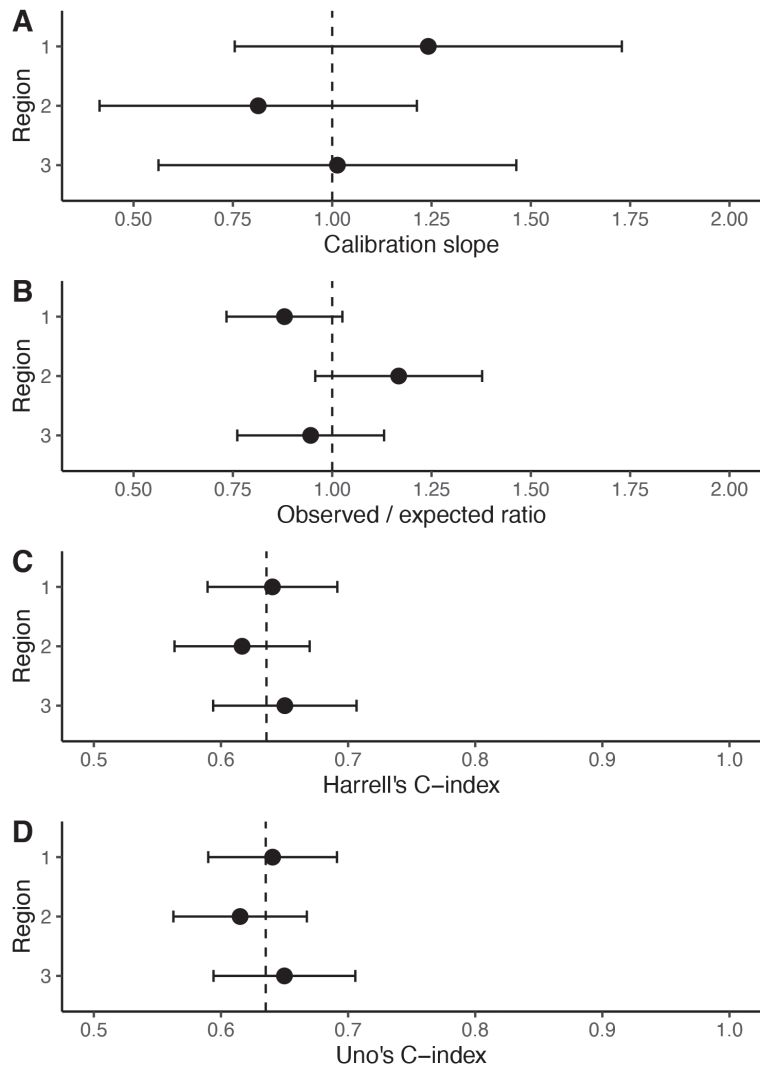
**Supplementary Figure 6.** Time-dependent ROC curve for the EHRFS model indicating the true positives and false positives 6 and 12 months after local treatment for CRLM



Suppl. Fig. 6. In a cumulative case/ dynamic control ROC analysis, the time-dependent receiver operator curve is shown with on the y-axis the true positive and the x-axis the false positive based on the model's linear predictor for each individual compared to the observed EHRFS at the given timepoint. The plot indicates how well the model predicts the survival time for the patients for 6 and 12 months after local treatment for CRLM, respectively. The AUC with 95% C.I. is indicated in each plot. The confidence intervals for AUC were calculated using 1000 bootstrap samples.

*Abbreviations:* AUC (area under the (ROC) curve), C.I. (confidence interval), CRLM (colorectal liver metastasis), FP (false positive), ROC (receiver operator curve), TP (true positive).

**Supplementary Figure 7.** Internal-external cross-validation results



Suppl. Fig. 7. The results for internal-external cross-validation are shown, including the calibration slope (A), observed/expected ratio (B), Harrell's C-index (C) and Uno's C-index (D). We used internal-external cross-validation to evaluate the generalizability of the model. The data were split in three geographic regions and all modeling steps including backward selection of variables and internal validation were repeated in two of three regions, after which the performance of the overfitting-adjusted model was evaluated in the left-out geographical region (C-index, calibration slope and intercept). Each geographic region was left-out of model development once, resulting in three estimates of external validation. The reference line in plot C & D indicates the mean C-index.

*Abbreviations:* C-index (calibration-index).

## Supplemental Methods

### *Early EHR as primary endpoint for the prediction model*

By consensus of experts in the field (JR, RJS, KB, EW, JH, MK, CJAP), early EHR (within six months, conform previous publications<sup>1,2</sup>) was defined as the clinically relevant primary endpoint of the model, due to the poor prognosis in patients with early EHR and lower chance of repeat local treatment, in contrast to patients with liver-only recurrences. Thus, the added value of local treatment of CRLM may not be justified in patients with a rapid EHR after local treatment of CRLM by consensus of experts in the field.

### *Statistical analysis*

We used standard descriptive statistics to describe baseline characteristics of the study population, including medians and interquartile intervals (IQI) for continuous variables, and frequency and percentages for categorical variables. Follow-up data and patient outcomes were described using (reverse) Kaplan Meier approaches. We assessed the prognostic impact of our primary endpoint occurrence of EHR  $\leq 6$  months after CRLM treatment using landmark analysis - which prevents immortal time bias - at six months after CRLM treatment and comparing the subsequent survival outcomes of three groups based on site of recurrence  $\leq 6$  months: no recurrence, intrahepatic only and EHR (which includes patients with intra- and extrahepatic recurrences).

We applied the recommendations published by Riley *et al.*<sup>3</sup>, to determine the number and complexity of the candidate predictors (together amounting to the number of coefficients) to be evaluated in our prediction model. We used the C-index for the Comprehensive Evaluation of Relapse Risk (CERR) score<sup>4</sup> as the anticipated minimum C-index for our model, since it most closely represents our primary end-point. With an expected C-index of  $\geq 0.695$ , and the observed EHR event rate within 12 months in our cohort, we had sufficient data to model 17 coefficients and fulfill the 3 criteria set by Riley *et al.*<sup>3</sup>.

Predictors were selected by assessment of a multidisciplinary team based on factors used in previous prediction models<sup>5-8</sup> and newly recognized prognostic factors<sup>9,10</sup>. Candidate predictors were blindly selected, prior to having analyzed the data. Nine candidate predictors were selected for model development, including 4 continuous variables that we aimed to model using three-knot restricted cubic splines (rcs) to allow for non-linearity (resulting in 17 coefficients): neoadjuvant systemic treatment, primary tumour location (left-sided, right-sided, rectum), T-status (T1-2, T3, T4), N-status (N0, N1, N2), *RAS/BRAF* mutational status (*RAS/BRAF*-wildtype, *RAS*-

mutant, *BRAF*-mutant), number of liver metastases (continuous), size of largest liver metastasis (continuous), pre-operative CEA (continuous) and DFI (continuous). Continuous variables were Winsorized at the 95<sup>th</sup> percentile before analyses to decrease influential points. As certain candidate predictors had missing data, and merely removing patients with missing data leads to loss of information and potentially also to selection bias, we used multiple imputation using multivariate imputation by chained equations (MICE)<sup>11</sup>, assuming missingness at random.

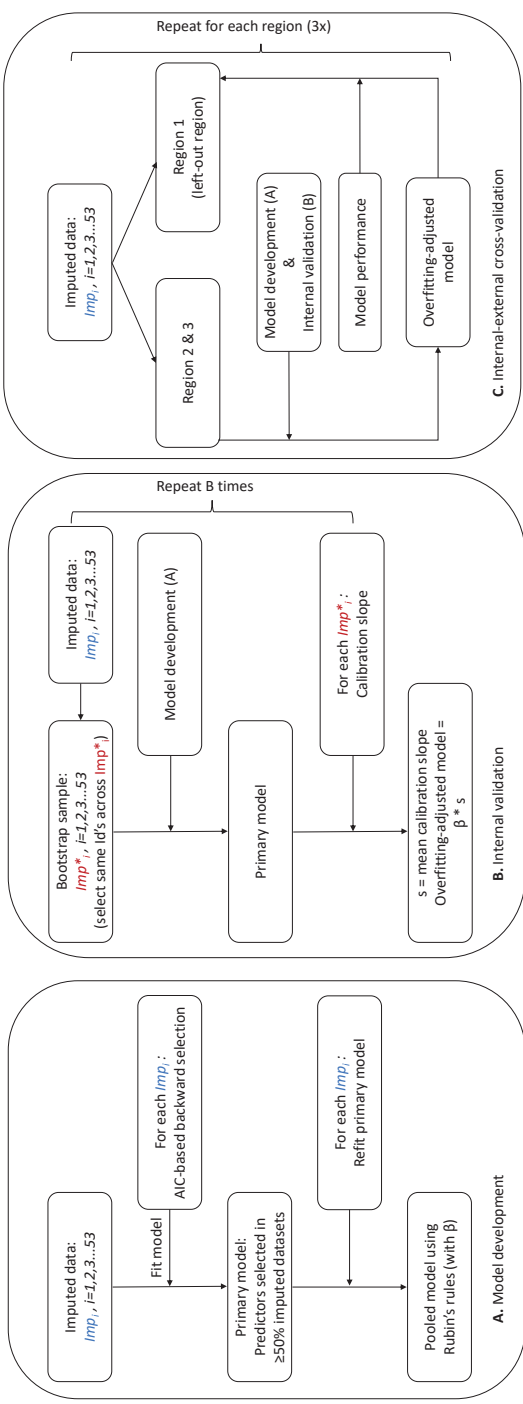
The imputation model contained all above selected candidate predictor variables including rcs transformations to accommodate congeniality, and included sidedness of the primary tumour, age, T-status, N-status, *RAS/BRAF*<sup>V600E</sup> mutational status, number and size of liver metastases, serum CEA, DFI, systemic perioperative treatment type, R-status, tumour burden score (TBS)<sup>12</sup>, passively imputed in the model as auxiliary variables, as well as the primary outcome (EHRFS within 12 months using a Nelson-Aalen estimator and event indicator). We generated 53 imputed datasets, based on the percentage of patients with at least one missing variable in the candidate predictor variables set. To enable internal-external cross-validation without information leakage, we imputed the data separately for each geographical region.

### ***Developing, validating and assessing performance of clinical risk score***

Following multiple imputation, a prediction model for EHRFS within 12 months after local treatment of CRLM (EHRFS model) was created using Cox regression, ignoring follow-up information beyond 12 months. This 12-month time-horizon was chosen to allow a sufficient number of events for robust model-development based on the criteria by Riley *et al.*<sup>3</sup>. The primary evaluation of resulting model's performance was early EHR ( $\leq 6$  months). Because of the limited evaluated follow-up period we did not evaluate deviations from the proportionality assumption.

The primary prediction model was developed in the whole cohort, using a Cox proportional hazards model with Akaike Information Criterion (AIC)-based backward selection in each imputed dataset leading to a primary model only including predictors selected in  $\geq 50\%$  of imputed datasets, which was then refitted in each imputed dataset to obtain a pooled selection model using Rubin's rules (Supplementary Methods Figure 1). The only variable that was not subjected to the above was adjuvant systemic therapy, which was included in all models using an offset for expected therapeutic efficacy (i.e. this effect was not estimated from the data but was imposed on the model by the offset). For the expected adjuvant systemic treatment effect we used the pooled random effects hazard ratio (HR) from known randomized controlled trials in the literature<sup>13,14</sup>, resulting in a HR of 0.73.

Supplementary Methods Figure 1. Model development, internal validation and internal-external cross-validation for development of prediction model



Suppl. Methods Fig. 1. A flowchart illustrating the steps in model development (A), internal validation (B) and internal-external cross-validation (C), adapted from <sup>15,16</sup>. The EHRFS model was developed by performing AIC-based backward selection of all candidate predictors in each imputed dataset. The selection model included variables which were selected in  $\geq 50\%$  of the imputed datasets. The EHRFS model regression coefficients were pooled using Rubin's rules. Internal validation was performed using 500 bootstrap samples to determine the shrinkage factor and overfitting-adjusted model. The model development and internal validation was repeated for internal-external cross-validation, using geographical regions in the development and validation cohorts. The cross-validation was repeated for all regions, resulting in three performance measures.

Abbreviations: AIC (Akaike Information Criteria), B (number of bootstrap samples),  $\beta$  (regression coefficient),  $Imp^*$  (bootstrap object of an imputed dataset),  $Imp$  (imputed dataset), s (shrinkage factor)

Model performance was assessed using calibration plots for 6 and 12 month EHR risk, discrimination (Harrell's C-index, Uno's C-index through 6 and 12 months), time-dependent receiver operator characteristic (ROC) curve, Nagelkerke's  $R^2$  and decision-curve analysis. Each measure was determined for each imputed dataset separately and pooled using Rubin's rules (Supplementary Methods Figure 1). Model-predicted and Kaplan-Meier observed survival estimates (and 95% CI boundaries) were pooled after complementary log-log transformation, and Nagelkerke's  $R^2$  was pooled after Fisher z-transformation. Decision curve analysis was used to assess the net benefit associated with CRLM treatment decisions based on a given threshold value for 6-month or 12-month EHRFS probability<sup>17</sup>. To visualize the potential relevance of the developed model we used Kaplan-Meier curves for EHRFS, RFS and OS, categorizing patients based on quartiles of (across-imputation dataset pooled) predicted EHR risk.

To quantify the overoptimism of the model regarding predicted risks and discriminative ability, we used internal validation by 500-fold bootstrap resampling, repeating all model-development steps in each bootstrap sample and testing the performance of the resulting models from each bootstrap in the original data. We derived a uniform shrinkage factor from internal validation that we applied to the apparent pooled regression coefficients of the primary model as fitted in the original data to create an overoptimism-corrected model (the offset for adjuvant systemic therapy was not shrunk). This overoptimism-corrected model yields predicted EHR probabilities that will agree more with actual risk in new patients. We similarly obtained overoptimism-corrected C-indexes that likely better reflect the actual discrimination of our model in new patients.

We used internal-external cross-validation to evaluate the generalizability of the model (Supplementary Methods Figure 1). The data were split in three geographic regions and all above described modeling steps including internal validation were repeated in two of three regions, after which the performance of the overfitting-adjusted model was evaluated in the left-out geographical region (C-index, calibration slope and intercept). Each geographic region was left-out of model development once, resulting in three estimates of external validation.

As an exploratory additional analysis, we tested whether the prognostic value of *RAS* mutation for EHRFS depended on the administration of preoperative systemic treatment which was reported<sup>18,19</sup>, by using a D1 test between a multivariable model without and with a *RAS*\*preoperative systemic treatment interaction term. To avoid bias, two separate multiple imputation models were created for patients based on preoperative systemic treatment status solely for the exploratory analysis.

## REFERENCES

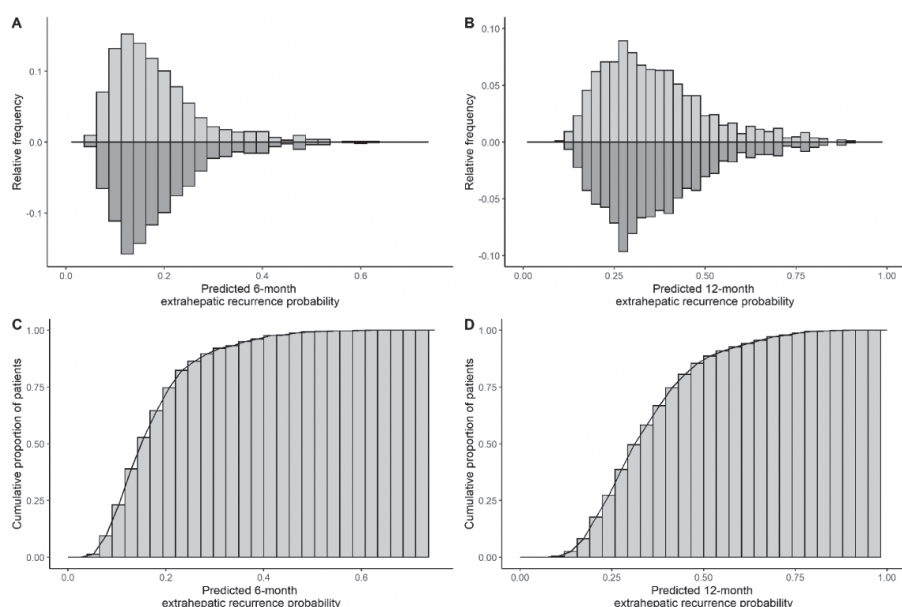
1. Malik H.Z., Gomez D., Wong V., Al-Mukthar A., Toogood G.J., Lodge J.P.A., et al. Predictors of early disease recurrence following hepatic resection for colorectal cancer metastasis. *Eur J Surg Oncol.* (2007);**33**(8):1003–9.
2. Vigano L., Capussotti L., Lapointe R., Barroso E., Hubert C., Giuliani F., et al. Early recurrence after liver resection for colorectal metastases: risk factors, prognosis, and treatment. A LiverMetSurvey-based study of 6,025 patients. *Ann Surg Oncol.* (2014);**21**(4):1276–86.
3. Riley R.D., Snell K.I.E., Ensor J., Burke D.L., Harrell F.E., Moons K.G.M., et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med.* (2019);**38**(7):1276–96.
4. Chen Y., Chang W., Ren L., Chen J., Tang W., Liu T., et al. Comprehensive Evaluation of Relapse Risk (CERR) Score for Colorectal Liver Metastases: Development and Validation. *Oncologist.* (2020);**25**(7):e1031–41.
5. Fong Y., Fortner J., Sun R.L., Brennan M.F., Blumgart L.H. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. *Ann Surg.* (1999);**230**(3):309–21.
6. Margonis G.A., Sasaki K., Gholami S., Kim Y., Andreatos N., Rezaee N., et al. Genetic And Morphological Evaluation (GAME) score for patients with colorectal liver metastases. *Br J Surg.* (2018);**105**(9):1210–20.
7. Paredes A.Z., Hyer J.M., Tsilimigras D.I., Moro A., Bagante F., Guglielmi A., et al. A Novel Machine-Learning Approach to Predict Recurrence After Resection of Colorectal Liver Metastases. *Ann Surg Oncol.* (2020);**27**(13):5139–47.
8. Brudvik K.W., Jones R.P., Giuliani F., Shindoh J., Passot G., Chung M.H., et al. RAS Mutation Clinical Risk Score to Predict Survival After Resection of Colorectal Liver Metastases. *Ann Surg.* (2019);**269**(1):120–6.
9. Buisman F.E., Galjart B., Buettner S., Groot Koerkamp B., Grunhagen D.J., Verhoef C. Primary tumor location and the prognosis of patients after local treatment of colorectal liver metastases: a systematic review and meta-analysis. *HPB.* (2020);**22**(3):351–7.
10. Yamashita S., Chun Y.S., Kopetz S.E., Vauthey J.N. Biomarkers in colorectal liver metastases. *Br J Surg.* (2018);**105**(6):618–27.
11. van Buuren S., Groothuis-Oudshoorn K., Vink G., Doove L., Jolani S., Schouten R., et al. mice: Multivariate Imputation by Chained Equations in R. R package. *J Stat Softw.* (2011);**45**(3):1–67.
12. Sasaki K., Morioka D., Conci S., Margonis G.A., Sawada Y., Ruzzenente A., et al. The Tumor Burden Score: A New “Metro-ticket” Prognostic Tool For Colorectal Liver Metastases Based on Tumor Size and Number of Tumors. *Ann Surg.* (2018);**267**(1):132–41.
13. Nordlinger B., Sorbye H., Glimelius B., Poston G.J., Schlag P.M., Rougier P., et al. Perioperative FOLFOX4 chemotherapy and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC 40983): long-term results of a randomised, controlled, phase 3 trial. *Lancet Oncol.* (2013);**14**(12):1208–15.

14. Kanemitsu Y., Shimizu Y., Mizusawa J., Inaba Y., Hamaguchi T., Shida D., et al. A randomized phase II/III trial comparing hepatectomy followed by mFOLFOX6 with hepatectomy alone for liver metastasis from colorectal cancer: JCOG0603 study. *J Clin Oncol.* (2020);**38**(15\_suppl):4005.
15. Steyerberg E.W., Harrell F.E. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol.* (2016);**69**:245–7.
16. Musoro J.Z., Zwinderman A.H., Puhan M.A., Ter Riet G., Geskus R.B. Validation of prediction models based on lasso regression with multiply imputed data. *BMC Med Res Methodol.* (2014);**14**(1):1–13.
17. Rousson V., Zumbo T. Decision curve analysis revisited: Overall net benefit, relationships to ROC curve analysis, and application to case-control studies. *BMC Med Inform Decis Mak.* (2011);**11**(45):1–9.
18. Takeda Y., Mise Y., Takahashi Y., Ito H., Inoue Y., Yoshioka R., et al. Limited Prognostic Value of KRAS in Patients Undergoing Hepatectomy for Colorectal Liver Metastases. *Ann Surg Oncol.* (2022);**29**(4):2383–91.
19. Kobayashi S., Takahashi S., Takahashi N., Masuishi T., Shoji H., Shinozaki E., et al. Survival Outcomes of Resected BRAF V600E Mutant Colorectal Liver Metastases: A Multicenter Retrospective Cohort Study in Japan. *Ann Surg Oncol.* (2020);**27**(9):3307–15.

## Appendix 1. Formulas to predict extrahepatic recurrence risk at 6 and 12 months following local treatment of CRLM

This appendix is added to the manuscript in line with the TRIPOD recommendations.

The distribution of patients with 6-month (A) and 12-month (B) predicted extrahepatic recurrence risk according to the apparent model is illustrated below. The top light grey values include an offset for patients who received adjuvant systemic treatment, whereas the lower dark grey values indicate the predicted probabilities if patients had not received adjuvant systemic treatment.



Adjuvant systemic therapy was included in the model using an offset for expected therapeutic efficacy (i.e. this effect was not estimated from the data but was imposed on the model by the offset) based on the published hazard ratio, resulting in a HR of 0.73.

### A. As observed in the analyzed dataset.

Baseline cumulative hazard at 6-months:

$$H_0(t_{6\text{ months}}) = 0.0416$$

Baseline cumulative hazard at 12-months:

$$H_0(t_{12\text{ months}}) = 0.0947$$

Prognostic index (PI; linear predictor):

$$\begin{aligned} \text{PI} = & -0.061 * X_{\text{Left-sided primary tumour location}} + 0.212 * X_{\text{Rectum primary tumour location}} \\ & + 0.198 * X_{\text{T3 upon diagnosis}} + 0.543 * X_{\text{T4}} \\ & + 0.200 * X_{\text{N1 upon diagnosis}} + 0.507 * X_{\text{N2}} \\ & + 0.758 * X_{\text{BRAF-mutant}} + 0.510 * X_{\text{RAS-mutant}} \\ & + f(\text{Number of liver metastases}) \\ & + f(\text{Size of largest liver metastasis}) \\ & - 0.314 * X_{\text{Adjuvant systemic treatment}} \end{aligned}$$

Where number of liver metastases and size of largest liver metastasis are described with a restricted cubic spline function:

$$\begin{aligned} f(\text{Number of liver metastases}) = & + 0.101 * \text{Number of liver metastases} \\ & + 0.00194 * \max(\text{Number of liver metastases} - 1, 0)^3 \\ & - 0.0023 * \max(\text{Number of liver metastases} - 2, 0)^3 \\ & + 0.000389 * \max(\text{Number of liver metastases} - 7, 0)^3 \end{aligned}$$

$$\begin{aligned} f(\text{Size of largest liver metastasis}) = & + 0.017 * \text{Size of largest liver metastasis} \\ & - 3.332 * 10^{-6} * \max(\text{Size of largest liver metastasis} - 11, 0)^3 \\ & + 4.535 * 10^{-6} * \max(\text{Size of largest liver metastasis} - 24, 0)^3 \\ & - 1.203 * 10^{-6} * \max(\text{Size of largest liver metastasis} - 60, 0)^3 \end{aligned}$$

The absolute predicted extrahepatic recurrence risk at time  $t$ :

$$\text{Risk} = 1 - \exp(-(\exp(\text{PI}) * H_0(t)))$$

Example:

The 6-month predicted extrahepatic recurrence risk for a patient with a right-sided primary tumour location, T3 tumour stage and N1 nodal stage upon diagnosis, a RAS-mutation and with 1 liver metastasis upon diagnosis of CRLM (size 23 mm), who received adjuvant systemic treatment:

$$\begin{aligned}
H_0(t=6) &= 0.0416 \\
\text{PI} &= -0.061*0 + 0.212*0 + 0.198*1 + 0.543*0 + 0.200*1 + 0.507*0 + \\
&\quad 0.758*0 + 0.510*1 \\
&\quad + 0.101*1 + 0.00194*\max(1-1,0)^3 - 0.0023*\max(1-2,0)^3 + \\
&\quad 0.000389*\max(1-7,0)^3 \\
&\quad + 0.017*23 - 3.332*10^{-6}*\max(23-11,0)^3 + 4.535*10^{-6}*\max(23-24,0)^3 \\
&\quad - 1.203*10^{-6}*\max(23-60,0)^3 - 0.314*1 \\
&= 1.08 \\
\text{Risk} &= 1 - \exp(-(\exp(1.08)*0.0416)) \\
&= 0.115 = 11.5\%
\end{aligned}$$

## B. Following correction for overoptimism.

Correction for overfitting was by 500-fold bootstrap resampling as internal validation. Since adjuvant systemic treatment was modelled using an offset term, its coefficient does not undergo shrinkage.

Baseline cumulative hazard at 6-months:

$$H_0(t_{6 \text{ months}}) = 0.1882$$

Baseline cumulative hazard at 12-months:

$$H_0(t_{12 \text{ months}}) = 0.4249$$

Prognostic index (PI; linear predictor):

$$\begin{aligned}
\text{PI} &= -0.053*X_{\text{Left-sided primary tumour location}} + 0.183*X_{\text{Rectum primary tumour location}} \\
&\quad + 0.171*X_{\text{T3 upon diagnosis}} + 0.469*X_{\text{T4}} \\
&\quad + 0.172*X_{\text{N1 upon diagnosis}} + 0.438*X_{\text{N2}} \\
&\quad + 0.654*X_{\text{BRAF-mutant}} + 0.440*X_{\text{RAS-mutant}} \\
&\quad + f(\text{Number of liver metastases}) \\
&\quad + f(\text{Size of largest liver metastasis}) \\
&\quad - 0.314*X_{\text{Adjuvant systemic treatment}}
\end{aligned}$$

Where number of liver metastases and size of largest liver metastasis are described with a restricted cubic spline function:

$$\begin{aligned}
f(\text{Number of liver metastases}) &= +0.087*\text{Number of liver metastases} \\
&\quad + 0.00167*\max(\text{Number of liver metastases}-1,0)^3 \\
&\quad - 0.0020*\max(\text{Number of liver metastases}-2,0)^3 \\
&\quad + 0.000333*\max(\text{Number of liver metastases}-7,0)^3
\end{aligned}$$

$$\begin{aligned}
 f(\text{Size of largest liver metastasis}) &= + 0.015 * \text{Size of largest liver metastasis} \\
 &\quad - 2.915 * 10^{-6} * \max(\text{Size of largest liver metastasis} - 11, 0)^3 \\
 &\quad + 3.968 * 10^{-6} * \max(\text{Size of largest liver metastasis} - 24, 0)^3 \\
 &\quad - 1.053 * 10^{-6} * \max(\text{Size of largest liver metastasis} - 60, 0)^3
 \end{aligned}$$

The absolute predicted extrahepatic recurrence risk at time  $t$ :

$$\text{Risk} = 1 - \exp(-(\exp(\text{PI}) * H_0(t)))$$





**EXTERNAL VALIDATION OF  
THE COLON LIFE NOMOGRAM  
FOR PREDICTING 12-WEEK  
MORTALITY IN DUTCH METASTATIC  
COLORECTAL CANCER PATIENTS  
TREATED WITH TRIFLURIDINE-  
TIPIRACIL IN DAILY PRACTICE**

Submitted to *Cancers*.

Emerens Wensink<sup>†</sup>  
Patricia Hamers<sup>†</sup>  
Maarten van Smeden  
Geraldine Vink  
Lidwien Smabers  
Renee Lunenberg  
Miangela Laclé  
Miriam Koopman  
Anne May  
Jeanine Roodhart

<sup>†</sup> Both authors share equal contribution.

## ABSTRACT

Predicting prognosis in refractory metastatic colorectal cancer (mCRC) patients is needed to guide decision-making. The Colon Life nomogram was developed to predict 12-week mortality in refractory mCRC patients. The aim of this study is to validate the Colon Life nomogram in last line/refractory mCRC patients receiving trifluridine/tipiracil (FTD/TPI) in daily practice.

The validation cohort consists of 150 QUALITAS study patients, an observational substudy of the Prospective Dutch CRC cohort, who were treated with FTD/TPI between 2016-2019. Model performance was assessed on discrimination, calibration, and clinical usefulness. The additional prognostic value of baseline quality of life (QoL) and thymidine kinase (TK1) expression in tissue was explored.

Of the 150 patients, 25 (16.7%) died within 12 weeks of starting FTD/TPI treatment. The C-statistic was 0.63 (95% C.I. 0.56-0.70). The observed/expected ratio was 0.52 (0.37-0.73). The calibration intercept and slope were -1.06 (-1.53 to -0.58) and 0.41 (0.01-0.81), respectively, which indicated overestimation of 12-week mortality by the nomogram. Decision curve analysis showed the nomogram did not yield a positive net benefit at clinically meaningful thresholds for predicted 12-week mortality. Addition of QoL to the nomogram improved the C-statistic to 0.85 (0.81-0.89). TK1 expression was associated with progression-free survival but not with overall survival.

We demonstrated evident miscalibration of the Colon Life nomogram upon external validation, which hampers its use in clinical practice. We recommend conducting a study with a sufficiently large sample size to update the Colon Life nomogram or to develop a new model including QoL.

## INTRODUCTION

Trifluridine/tipiracil (FTD/TPI) is a treatment available for refractory metastatic colorectal cancer (mCRC) patients<sup>1</sup>. The phase III RECURSE trial demonstrated a statistically significant but modest survival benefit for FTD/TPI over placebo in pretreated mCRC patients with a median overall survival (OS) of 7.1 months versus 5.3 months, respectively<sup>2</sup>. The median progression-free survival (PFS) after FTD/TPI initiation was 2-3 months in clinical trial and real-world study patients<sup>2-4</sup>. The limited median PFS and low response rate (44%<sup>2</sup>) imply that many patients do not benefit from FTD/TPI treatment, resulting in an urgent need for adequate patient selection to avoid unnecessary treatment-related toxicity and to increase cost-effectiveness.

The Colon Life nomogram aims to identify patients eligible for later treatment lines or inclusion in clinical trials by estimating life expectancy<sup>5</sup>. The suggested approach is that in patients with a predicted limited life expectancy, oncologists should consider withholding further treatment<sup>5</sup>. Pietrantonio *et al.*<sup>5</sup> developed the nomogram to predict 12-week mortality from diagnosis of refractory mCRC in 411 Italian mCRC patients who received FTD/TPI (n=27, 6.6%), regorafenib (n=113, 27.5%), or other treatments (n=271, 65.9%) between 2006-2015. The nomogram includes primary tumour resection, ECOG performance status (PS), lactate dehydrogenase (LDH) value, and peritoneal involvement as predictor variables, with the latter three variables determined upon diagnosis of refractory disease. The authors who developed the nomogram also performed three external validations: in a cohort similar to the development cohort<sup>5</sup>, in patients treated with FTD/TPI in the Italian compassionate use program<sup>6</sup>, and in RECURSE trial participants<sup>7</sup>. External validation of the Colon Life nomogram in patients receiving FTD/TPI in daily practice has not yet been described. Based on the nomogram's discriminatory ability in the external validation studies, the Colon Life nomogram was found to be "an accurate tool for estimating life expectancy and the risk of death" in patients with refractory mCRC<sup>5-7</sup>. The calibration of a prediction model, which measures the agreement between predicted and observed outcome risks, is a second important aspect of model performance. Like discrimination, it also influences the clinical usefulness of a prediction model<sup>8,9</sup>.

Two parameters not considered in the development of the Colon Life nomogram which have since received attention as possible prognostic markers in mCRC patients receiving FTD/TPI treatment are QoL<sup>3</sup> and TK1 expression<sup>10,11</sup>. In a recent study, a low QoL at start of treatment was independently associated with shorter OS and PFS<sup>3</sup>. In addition to QoL, TK1 may also be of prognostic value as it has a role in the metabolism of trifluridine.

Trifluridine, the active component of FTD/TPI, is incorporated into tumour DNA, causing DNA damage after phosphorylation by the enzyme TK1<sup>12</sup>. In patients who were treated with FTD/TPI, high levels of TK1 expression (>10%) was associated with longer OS and PFS<sup>10</sup>. Similarly, patients with high TK1 expression (>15%) demonstrated improved OS when treated with FTD/TPI versus placebo<sup>11</sup>. Therefore, we analyzed whether adding QoL and/or TK1 to the Colon Life nomogram increased its prognostic value.

In this study, we aimed to externally validate the Colon Life nomogram in a Dutch real-world mCRC patient population receiving FTD/TPI. Moreover, we explored the additional prognostic value of QoL and TK1 expression to the nomogram.

## METHODS

### External validation cohort

We externally validated the Colon Life nomogram in 150 patients who participated in the QUALITAS study, a substudy of the Prospective Dutch CRC cohort (PLCRC)<sup>3,13</sup>. The QUALITAS study assessed QoL and survival of mCRC patients who were treated with FTD/TPI in daily practice in the Netherlands between 2016-2019. Details on study design, inclusion criteria, data collection, baseline characteristics, assumptions, and definitions were published previously<sup>3</sup>. For the current study, the following additional variables, determined after stopping the previous treatment line and prior to starting FTD/TPI treatment, were collected: LDH value, CEA level, leucocyte count, and ECOG PS (Eastern Cooperative Oncology Group Performance Status).

### External validation of the Colon Life nomogram

#### *Sample size calculations*

We applied the recommendations published by Riley et al.<sup>14</sup> to determine the minimum sample size recommended for external validation for the Colon Life nomogram. For this, we used the reported C-statistic for the Colon Life nomogram in published external validation studies as the anticipated minimum C-statistic for the Colon Life nomogram and an observed/expected (O/E) ratio of 1.0 to determine the expected confidence interval width within our cohort. We calculated that in our cohort of 150 patients and 25 events, we would be able to detect an O/E ratio of 1.0 with a 95% confidence interval (C.I.) of 0.64-1.36 and a C-statistic of 0.72 with a 95% C.I. of 0.61-0.83, assuming that the linear predictor distributions are normal with a common variance in each outcome group.

### ***Missing data and imputation***

Certain candidate predictors had missing data. To avoid loss of information and selection bias, we used multiple imputation using Multivariate Imputation by Chained Equations (MICE)<sup>15</sup>, assuming missingness at random. The imputation model contained the Colon Life nomogram predictor variables including restricted cubic spline transformations (to accommodate congeniality), the outcome (12-week mortality), and joint model-derived baseline EORTC QLQ-C30 Summary Score (QoL-SS) for QoL, time from diagnosis of mCRC to starting date of FTD/TPI, number of metastatic sites, age, CEA and leukocyte count (with the latter four prior to starting FTD/TPI) as auxiliary variables. We generated 29 imputed datasets, based on the percentage of patients with at least one missing key variable.

### ***Assessing model performance***

The full multivariable binary logistic regression model, including the coefficients, intercept, and splines were provided by Pietrantonio *et al.* who developed the nomogram (Appendix 1).

Following multiple imputation, 12-week mortality probabilities were estimated for all patients using this model. Performance of the Colon Life nomogram in QUALITAS patients was assessed through discrimination, calibration, and clinical usefulness. Each performance measure was determined for each imputed dataset separately and pooled using Rubin's rules. Predicted 12-week mortality rates were pooled after complementary log-log transformation.

First, discrimination was determined by the C-statistic. The C-statistic, equivalent to the area under the Receiver Operating Characteristic (ROC) curve, is a measure of how well the model discriminates between those who develop the event (die) and those who do not (survive). The C-statistic is interpreted as the probability that a randomly selected patient who died will have a higher probability of death than a randomly selected subject who survived. Values close to one indicate perfect discrimination ability, while values close to 0.5 indicate poor discrimination ability. Second, calibration was assessed by multiple methods<sup>16,17</sup>. The ratio of observed versus expected number of events (O/E-ratio) was pooled after logarithmic transformation. This ratio should be close to one if the model calibrates well in the validation dataset. Then, model calibration was assessed visually with a calibration plot (a scatter plot of predicted versus observed outcome probabilities) for which well calibrated predictions will lie along the 45° line<sup>16,17</sup>. Finally, calibration was assessed by fitting a logistic regression model with the observed outcome as the dependent variable, and

the log-odds of transformed model predictions as the independent variable. For a perfectly calibrated model, the model intercept and slope would be 0 and 1, respectively. Note that our sample size was insufficient for updating the model in case of miscalibration. As the third step of model performance evaluation, clinical usefulness was assessed by decision curve analysis to determine whether use of a prediction model to inform decision-making in clinical practice may result in net benefit for patients<sup>18–20</sup>. A model can be recommended for clinical use if it has the highest level of benefit across a range of clinically reasonable thresholds<sup>19</sup>. We examined the relative benefit of treating patients with FTD/TPI who did not die within 12 weeks against the harms of treating patients who died within 12 weeks. Our study is reported following the TRIPOD guidelines for prognostic model validation<sup>8</sup>.

### ***Additional prognostic value of baseline QoL***

In the QUALITAS study population, baseline (i.e. at FTD/TPI initiation) overall QoL was measured with the QoL-SS, which encompasses all functional and symptom scales of the EORTC-QLQ-C30 except for financial difficulties<sup>21</sup>. In an exploratory analysis, we re-estimated the logistic regression coefficients before and after adding QoL-SS to the Colon Life nomogram, in order to assess the additional prognostic value of QoL-SS. The additional prognostic value of QoL-SS to the Colon Life variables was assessed by (1) a likelihood ratio test comparing model fit of the Colon Life nomogram with re-estimated coefficients to the ‘new’ nomogram (including QoL-SS), and by (2) comparing performance in terms of discrimination (C-statistic).

### ***Prognostic value of TK1 expression***

The association between TK1 expression in tumour tissue and survival (PFS and OS) was examined. For QUALITAS patients who consented to using archival tissue for research, formalin-fixed, paraffin-embedded (FFPE) tumour tissue specimens, either biopsies or surgical specimens, were requested via PALGA, the Dutch nationwide pathology archive for histo- and cytopathology<sup>22</sup>. If available, a primary tumour sample and the most recent metastatic sample prior to starting FTD/TPI were requested for each patient. Immunohistochemical staining of TK1 was performed by the University Medical Center Utrecht pathology laboratory using the BenchMark ULTRA fully automated slide staining system (Roch, Rotkreuz, Switzerland). Antigen retrieval was performed with EDTA at pH 9 for 24 minutes at 100°C, prior to incubating the slides with antibody (anti-human TK1 rabbit monoclonal antibody, Abcam, Cambridge, UK, ab239885, 1:200 dilution) for 32 minutes at room temperature. After staining, digital images were captured using the Nano-

zoomer XR scanner (Hamamatsu, Hamamatsu, Japan). TK1 expression was quantified using QuPath v0.2.3 (QuPath, Edinburgh, Scotland)<sup>23</sup>. To automate TK1 expression level analysis, quantification algorithms were generated and repetitively improved until each command was optimized. Selected commands included performing positive cell detection (Supplementary Table 1) and tumour versus non-tumour tissue classification. Intensity of TK1 expression was categorized by degree of staining, as 1+, 2+, 3+ for weak, moderate, and strong, respectively. An object classifier was made to classify tumour and non-tumour cells and optimized by assessing its performance in distinguishing tumour and non-tumour cells. In case of uncertainty between tumour and non-tumour parts, the image was assessed by a pathologist. Using the QuPath-obtained values, TK1 expression levels were quantified using two parameters. First, as the percentage of tumour cells with a moderate or strong degree (+2 or +3) of TK1 staining, consistent with previously published methods<sup>11</sup> and referred to as the 'percentage TK1-expressing tumour cells'. Secondly, as the H-score (multiplying the percentage of cells in each staining intensity category by the number associated with the category (0, 1, 2, or 3), and summing the results). The H-score results range from 0-300, where 0 indicates that 'all cells are negative' and 300 indicates 'all cells are strongly positive'<sup>24</sup>. Differences in survival in patients categorized per TK1 expression level were compared using log-rank tests in groups based on exploratory cut-off values. A range of cut-off values for the percentage of TK1 expressing cells and the H-score were selected to ensure that it contained the median value, conform previous studies<sup>11</sup>, and spanned the interquartile range values.

## Statistical analysis

*P* values <0.05 were considered statistically significant and all tests were two-sided. Analyses were carried out using SPSS version 26.0 and R version 4.0.3. (packages mice, rms, pROC, survival)<sup>25</sup>.

## RESULTS

### External validation cohort

Baseline characteristics of the QUALITAS validation cohort compared with the development cohort and previous external validation cohorts are presented in Table 1. The observed 12-week mortality event rate in the development cohort (30.2%) was considerably higher than our observed 12-week mortality rate (16.7%, 25/150). All patients in the QUALITAS validation cohort and only 6.6% of the patients in the development cohort were treated with FTD/TPI.

**Table 1.** Characteristics of the QUALITAS study patients compared to development and external validation cohort patients

	QUALITAS <sup>3</sup>	Development dataset <sup>5</sup>	External validation dataset <sup>5</sup>	Italian CUP <sup>6</sup>	RECOURSE <sup>7</sup>
<b>Total study population</b>	150 (100%)	411	410	341	161 FTD/TPI
<b>Number of deaths within 12 weeks</b>	25 (17%)	124 (30%)	89 (22%)	60 (18%)	22 (14%)
<b>Inclusion period</b>	2017-2019	2006-2015	2010-2016	2015-2016	2012-2013
<b>Country</b>	the Netherlands	Italy	Italy	Italy	Japan, USA, EU, Australia
<b>Male sex</b>	102 (68%)	242 (59%)	251 (61%)	212 (62%)	109 (68%)
<b>Age</b>					
- Mean ( $\pm$ SD)	65.0 ( $\pm$ 9.1)	-	-	-	-
- Median (IQR)	65 (59-72)	66 (58-72)	65 (55-71)	61 (?)	-
- Range	33-86	-	-	33-81	-
<65	68 (45%)	-	-	-	96 (60%)
65-74	60 (40%)	-	-	-	7 (4%)
$\geq$ 75	22 (15%)	-	-	-	58 (36%)
<b>ECOG Performance Status</b>					
0	34 (23%)	194 (47%)	210 (51%)	200 (59%)	87 (54%)
1	63 (42%)	154 (38%)	200 (49%)	132 (39%)	74 (46%)
2	14 (9%)	63 (15%)	-	9 (2%)	-
3	-	-	-	-	-
Missing	39 (26%)				
<b>Primary tumour site</b>					
Right-sided colon	44 (29%)	128 (31%)	135 (33%)	99 (29%)	-
Left-sided colon	54 (36%)	156 (38%)	179 (44%)	-	-
Rectal	52 (35%)	127 (31%)	96 (23%)	-	68 (42%)
<b>Primary tumour resection</b>	113 (75%)	348 (85%)	358 (87%)	312 (91%)	124 (77%)
Synchronous mCRC	98 (65%)	292 (71%)	272 (66%)	209 (61%)	NR
Metachronous mCRC	52 (35%)	119 (29%)	138 (34%)	132 (39%)	

	QUALITAS <sup>3</sup>	Development dataset <sup>5</sup>	External validation dataset <sup>5</sup>	Italian CUP <sup>6</sup>	RECOURSE <sup>7</sup>
<b>Molecular pathology<sup>a</sup></b>					NR
<i>BRAF</i> mutation	4 (3%)	13 (3%)	17 (4%)	17 (5%)	
<i>BRAF</i> wildtype	108 (72%)	219 (53%)	224 (55%)	-	
<i>BRAF</i> status unavailable	38 (25%)	179 (44%)	169 (41%)	-	
<i>RAS</i> mutation	67 (45%)	167 (41%)	198 (49%)	200 (59%)	
<i>RAS</i> wildtype	49 (33%)	173 (42%)	133 (32%)	-	
<i>RAS</i> status unavailable	34 (23%)	71 (17%)	79 (19%)	-	
MSI	2 (1%)	-	-	6 (2%)	
MSS	87 (58%)	-	-	172 (50%)	
MS status unavailable	61 (41%)	-	-	163 (48%)	
<b>Number of metastatic sites</b>					NR
No distant metastasis	1 (0.7%)	-	-	-	
1 organ	20 (13%)	87 (21%)	81 (20%)	14 (4%)	
2 organs	57 (38%)	172 (42%)	147 (36%)	64 (19%)	
>2 organs	-	-	-	-	
3 organs	48 (32%)	-	-	122 (36%)	
≥3 organs	-	152 (37%)	182 (44%)	-	
≥4 organs	24 (16%)	-	-	141 (41%)	
<b>Localization of metastases</b>					
Liver	115 (77%)	313 (76%)	308 (75%)	267 (78%)	-
Liver-only	10 (7%)	-	-	12 (4%)	-
Lung	102 (68%)	258 (63%)	254 (62%)	255 (75%)	-
Lung-only	6 (4%)	-	-	-	-
Peritoneal	31 (21%)	95 (23%)	102 (25%)	82 (24%)	30 (19%)
Peritoneal-only	4 (3%)	-	-	-	-
Bone	28 (19%)	36 (9%)	29 (7%)	-	-
Brain	3 (2%)	10 (2%)	11 (3%)	10 (3%)	-

	QUALITAS <sup>3</sup>	Development dataset <sup>5</sup>	External validation dataset <sup>5</sup>	Italian CUP <sup>6</sup>	RECOURSE <sup>7</sup>
<b>No. of prior systemic treatment regimens<sup>b</sup></b>					
Median (range)	-	3 (1-7)	3 (1-9)	-	-
0	1 (1%)	-	-	-	-
1	18 (12%)	-	-	21 (6%)	-
0-1	-	-	-	-	-
2	76 (51%)	-	-	93 (27%)	-
>2	-	-	-	-	-
3	41 (27%)	-	-	96 (28%)	-
4	14 (9%)	-	-	78 (23%)	-
5	-	-	-	31 (9%)	-
6	-	-	-	18 (5%)	-
≥7	-	-	-	4 (1%)	-
<b>FTD/TPI-treated patients</b>	150 (100%)	27 (6.6%)	100 (24%)	341 (100%)	161 (100%)
<b>Regorafenib-treated patients</b>	0 (0%)	113 (27%)	91 (22%)	-	-
<b>Exposure to prior systemic anticancer agents<sup>c</sup></b>		NR	NR		
fluoropyrimidine	150 (100%)			337 (99%)	161 (100%)
irinotecan	82 (55%)			334 (98%)	161 (100%)
oxaliplatin	132 (88%)			312 (91%)	161 (100%)
bevacizumab	95 (63%)			294 (86%)	-
afibercept	-			31 (9%)	-
anti-EGFR	47 (31%)			143 (42%)	-
regorafenib	0 (0%)			121 (35%)	-
<b>Exposure to treatments:</b>					
Standard chemotherapy agents <sup>c</sup>	75 (50%)			-	161 (100%)
Standard chemotherapy agents <sup>c</sup> + bevacizumab <sup>d</sup>	55 (37%)			-	-

	QUALITAS <sup>3</sup>	Development dataset <sup>5</sup>	External validation dataset <sup>5</sup>	Italian CUP <sup>6</sup>	RECOURSE <sup>7</sup>
<b>Time from diagnosis mCRC to start FTD/TPI (months)</b>					
Median (IQR)	26.2 (16.8-40.8)	19 (13-29) <sup>c</sup>	26 (17-40) <sup>c</sup>	-	NR
<18 months	43 (29%)	-	-	90 (26%)	
≥18 months	107 (71%)	-	-	248 (73%)	
<b>CEA (ng/mL)</b>					
Median (IQR)	46 (16-259)	42 (7-190)	58 (15-252)	NR	NR
Mean (SD)	427	-	-		
Missing	10 (7%)				
<b>Leucocytes (/uL)</b>					
< 10,000	122 (81%)	345 (84%)	336 (82%)	NR	NR
≥ 10,000	21 (14%)	66 (16%)	74 (18%)		
Missing	7 (5%)				
<b>Lactate dehydrogenase (U/L)</b>					
Median (IQR)	272 (210-391)	271 (191-480)	353 (215-529)	NR	351 (245-561)
Missing	7 (5%)				

The baseline characteristics of the external validation cohort (QUALITAS study patients) is shown alongside the published characteristics of the development and previously published external validation cohorts. Synchronous disease was defined as stage IV CRC at diagnosis. Age, performance status, number of metastatic sites, localization of metastases, number of prior treatment regimens, exposure to prior systemic anticancer agents, CEA, leukocytes and LDH were all determined at refractory disease, prior to starting FTD/TPI treatment.

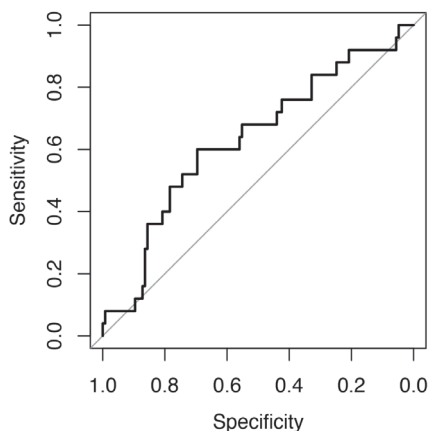
**Abbreviations:** CRC (colorectal cancer), CUP (Compassionate Use Program), FTD/TPI (trifluridine/tipiracil), IQR (interquartile range), mCRC (metastatic colorectal cancer), No. (number), NR (not reported), SD (standard deviation).

- a. We assumed that RAS and BRAF mutations are mutually exclusive.
- b. If a tumour recurred within 6 months after the last administration of adjuvant systemic therapy, this adjuvant systemic therapy was counted as a prior treatment regimen for metastatic disease.
- c. Prior exposure to fluoropyrimidine, oxaliplatin and irinotecan. Adjuvant chemotherapy was counted as prior exposure to fluoropyrimidine/oxaliplatin regardless of the interval between last administration of adjuvant chemotherapy and recurrence.
- d. Most likely in the context of the TASCO1 study.
- e. Time to chemorefractoriness.

## Assessment of Colon Life nomogram performance in the QUALITAS dataset

In our cohort, the C-statistic, which assesses discriminative ability, was 0.63 (95% C.I. 0.56-0.70) (Figure 1). The O/E-ratio was 0.52 (95% C.I. 0.37-0.73). The calibration plot indicated systematic evident overestimation of the observed mortality (Figure 2). For comparison, the calibration plot for our cohort is shown alongside the calibration plots for the published external validation cohorts (Supplementary Figure 1). The calibration intercept was -1.06 (95% C.I. -1.53 to -0.58) and the slope was 0.41 (95% C.I. 0.01 – 0.81). Figure 3 shows the decision curves for informing clinical decision-making regarding FTD/TPI treatment using 12-week mortality risk estimated by the Colon Life nomogram. The net benefit for patients is determined by comparing across 3 scenarios: non-informed decision-making, by (1) treating all patients or (2) treating no patients, and (3) informed decision-making, by treating patients according to the nomogram's predicted 12-week mortality for a given range of thresholds. In the third scenario, patients with a predicted 12-week mortality below the threshold are treated with FTD/TPI, while patients with a predicted 12-week mortality above the threshold are not. The overall benefit is the sum of the net benefit for the treated and untreated patients. Colon Life nomogram guided FTD/TPI treatment, compared to non-informed decision-making by treating all or no patients, resulted in a consistently lower overall net benefit for clinically meaningful thresholds (predicted 12-week mortality risk thresholds above 19%). Using nomogram-informed risk to guide clinical decision-making for patients who did not receive FTD/TPI treatment, resulted in a net benefit <0 (harm) for clinically relevant thresholds.

**Figure 1.** Receiver-operator curve for Colon Life nomogram in QUALITAS patients



**Figure 2.** Calibration plot of the Colon Life nomogram: predicted versus observed 12-week mortality in QUALITAS patients

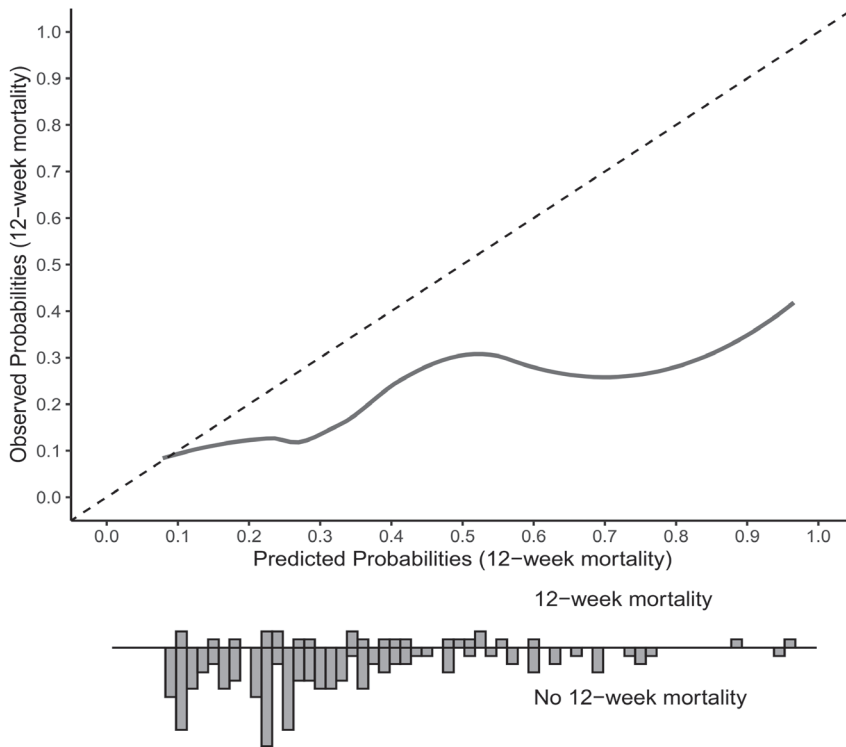


Fig 2. The Colon Life nomogram predicted 12-week mortality versus the observed 12-week mortality is shown, with the diagonal line indicating performance of a well-calibrated model. The histogram displays the predicted probability distribution for patients who died within 12 weeks of initiating FTD/TPI treatment versus who survived.

*Abbreviations:* FTD/TPI (trifluridine/tipiracil).

**Figure 3.** Decision curve analysis indicating the net benefit when using Colon Life nomogram-informed decision-making for FTD/TPI treatment

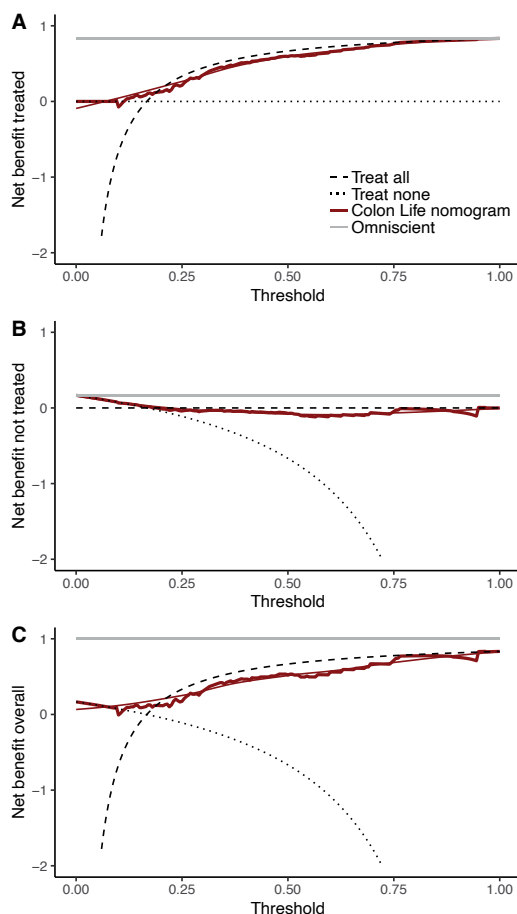


Fig 3. Decision curves are shown for informing clinical decision-making regarding FTD/TPI treatment using predicted 12-week mortality risk estimated by the Colon Life nomogram. The net benefit for patients is compared across 3 scenarios: non-informed decision-making, by treating all patients (“treat all”) or treating no patients (“treat none”), and informed decision-making, by treating patients according to the models’ predicted 12-week mortality risk for a given range of thresholds (“Colon Life nomogram”, thick red line is the Loess smoothed line and the thin red line is the non-smoothed line). The “Omniscient” line indicates the line for a hypothetical all-knowing model. The overall benefit is the sum of the net benefit for the treated and untreated patients. The thresholds shown represent values for 12-week mortality rate. Patients receiving FTD/TPI treatment benefit if the predicted 12-week survival probability is above the threshold value and the patient did not die within 12 weeks (versus harm if they did). Similarly the net benefit for not receiving FTD/TPI is analyzed, where patients benefit if the predicted 12-week survival probability is below the threshold value and the patient did die within 12 weeks (versus harm if they did not).

*Abbreviations:* FTD/TPI (trifluridine/tipiracil).

## Additional prognostic value of baseline QoL

A low QoL-SS at start of treatment (baseline) was previously found to be independently prognostic for shorter survival in QUALITAS patients<sup>3</sup>. We examined the additional prognostic value of adding baseline QoL-SS to the Colon Life nomogram. After adding baseline QoL-SS to the logistic regression model, the C-statistic improved from 0.69 (95% C.I. 0.62-0.75) for a Colon Life nomogram with re-estimated coefficients to 0.85 (95% C.I. 0.81-0.89; Supplementary Figure 2). The likelihood ratio test demonstrated a statistically significant improvement in model fit ( $p < 0.005$ ). The distribution of the predicted 12-week mortality is displayed along with the calibration plot in Supplementary Figure 3.

## Prognostic value of TK1 expression

We examined the prognostic value of TK1 since high TK1 expression was previously found to be associated with improved PFS and OS in FTD/TPI-treated patients<sup>10,11</sup>. TK1 expression was analyzed in archival tissue samples of 110 patients with 146 tissue samples, including 93 primary tumour, 51 metastatic and 2 recurrence tissue samples (Supplementary Figure 4). In 36 patients with paired tissue samples, the TK1 expression levels varied substantially between primary versus metastatic tissue (paired sample t-test,  $p < 0.05$  for H-score and % TK1-expressing cells, Supplementary Figure 5). In all patients, the TK1 expression levels were significantly higher in primary tumour tissue samples compared to recurrent or metastatic tissue samples (Kruskal-Wallis test,  $p < 0.05$ , Supplementary Figure 6).

A significant difference in the median PFS was seen for patients across several TK1 cut-off values (Supplementary Table 2). The median PFS was significantly longer in patients with  $< 2.5\%$  TK1 expressing tumour cells, with a median PFS of 104 days (95% C.I. 86-158,  $n = 51$ ) versus 84 days (95% C.I. 72-98,  $n = 100$ ,  $p < 0.05$ ) in patients with  $\geq 2.5\%$  TK1 expressing tumour cells. Similarly, the median PFS was longer in patients with a H-score  $< 30$  (98 days, 95% C.I. 90-141,  $n = 71$ ) compared to patients with a H-score  $\geq 30$  (77 days, 95% C.I. 67-97,  $n = 80$ ,  $p < 0.05$ ; Supplementary Table 2). The results remained significant when analyzing only primary tissue samples. A similar trend was seen when analyzing only metastatic tissue samples, although the result was significant at a higher TK1 cut-off value (15% for percentage TK1 expressing tumour cells and 45 for H-score). For OS, no prognostic value was seen for TK1 expression level (Supplementary Table 3).

## DISCUSSION

The Colon Life nomogram showed evident miscalibration in our real-world population of mCRC patients who were treated with FTD/TPI. Observed and predicted 12-week mortality did not concur. Using the Colon Life nomogram to aid in deciding whether to start FTD/TPI treatment in daily practice would do more harm than good. Guiding clinical decision-making with the nomogram resulted in a lower net benefit for clinically relevant 12-week mortality thresholds compared to non-informed decision-making through treating all patients. Although the nomogram demonstrated a higher net benefit for 12-week mortality thresholds below 19%, this is not likely to be clinically relevant, since patients will not be denied FTD/TPI treatment based on at most a 19% predicted risk of 12-week mortality. The miscalibration of the Colon Life nomogram in our validation cohort is consistent with previous external validation studies in patients receiving FTD/TPI treatment<sup>6,7</sup>. Yet, the miscalibrated nomogram was deemed a reliable and accurate prognostic model based on its discriminative ability in earlier studies<sup>5-7</sup>. This is likely due to the general focus in prediction model studies on the discriminative ability, ignoring the importance of adequate calibration<sup>9,26,27</sup>. Miscalibration can be the result of methodological issues during model development, such as overfitting, or due to differences between the development and external validation cohorts, including patient characteristics, mortality rate, or differences resulting from advancements in treatment for refractory mCRC over time<sup>27</sup>. Without recalibration and further external validation of the updated model, the Colon Life nomogram cannot be used in daily practice to guide decision-making for patients who are eligible for FTD/TPI.

In an era in which newly developed prognostic models are increasingly being published each year, external validation and clinical impact studies of existing models are needed to ensure that the models can be used in daily practice<sup>9</sup>. Strengths of our study include external validation of the Colon Life nomogram in patients treated in daily practice with FTD/TPI, in a different geographical area than the original model cohort and adherence to the TRIPOD guidelines. Despite a small sample size with low number of events, we were able to demonstrate evident miscalibration of the nomogram in our cohort.

Given the limited efficacy of FTD/TPI on a group level, methods to improve patient selection are needed to avoid unnecessary exposure to treatment-related toxicity and increase cost-effectiveness. An accurate tool to predict short term mortality is useful

to decide when to refrain from starting FTD/TPI since chemotherapy use near death is likely not beneficial<sup>28,29</sup>. Moreover, the tool can confirm that treatment may be beneficial for patients who are at low risk of short term mortality. We therefore recommend conducting a new well-powered study, complying to the TRIPOD guidelines to either update/recalibrate the Colon Life nomogram or develop a new prediction model for refractory mCRC patients. Updating prediction models and developing a new model requires a large sample size<sup>14,30,31</sup>. A sufficient sample size might be realized by using individual patient data (IPD) of patients treated with FTD/TPI. In this new or updated model we recommend adding baseline QoL as a predictor for survival, given its strong and independent association with survival<sup>13,32,33</sup> and the large improvement in discrimination we saw in our exploratory analysis. Baseline QoL-SS likely better reflects patient functioning and has superior prognostic value than physician-reported indicators, such as ECOG PS<sup>32,34</sup>. With increasing use of patient-reported outcome measures (PROMs), a prognostic tool which incorporates baseline QoL-SS can be feasible in daily practice. Ideally, this new model will be able to predict mortality in refractory mCRC patients, regardless of the treatment type. However, due to differences in patient populations and treatment effect, models may differ for standard-of-care treatments versus experimental treatments (e.g. phase I studies).

TK1 expression was identified in previous literature as a prognostic biomarker for survival in FTD/TPI-treated patients, although results were conflicting<sup>10,11</sup>. Kuboki *et al.*<sup>10</sup> reported improved OS and PFS for patients with high TK1 expression. Yoshino *et al.*<sup>11</sup> found that OS was worse in patients with high TK1 expression who received placebo, whereas patients with high TK1 expression showed an improvement in OS when treated with FTD/TPI. These discordant results might be explained by the difference in cut-off values and method of quantification between studies. Moreover, the dual role of TK1 as a predictive and prognostic biomarker complicates the expected effect. TK1 is upregulated during the cell cycle<sup>35</sup>, thus a high TK1 expression level may reflect a poor prognosis in mCRC patients. Yet, the incorporation of trifluridine in the nucleus is dependent on TK1 activity<sup>12</sup>, so a high TK1 expression may also be predictive for a better response to FTD/TPI treatment. Our results show that TK1 expression levels vary per tissue type, with substantially lower TK1 expression level in metastatic tissue samples than primary samples. A high TK1 expression level was associated with poor PFS, with differing cut-off values for primary versus metastatic tissue samples. In univariable analysis, TK1 expression levels were not associated with OS. The feasibility

of TK1 as a biomarker is limited considering the conflicting results and varying cut-off values per tissues sample type, rather than having a consistent accuracy in available archival tissue. Given the differing expression levels, heterogeneous tissue sample types and low number of metastatic tissue samples, we were unable to examine the prognostic value of TK1 in multivariable analysis. Ideally, the prognostic value of TK1 as a biomarker in FTD/TPI-treated patients will be confirmed per tissue sample type.

In conclusion, our external validation study showed evident miscalibration of the Colon Life nomogram, which hampers its use in clinical practice. There is an unmet need for an accurate tool predicting short-term mortality in patients eligible for FTD/TPI treatment, which may be realized by conducting an IPD study with sufficient sample size to enable robust model development to update the Colon Life nomogram or to develop a new model including baseline QoL as a predictor.

## ACKNOWLEDGEMENTS

The authors would like to thank all patients who participated in QUALITAS and the staff of the 26 participating hospitals. We also thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry as well as IKNL staff for scientific advice. We gratefully acknowledge Dr. Pietrantonio and Dr. Miceli for providing the prediction equation of the Colon Life nomogram. We would like to thank the Pathology Laboratory of the University Medical Centre Utrecht (UMCU), including Domenico Castigliero and Erica Siera-de Koning, for immunohistochemistry of tissue samples.

## FUNDING

This study was funded by an unrestricted research grant from Servier.

## REFERENCES

1. Van Cutsem E., Cervantes A., Adam R., Sobrero A., Van Krieken J.H., Aderka D., et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol.* (2016);**27**(8):1386–422.
2. Mayer R.J., Van Cutsem E., Falcone A., Yoshino T., Garcia-Carbonero R., Mizunuma N., et al. Randomized Trial of TAS-102 for Refractory Metastatic Colorectal Cancer. *N Engl J Med.* (2015);**372**(20):1909–19.
3. Hamers P.A.H., Vink G.R., Elferink M.A.G., Stellato R.K., Dijksterhuis W.P.M., Punt C.J.A., et al. Quality of life and survival of metastatic colorectal cancer patients treated with trifluridine-tipiracil (QUALITAS). *Clin Colorectal Cancer.* (2022);**21**(2):154–66.
4. Bachet J.B., Wyrwicz L., Price T., Cremolini C., Phelip J.M., Portales F., et al. Safety, efficacy and patient-reported outcomes with trifluridine/tipiracil in pretreated metastatic colorectal cancer: results of the PRECONNECT study. *ESMO open.* (2020);**5**(3):1–10.
5. Pietrantonio F., Miceli R., Rimassa L., Lonardi S., Aprile G., Mennitto A., et al. Estimating 12-week death probability in patients with refractory metastatic colorectal cancer: The Colon Life nomogram. *Ann Oncol.* (2017);**28**(3):555–61.
6. Cremolini C., Rossini D., Martinelli E., Pietrantonio F., Lonardi S., Noventa S., et al. Trifluridine/Tipiracil (TAS-102) in Refractory Metastatic Colorectal Cancer: A Multicenter Register in the Frame of the Italian Compassionate Use Program. *Oncologist.* (2018);**23**(10):1178–87.
7. Pietrantonio F., Fucà G., Manca P., Pagani F., Raimondi A., Prisciandaro M., et al. Validation of the Colon Life nomogram in patients with refractory metastatic colorectal cancer enrolled in the RECURSE trial. *Tumori.* (2021);**107**(4):353–9.
8. Moons K.G.M., Altman D.G., Reitsma J.B., Ioannidis J.P.A., Macaskill P., Steyerberg E.W., et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med.* (2015);**162**(1):W1–73.
9. Collins G.S., De Groot J.A., Dutton S., Omar O., Shanyinde M., Tajar A., et al. External validation of multivariable prediction models: A systematic review of methodological conduct and reporting. *BMC Med Res Methodol.* (2014);**14**(1):1–11.
10. Kuboki Y., Nishina T., Shinozaki E., Yamazaki K., Shitara K., Okamoto W., et al. TAS-102 plus bevacizumab for patients with metastatic colorectal cancer refractory to standard therapies (C-TASK FORCE): an investigator-initiated, open-label, single-arm, multicentre, phase 1/2 study. *Lancet Oncol.* (2017);**18**(9):1172–81.
11. Yoshino T., Yamazaki K., Shinozaki E., Komatsu Y., Nishina T., Baba H., et al. Relationship Between Thymidine Kinase 1 Expression and Trifluridine/Tipiracil Therapy in Refractory Metastatic Colorectal Cancer: A Pooled Analysis of 2 Randomized Clinical Trials. *Clin Colorectal Cancer.* (2018);**17**(4):e719–32.
12. Sakamoto K., Yokogawa T., Ueno H., Oguchi K., Kazuno H., Ishida K., et al. Crucial roles of thymidine kinase 1 and deoxyUTPase in incorporating the antineoplastic nucleosides trifluridine and 2'-deoxy-5-fluorouridine into DNA. *Int J Oncol.* (2015);**46**(6):2327–34.
13. Burbach J.P.M., Kurk S.A., Coebergh van den Braak R.R.J., Dik V.K., May A.M., Meijer G.A., et al. Prospective Dutch colorectal cancer cohort: an infrastructure for long-term observational, prognostic, predictive and (randomized) intervention research. *Acta Oncol (Madr).* (2016);**55**(11):1273–80.

14. Riley R.D., Debray T.P.A., Collins G.S., Archer L., Ensor J., van Smeden M., et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. *Stat Med.* (2021);**40**(19):4230–51.
15. van Buuren S., Oudshoorn C.G.M. MICE: multivariate imputation by chained equations. *R Packag version.* (2011);**1**(3):2007.
16. Shanahan T.A.G., Fuller G.W., Sheldon T., Turton E., Quilty F.M.A., Marincowitz C. External validation of the Dutch prediction model for prehospital triage of trauma patients in South West region of England, United Kingdom. *Injury.* (2021);**52**(5):1108–16.
17. Grant S.W., Collins G.S., Nashef S.A.M. Statistical primer: Developing and validating a risk prediction model. *Eur J Cardio-thoracic Surg.* (2018);**54**(2):203–8.
18. Vickers A.J., Elkin E.B. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak.* (2006);**26**(6):565–74.
19. Vickers A.J., van Calster B., Steyerberg E.W. A simple, step-by-step guide to interpreting decision curve analysis. *Diagnostic Progn Res.* (2019);**3**(1):1–8.
20. Rousson V., Zumbo T. Decision curve analysis revisited: Overall net benefit, relationships to ROC curve analysis, and application to case-control studies. *BMC Med Inform Decis Mak.* (2011);**11**(45):1–9.
21. Giesinger J.M., Kieffer J.M., Fayers P.M., Groenvold M., Petersen M.A., Scott N.W., et al. Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust. *J Clin Epidemiol.* (2016);**69**:79–88.
22. Casparie M., Tiebosch A.T., Burger G., Blauwgeers H., van de Pol A., van Krieken J.H., et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol.* (2007);**29**(1):19–24.
23. Bankhead P., Loughrey M.B., Fernández J.A., Dombrowski Y., McArt D.G., Dunne P.D., et al. QuPath: Open source software for digital pathology image analysis. *Sci Rep.* (2017);**7**(1):1–7.
24. Fedchenko N., Reifenrath J. Different approaches for interpretation and reporting of immunohistochemistry analysis results in the bone tissue - a review. *Diagn Pathol.* (2014);**9**:221.
25. R: A language and environment for statistical computing. <https://www.r-project.org/>. Vienna, Austria: R Core Team. R Foundation for Statistical Computing;
26. Mahar A.L., Compton C., Halabi S., Hess K.R., Weiser M.R., Groome P.A. Personalizing prognosis in colorectal cancer: A systematic review of the quality and nature of clinical prognostic tools for survival outcomes. *J Surg Oncol.* (2017);**116**(8):969–82.
27. Van Calster B., McLernon D.J., Van Smeden M., Wynants L., Steyerberg E.W., Bossuyt P., et al. Calibration: The Achilles heel of predictive analytics. *BMC Med.* (2019);**17**(1):1–7.
28. Prigerson H.G., Bao Y., Shah M.A., Elizabeth Paulk M., LeBlanc T.W., Schneider B.J., et al. Chemotherapy use, performance status, and quality of life at the end of life. *JAMA Oncol.* (2015);**1**(6):778–84.
29. Akhlaghi E., Lehto R.H., Torabikhah M., Sharif Nia H., Taheri A., Zaboli E., et al. Chemotherapy use and quality of life in cancer patients at the end of life: An integrative review. *Health Qual Life Outcomes.* (2020);**18**(1):1–9.

30. Vergouwe Y., Steyerberg E.W., Eijkemans M.J.C., Habbema J.D.F. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol.* (2005);**58**(5):475–83.
31. Collins G.S., Ogundimu E.O., Altman D.G. Sample size considerations for the external validation of a multivariable prognostic model: A resampling study. *Stat Med.* (2016);**35**(2):214–26.
32. Gotay C.C., Kawamoto C.T., Bottomley A., Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J Clin Oncol.* (2008);**26**(8):1355–63.
33. Bonnetain F., Borg C., Adams R.R., Ajani J.A., Benson A., Bleiberg H., et al. How health-related quality of life assessment should be used in advanced colorectal cancer clinical trials. *Ann Oncol.* (2017);**28**(9):2077–85.
34. Mol L., Ottevanger P.B., Koopman M., Punt C.J.A. The prognostic value of WHO performance status in relation to quality of life in advanced colorectal cancer patients. *Eur J Cancer.* (2016);**66**:138–43.
35. Jagarlamudi K.K., Shaw M. Thymidine kinase 1 as a tumor biomarker: Technical advances offer new potential to an old biomarker. *Biomark Med.* (2018);**12**(9):1035–48.

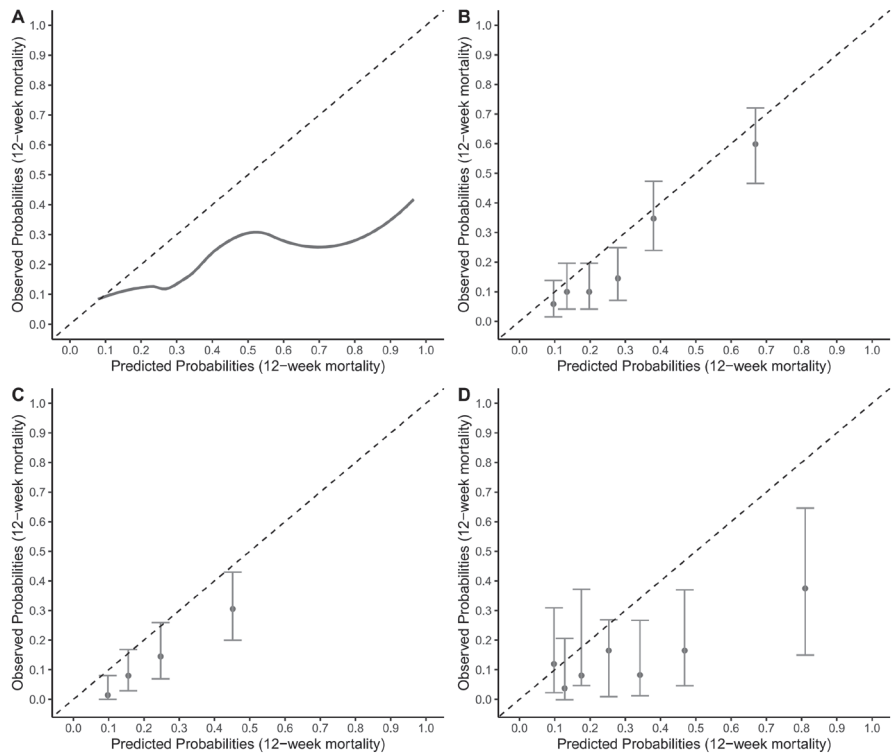
# SUPPLEMENTARY MATERIAL

**Supplementary Table 1.** QuPath specifications for TK1 expression analysis

<b>Positive cell detection</b>	Using the optical density sum and requested pixel size 0.5 $\mu\text{m}$ .
<b>Nucleus parameters</b>	Background radius 8 $\mu\text{m}$ , median filter radius 0 $\mu\text{m}$ . sigma 1.5 $\mu\text{m}$ , minimum area 10 $\mu\text{m}^2$ and maximum area 400 $\mu\text{m}^2$ .
<b>Intensity parameters</b>	Threshold 0.07, maximum background intensity 2, split by shape.
<b>Cell parameters</b>	Cell expansion up to 7.4468 $\mu\text{m}$ and include cell nucleus.
<b>Intensity threshold parameters</b>	Score compartment: cell DAB OD mean. Threshold +1 (0.081), +2 (0.19), +3 (0.5).

The positive cell detection parameters used for the TK1 expression analysis in QuPath are specified in the table.

**Supplementary Figure 1.** Calibration plots of our external validation cohort alongside previously published external validation cohorts

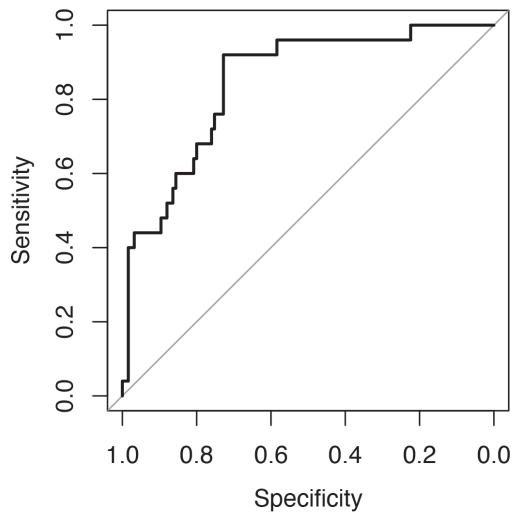


Suppl. Fig. 1. The calibration plots of predicted 12-week mortality probability versus observed 12-week mortality probability are shown for our external validation cohort (A), and as published for the Italian refractory mCRC patients from 12 institutions (external validation cohort)<sup>1</sup> (B), Italian compassionate use program cohort<sup>2</sup> (C), and the RECOURSE trial participants<sup>3</sup> (D). The published calibration plots were digitized using WebPlotDigitizer version 4.4.

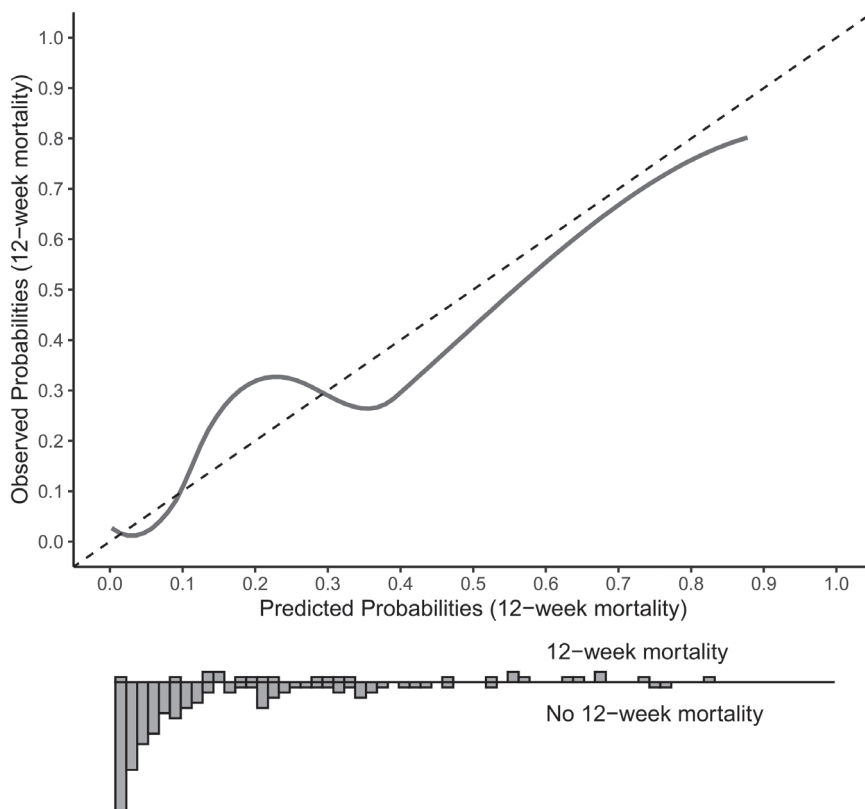
## References:

1. Pietrantonio F, Miceli R, Rimassa L, Lonardi S, Aprile G, Mennitto A., et al. Estimating 12-week death probability in patients with refractory metastatic colorectal cancer: The Colon Life nomogram. *Ann Oncol.* (2017);**28**(3):555–61.
2. Cremolini C., Rossini D., Martinelli E., Pietrantonio F, Lonardi S., Noventa S., et al. Trifluridine/Tipiracil (TAS-102) in Refractory Metastatic Colorectal Cancer: A Multicenter Register in the Frame of the Italian Compassionate Use Program. *Oncologist.* (2018);**23**(10):1178–87.
3. Pietrantonio F, Fucà G., Manca P., Pagani F, Raimondi A., Prisciandaro M., et al. Validation of the Colon Life nomogram in patients with refractory metastatic colorectal cancer enrolled in the RECOURSE trial. *Tumori.* (2021);**107**(4):353–9.

**Supplementary Figure 2.** Receiver-operator curve for the addition of QoL-SS to the Colon Life nomogram in QUALITAS patients



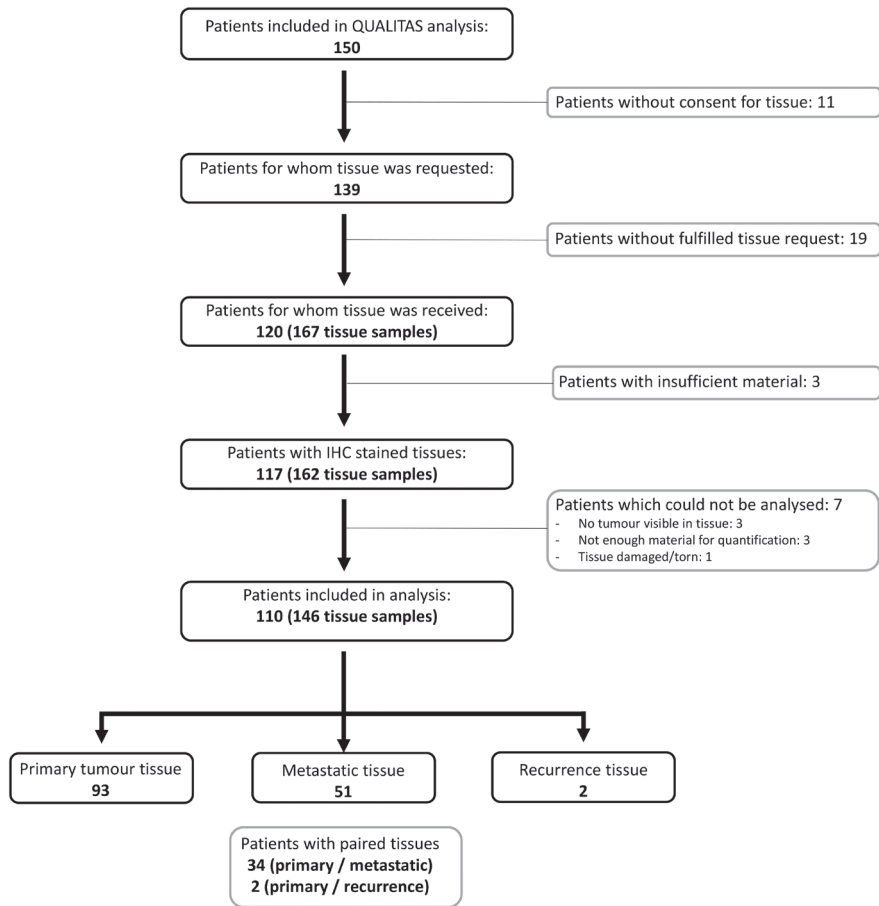
**Supplementary Figure 3.** Calibration plot of the Colon Life nomogram with QoL-SS predicted versus observed 12-week mortality in QUALITAS patients



Suppl. Fig. 3. The Colon Life model with baseline QoL-SS predicted 12-week mortality versus the observed 12-week mortality is shown, with the diagonal line indicating performance of a well-calibrated model. The histogram displays the predicted probability distribution for 12-week mortality risk for patients who died within 12 weeks of initiating FTD/TPI treatment versus who survived.

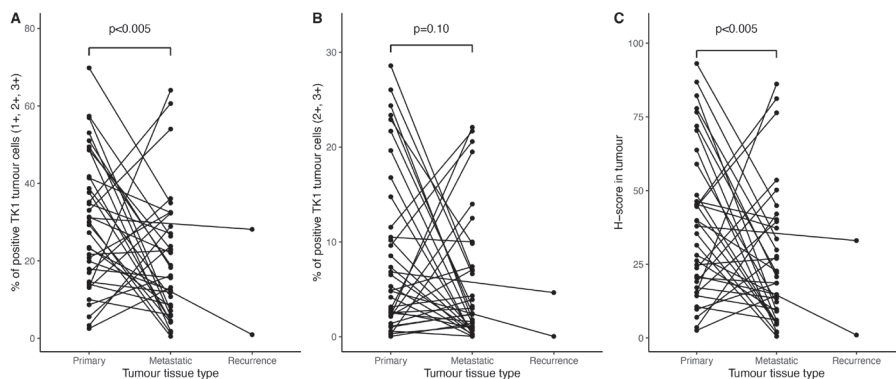
*Abbreviations:* FTD/TPI (trifluridine/tipiracil).

**Supplementary Figure 4.** Flow diagram for QUALITAS patient tissue samples analyzed for TK1 expression



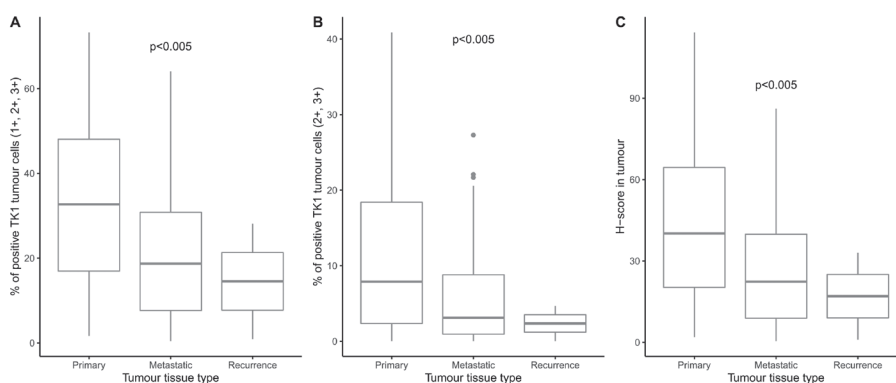
Suppl. Fig. 4. Flow diagram illustrating the QUALITAS patients for which archival tissue could be requested from PALGA, tissue received and tissue which could be analyzed for TK1 expression levels.

**Supplementary Figure 5.** TK1 expression levels in paired primary – metastatic – recurrence tissue samples



Suppl. Fig. 5. Paired scatterplots indicating the TK1 expression level in patients for which paired tissue samples were available (primary tumour, metastatic or recurrence tissue samples). The percentage of TK1 expressing tumour cells (+1,+2, +3), percentage of TK1 expressing tumour cells (+2, +3) and H-score are represented in the y-axis for plots A, B and C, respectively. The paired samples t-test  $p$ -value is indicated in the plot for the primary and metastatic tissue samples.

**Supplementary Figure 6.** TK1 expression levels per tissue type



Suppl. Fig. 6. Boxplots indicating the TK1 expression level for all patients for which tissue samples were available, categorized per type of tissue sample (primary tumour, metastatic or recurrence tissue samples). The  $p$ -values for the Kruskal-Wallis test are shown. The percentage of TK1 expressing tumour cells (+1,+2, +3), percentage of TK1 expressing tumour cells (+2, +3) and H-score are represented in the y-axis for plots A, B and C, respectively.

**Supplementary Table 2.** Progression-free survival in patients with low versus high TK1 expression levels based on different cut-off values

	All tissue samples			Primary tumour			Metastatic tissue		
	<i>n</i>	PFS (median in days; 95% C.I.)	<i>p</i> -value	<i>n</i>	PFS (median in days; 95% C.I.)	<i>p</i> -value	<i>n</i>	PFS (median in days; 95% C.I.)	<i>p</i> -value
% TK1 expressing tumour cells (+2,+3)	51	104 (86-158)	0.032	25	133 (89-238)	0.021	25	97 (67-186)	0.718
≥ 2.5%									
% TK1 expressing tumour cells (+2,+3)	100	84 (72-98)	0.032	72	84 (72-98)	0.021	27	90 (70-180)	0.718
< 5%	69	98 (87-136)	0.071	36	98 (89-192)	0.113	31	104 (79-211)	0.413
% TK1 expressing tumour cells (+2,+3)	82	83 (71-98)	0.071	61	83 (71-108)	0.113	21	84 (65-180)	0.413
≥ 5%	97	97 (86-123)	0.104	55	98 (86-141)	0.300	41	96 (84-141)	0.209
< 10%									
% TK1 expressing tumour cells (+2,+3)	54	78 (68-104)	0.104	42	81 (68-112)	0.300	11	70 (65-NR)	0.209
≥ 10%	113	97 (85-115)	0.245	66	93 (79-123)	0.815	45	97 (84-180)	0.017
< 15%									
% TK1 expressing tumour cells (+2,+3)	38	81 (64-112)	0.245	31	84 (64-163)	0.815	7	65 (51-NR)	0.017
≥ 15%	127	97 (85-112)	0.406	77	97 (79-115)	0.850	48	96.5 (84-141)	0.090
< 20%									
% TK1 expressing tumour cells (+2,+3)	24	81 (64-115)	0.406	20	83.5 (64-180)	0.850	4	61 (51-NR)	0.090
≥ 20%	146	92 (84-108)	0.001	91	92 (84-114)	0.097			
< 30%									
% TK1 expressing tumour cells (+2,+3)	5	55 (50-NR)	0.001	6	56 (50-NR)	0.097			
≥ 30%	37	108 (89-185)	0.222	15	141 (104-337)	0.250	21	85 (56-337)	0.652
H-score in tumour	114	86 (74-98)	0.222	82	84 (72-98)	0.250	31	96 (70-180)	0.652
≥ 15	71	98 (90-141)	0.025	37	104 (91-168)	0.034	33	97 (84-186)	0.530
H-score in tumour	80	77 (67-97)	0.025	60	79 (68-104)	0.034	19	70 (65-185)	0.530
≥ 30	96	98 (90-136)	0.013	52	98 (86-158)	0.174	42	106 (85-185)	0.001
H-score in tumour	55	72 (67-92)	0.013	45	79 (68-108)	0.174	10	65 (51-NR)	0.001
≥ 45	119	97 (85-112)	0.451	70	93 (79-115)	0.952	47	97 (84-146)	0.021
H-score in tumour	32	81 (64-115)	0.451	27	84 (67-163)	0.952	5	52 (51-NR)	0.021
≥ 60									

Log-rank test *p*-values are shown.

Abbreviations: C.I. (confidence interval), *n* (count), PFS (progression-free survival).



**Supplementary Table 3.** Overall survival in patients with low versus high TK1 expression levels based on different cut-off values

		All tissue samples			Primary tumour			Metastatic tissue		
		<i>n</i>	OS (median in days; 95% C.I.)	<i>p</i> -value	<i>n</i>	OS (median in days; 95% C.I.)	<i>p</i> -value	<i>n</i>	OS (median in days; 95% C.I.)	<i>p</i> -value
% TK1 expressing tumour cells (+2,+3)	< 2.5%	51	273 (214-416)	0.385	25	298 (210-503)	0.209	25	242 (166-425)	0.857
	≥ 2.5%	100	234 (176-266)	0.385	72	239 (191-278)	0.209	27	210 (164-424)	0.857
% TK1 expressing tumour cells (+2,+3)	< 5%	69	246 (210-400)	0.679	36	246 (199-425)	0.565	31	273 (210-422)	0.682
	≥ 5%	82	239 (171-278)	0.679	61	244 (191-293)	0.565	21	167 (155-452)	0.682
% TK1 expressing tumour cells (+2,+3)	< 10%	97	239 (199-281)	0.834	55	246 (199-298)	0.709	40	234 (166-416)	0.613
	≥ 10%	54	242 (171-408)	0.834	42	242 (171-408)	0.709	12	219 (164-NR)	0.613
% TK1 expressing tumour cells (+2,+3)	< 15%	113	242 (199-281)	0.672	66	244 (199-298)	0.397	45	239 (167-416)	0.343
	≥ 15%	38	244 (164-438)	0.672	31	249 (167-439)	0.397	7	164 (91-NR)	0.343
% TK1 expressing tumour cells (+2,+3)	< 20%	127	242 (199-281)	0.417	77	246 (199-295)	0.294	48	234 (167-408)	0.559
	≥ 20%	24	244 (167-452)	0.417	20	244 (167-NR)	0.294	4	209 (91-NR)	0.559
% TK1 expressing tumour cells (+2,+3)	< 30%	146	242 (211-281)	0.335	92	249 (211-295)	0.300			
	≥ 30%	5	111 (71-NR)	0.335	5	111 (71-NR)	0.300			
H-score in tumour	< 15	37	281 (210-422)	0.811	15	374 (210-NR)	0.524	21	273 (128-425)	0.965
	≥ 15	114	234 (197-266)	0.811	82	242 (199-281)	0.524	31	220 (164-420)	0.965
H-score in tumour	< 30	71	262 (214-400)	0.561	37	295 (210-425)	0.400	33	242 (167-420)	0.921
	≥ 30	80	234 (171-269)	0.561	60	239 (171-281)	0.400	19	176 (155-452)	0.921
H-score in tumour	< 45	96	242 (210-298)	0.649	52	244 (199-374)	0.893	42	242 (210-420)	0.105
	≥ 45	55	239 (171-293)	0.649	45	249 (191-330)	0.893	10	164 (120-NR)	0.105
H-score in tumour	< 60	119	242 (199-293)	0.553	70	246 (199-295)	0.338	47	239 (167-416)	0.261
	≥ 60	32	241 (167-439)	0.553	27	242 (171-NR)	0.338	5	155 (91-NR)	0.261

Log-rank test *p*-values are shown.Abbreviations: C.I. (confidence interval), *n* (count), OS (overall survival).

## Appendix 1. Formulas to predict 12-week mortality risk following diagnosis of refractory metastatic colorectal cancer (Colon Life nomogram).

The specifications for the Colon Life nomogram were obtained via the authors of the original publication (Pietrantonio *et al.* Estimating 12-Week Death Probability in Patients with Refractory Metastatic Colorectal Cancer: The Colon Life Nomogram. *Ann. Oncol.* **2017**, 28, 555-561.)

Prognostic index (PI):

$$\begin{aligned} \text{PI} &= -2.6025 + 0.6964 * X_{\text{No primary tumour resection}} \\ &+ 1.0033 * X_{\text{ECOG Performance Status} = 1} + 2.0567 * X_{\text{ECOG Performance Status} = 2} \\ &+ f(\text{LDH}) \\ &+ 0.4989 * X_{\text{Peritoneal metastasis}} \end{aligned}$$

Where LDH is described with a restricted cubic spline function:

$$\begin{aligned} f(\text{LDH}) &= 0.0018 * X_{\text{LDH}} - 8.3612 * 10^{-10} * \max(X_{\text{LDH}} - 165)^3 \\ &+ 1.0159 * 10^{-9} * \max(X_{\text{LDH}} - 271)^3 \\ &- 1.7977 * 10^{-10} * \max(X_{\text{LDH}} - 764)^3 \end{aligned}$$

The absolute predicted 12-week mortality risk:

$$\text{Risk} = 1 / (1 + \exp(-\text{PI}))$$

Example:

The 12-week predicted mortality risk for a patient who received a primary tumour resection and who has refractory metastatic colorectal cancer with a performance score of 2, LDH value of 792 and who was diagnosed with peritoneal metastases:

$$\begin{aligned} \text{PI} &= -2.6025 + 2.0567 * X_{\text{ECOG Performance Status} = 2} + 0.0018 * (792) \\ &- 8.3611807 * 10^{-10} * \max(792 - 165)^3 + 1.0158919 * 10^{-9} * \max(792 - 271)^3 \\ &- 1.7977387 * 10^{-10} * \max(792 - 764)^3 + 0.4989 * X_{\text{Peritoneal metastasis}} \\ &= -2.6025 + 2.0567 + 0.0018 * (792) \\ &- 8.3611807 * 10^{-10} * 627^3 + 1.0158919 * 10^{-9} * 521^3 - 1.7977387 * 10^{-10} * 28^3 \\ &+ 0.4989 \\ &= 1.3163 \\ \text{Risk} &= 1 / (1 + \exp(-1.3163)) \\ &= 0.7886 \end{aligned}$$



## SUMMARY (ENGLISH)

## CHAPTER 9

## SUMMARY (ENGLISH)

Recent advancements in the treatment for patients with metastatic colorectal cancer (mCRC) include new types of targeted treatment, immunotherapy and improvement in the local treatment of metastases. Despite these advances, however, patients with mCRC continue to have a poor prognosis. In a study examining the overall survival (OS) in real-life Dutch patients with mCRC, the median OS was 12 months and remained unchanged from 2008 to 2016, with the exception of the best-case patients and patients receiving local treatment. These results differ markedly from the results seen in patients enrolled in clinical trials, re-enforcing the need to ensure that study results can be translated to the general population of mCRC patients. One aspect which may aid in improving the survival of patients with mCRC is enabling personalized treatment, which ensures that patients receive effective treatment without potential toxic side effects from costly ineffective treatment. This thesis focused on improving the treatment of patients with mCRC using personalized treatment.

## PART I – OPTIMIZING TREATMENT FOR PATIENTS WITH DEFICIENT MISMATCH REPAIR COLORECTAL CANCER

### Chapter 1

In 5% of mCRC patients, tumours display a deficient mismatch repair (dMMR) system, resulting in a high mutational load. Patients with a dMMR mCRC tumour have a worse prognosis compared to patients with a proficient mismatch repair (pMMR) tumour. Recently, immunotherapy (immune checkpoint inhibitors; ICI) has been shown to be beneficial in dMMR mCRC patients. However, results were initially based on single-arm trials, lacking data for patients receiving standard systemic treatment. With a low incidence of dMMR tumours among mCRC patients, the survival results for patients while receiving later line chemotherapy and targeted treatment were largely unknown, complicating the interpretation of the ICI trial results. In **Chapter 1**, we examined the OS of patients in a large, comprehensive cohort of 281 dMMR mCRC patients, treated with or without systemic non-immunotherapy. We observed that 62% of dMMR mCRC patients received first line treatment and 26% second line treatment. Prognosis of dMMR mCRC patients was poor, with a median OS from diagnosis of 16.0 months (95% C.I. 13.8 – 19.6) with anti-tumour treatment and 2.5 months (95% C.I. 1.8 – 3.5) without. In patients receiving therapy, the median OS was 12.8 months

(95% C.I. 10.7 – 15.2) from first line treatment and 6.2 months (95% C.I. 5.4 – 8.9) from second line treatment. The resulting 12-month survival rates (17.2%) from second line treatment in our patients were much lower than the published 12-month survival rates for nivolumab (73%) and nivolumab/ipilimumab (85%) beyond first line in the CheckMate 142 trial. In population-based dMMR mCRC patients receiving anti-tumour treatment, the median OS was 7.6 months shorter compared to pMMR mCRC patients receiving anti-tumour treatment. In patients receiving first line systemic treatment, we observed a significant association between metastasectomy with better survival and right-sided primary tumour location at diagnosis ('sidedness') with worse survival. Metastasectomy is a known prognostic factor for OS in dMMR mCRC patients, while having a right-sided tumour was not a known prognostic factor in this subtype of CRC. Our data demonstrated a poor prognostic value for dMMR in mCRC patients. Given the poor outcome in our data set compared to the results of immunotherapy in dMMR mCRC patients, our data strongly supports a survival benefit of immunotherapy in dMMR mCRC patients.

## Chapter 2

Patients with dMMR mCRC have a long-lasting response to immune checkpoint inhibitors (ICI). Chemotherapy and targeted treatment remain a part of treatment, including for patients who did not respond to immunotherapy (20-30%) or in later line treatment. Although biologically distinct, patients with dMMR mCRC tumours are treated with the same chemotherapy and targeted treatment as patients with pMMR mCRC tumours. We do not know if these treatments have the same efficacy in patients with dMMR mCRC, which is relevant considering the worse prognosis and lower treatment response in these patients. Clarity is needed regarding the treatment efficacy of different types of chemotherapy and monoclonal antibodies in dMMR mCRC patients in order to avoid unnecessary exposure to ineffective treatments and toxicity. In **Chapter 2**, an evidence-based review, we summarized clinical trial results reporting treatment efficacy of different types of chemotherapy and monoclonal antibodies in dMMR mCRC patients. We identified 11 trials, however, due to the low amount of evidence, both in number of trials and number of dMMR patients, we could not draw a definitive conclusion regarding the relative efficacy of different types of systemic treatments in dMMR mCRC patients. We identified three trends which need to be confirmed prior to implementation in the clinic. Firstly, patients with dMMR

mCRC who received 5-fluorouracil (5-FU) + irinotecan combination treatment did not have increased benefit over 5-FU monotherapy. Secondly, patients with dMMR mCRC may have benefitted more from oxaliplatin combination treatment compared to irinotecan combination or 5-FU monotherapy, based on results from one study. Thirdly, addition of bevacizumab targeted treatment appeared to be more effective compared to anti-EGFR treatment with cetuximab, while cetuximab compared to no targeted treatment appeared not to offer significant benefit in PFS for *BRAF*-wildtype dMMR mCRC patients. However, there was insufficient data to definitively conclude which systemic agent is preferred. We reviewed biological rationale which may support differential benefit of a specific treatment type. Preclinical studies suggested that dMMR mCRC tumours may be less sensitive to 5-FU, more sensitive to irinotecan, but equally sensitive to oxaliplatin, which is in contrast to the trends observed in clinical studies. The preclinical studies were based on traditional models which may impede the translation of results, compared to more representative patient models (e.g. organoids). To accelerate research regarding the efficacy of chemotherapy and targeted treatment in dMMR mCRC patients, we advise researchers to use representative patient models, to publish patient-level data to enable meta-analyses and to increase testing of MMR status in daily clinical care to increase known dMMR mCRC patients in population registries. Hopefully, new data will help to reveal how best to treat dMMR mCRC patients.

### Chapter 3

In **Chapter 3**, we examined whether the poor response of dMMR mCRC tumours to systemic treatment is due to tumour cell-intrinsic treatment resistance. In a cohort of 12 patient-derived organoids (PDOs) from pMMR and dMMR tumours, we tested the response to different treatment types, including 5-fluorouracil (5-FU), oxaliplatin, SN-38, binimetinib, encorafenib, and cetuximab. We found no significant differences in sensitivity to any of the drugs between the groups of dMMR and pMMR PDOs. As such, the poor prognosis of dMMR mCRC patients did not seem to simply be caused by a tumour-cell intrinsic sensitivity difference to treatment. Potentially, distinct microenvironments in pMMR versus dMMR CRC tumours may influence treatment response. PDO co-culture models would be suited to model the distinct microenvironments. In our cohort we were also able to compare drug responses between subgroups based on *BRAF* and *KRAS* mutational status. We observed that

*BRAF*-mutant and *KRAS*-mutant PDOs were significantly more resistant to 5-FU and oxaliplatin, but not irinotecan (SN-38), compared to *BRAF/KRAS*-wildtype PDOs ( $p=0.017$  and  $p=0.008$ ). As well, *BRAF*-mutant PDOs were more sensitive to encorafenib (versus *KRAS*-mutant and *BRAF/KRAS*-wildtype PDOs), while *BRAF*- and *KRAS*-mutant PDOs were more sensitive to binimetinib (versus wildtype). Thus, mutant *KRAS* and *BRAF* impact the intrinsic sensitivity of tumour cells to chemotherapy and targeted treatment. To tailor treatment strategies for specific CRC subtypes, future drug screens should ideally be performed on rationally chosen cohorts of PDOs including defined by *BRAF/KRAS* mutational and MMR status, or new yet to be identified subtypes.

## PART II – PATIENT-DERIVED ORGANOID AS A PREDICTIVE BIOMARKER TO GUIDE CANCER TREATMENT

### Chapter 4

In personalized treatment, treatment is tailored to the individual characteristics of each patient to optimize a patient's benefit. However, only half of oncology patients are eligible for genetically-matched treatment and for the majority of anticancer agents, no genetic markers are available. As a result, many oncology patients are exposed to ineffective treatment with associated side effects and treatment costs. Patient-derived organoids (PDOs) enable individualized tumour response testing by screening anticancer agents *ex vivo* on PDOs to predict clinical response. In **Chapter 4**, we reviewed 17 publications which report results on using PDOs as a predictive biomarker in the treatment of cancer patients, focusing on evidence for analytical validity, clinical validity and clinical utility. Our review demonstrated that individualized tumour response testing using PDOs has clinically validity as a predictive biomarker for cancer patients. PDO-based screening discriminated clinical response with a pooled sensitivity and specificity of 0.81 (95% C.I. 0.69-0.89) and 0.74 (95% C.I. 0.64-0.82), respectively across different tumour types and treatments. The evidence was strongest for CRC patients, with smaller studies showing promising results for other tumours. Validation studies are needed in similar, larger patient cohorts before PDO-based drug screens can be implemented in the clinic. Future studies should show that using PDOs to guide treatment is feasible, both with a sufficient organoid establishment rate and in adequate time to PDO screen results. Lastly, the clinical utility of using PDOs to guide

treatment needs to be proven by showing that PDO-based individualized tumour response testing is cost effective and offers clinical benefit for patients. Ultimately, if we are able to establish PDOs for the majority of patients within a feasible time frame, PDOs may be a potential predictive biomarker to facilitate personalized treatment for a group of patients for whom there is a great need for valid predictive biomarkers.

## Chapter 5

The inability to predict to which treatment cancer patients will respond results in decreased treatment efficacy, survival and unnecessary toxicity. Drug screening using patient-derived organoids (PDOs) provides a promising predictive biomarker for systemic treatment efficacy. For CRC, it is unclear if PDOs predict for all treatment types, across all treatment lines and for clinically relevant outcomes. In **Chapter 5**, we presented results for our study, in which we examined if PDOs can be used to predict response to multiple standard-of-care (SOC) agents in mCRC patients, including different treatment lines and relevant outcomes.

Drug screens were performed on 19 PDOs derived from CRC patients from the HUB-Cancer Biobank and OPTIC study to measure the growth rate inhibition during exposure to chemotherapies and targeted agents. We quantified organoid sensitivity using the area under the curve of the drug response curve ( $GR_{AUC}$ ), the GR value at the highest concentration ( $GR_{max}$ ) and the concentration of 50% growth rate inhibition ( $GR_{50}$ ). Patient response was evaluated using percentage change in size of the target lesions (decrease or stable/increase in tumour size) on follow-up CT scans. We presented results for 5-fluorouracil (5-FU), 5-FU + irinotecan, irinotecan and panitumumab treatment. Results for oxaliplatin-based treatment will follow. In our study, associations were seen for all treatment types (5-FU, irinotecan-based treatment and panitumumab), with best association for  $GR_{AUC}$  (organoid response) and change in tumour size (patient response). A classifier based on normalized  $GR_{AUC}$  was able to predict patient response (change in tumour size during treatment) for chemotherapy correctly for 13/16 patients, with a sensitivity of 77.8%, specificity of 85.7% and AUC of 0.905. PDOs from non-chemonaïve patients (4/19) showed increased resistance to chemotherapy, but not to panitumumab. Organoids derived from left-sided colon *RAS/BRAF*-wildtype tumours were more sensitive to panitumumab versus *RAS/BRAF*-mutant and *RAS/BRAF*-wildtype right-sided or rectal tumours. Importantly, we observed varying sensitivity to panitumumab within organoids from left-sided

colon *RAS/BRAF*-wildtype tumours. These results suggest that PDOs may further aid in guiding anti-EGFR treatment compared to molecular pathology and sidedness only.

Our results demonstrated that organoid response correlates with patient response to commonly given CRC treatment, suggesting that organoids can serve as a predictive biomarker and guide personalized cancer treatment. Associations were seen for all treatment types (5-FU, irinotecan, 5-FU + irinotecan and panitumumab). Interestingly, we observed varying organoid sensitivities to panitumumab for left-sided colon *RAS/BRAF*-wildtype tumours, suggesting that organoids may further aid in guiding anti-EGFR treatment.

## PART III – PREDICTION MODELS FOR mCRC PATIENTS

### Chapter 6

For patients with colorectal liver metastases (CRLM), local treatment of the metastases through surgical resection or ablation can result in long-term survival, with 5-year OS rates up to 55%. With optimized surgical techniques, more lenient resection criteria, and optimization of induction systemic therapy, more patients with CRLM are eligible for local treatment. However not all patients benefit from this treatment. Relapse after liver resection occurs in up to 75% of patients and a subgroup of patients have no long-term OS benefit. To increase postoperative survival, we need to stratify patients to customize therapy.

Multiple prediction models have been developed to predict prognosis after CRLM resection. Among earlier prediction models for patients with CRLM, the Fong score is still used most frequently to predict prognosis after liver resection. The Fong prediction model incorporated lymph node stage, CEA value, disease-free interval (DFI), and size and number of liver metastases. A more recently developed prediction model for patients with CRLM, the Genetic And Morphological Evaluation (GAME) score, incorporates recalibrated tumour markers such as *KRAS* mutational status, extrahepatic disease presence, and the Tumour Burden Score. The GAME score outperformed the Fong score in two single-institution patient cohorts but lacks external validation in more unselected patient cohorts. To assess if the prediction models can be used to guide treatment, external validation is needed. However, external validation efforts of these prediction models are scarce, especially in populations receiving modern

systemic therapies, improved surgical and ablative treatment options. In **Chapter 6**, we externally validated the Fong and GAME prediction models in a population-based cohort of patients who received local treatment of CRLM. We also externally validated the prediction models in two pre-specified subgroups:  $\leq 70 / > 70$  years and with/without perioperative systemic therapy.

A total of 1105 patients with local treatment of CRLM, diagnosed in 2015/2016, were included from the Netherlands Cancer Registry, a nationwide population-based database. In our cohort, the median OS was 51.3 months and the median disease-free survival (DFS) was 10.1 months. The GAME and Fong score showed discriminative ability in the cohort and in the predefined subgroups. The median OS in the GAME risk groups (high/moderate/low) was 32.4, 46.7, and 68.1 months, respectively ( $p < 0.005$ ). For the Fong model risk groups (high/moderate/low), median OS was 36.1, 48.5 months and not reached within the follow-up period, respectively ( $p < 0.005$ ). For discrimination, Harrell's concordance-index (C-index) was determined. The C-index reflects the ability of the model to differentiate between patients who do and do not experience an event, with 0.5 representing a model without any discriminatory ability beyond chance and 1 perfect discrimination. Overall, the discriminative ability of the GAME and the Fong score, as measured by the Harrell's C-index for OS, was weak, 0.596 (95% C.I. 0.572-0.621) and 0.577 (95% C.I. 0.554-0.601), respectively. The C-index was comparable in the subgroups. The overall predictive performance of the Fong and GAME prediction models was suboptimal, with a high prognostic uncertainty which limits its utility in clinical decision-making. Furthermore, 25% of patients identified as "high-risk" according to the GAME score did achieve long-term survival, which exceeded five years, and this rate was even higher in the Fong high-risk group. In conclusion, both prediction models showed predictive ability in a population-based cohort and in predefined subgroups. However, the limited discriminative ability of these prediction models results in insufficient preoperative risk stratification for clinical decision-making.

## Chapter 7

In **Chapter 7** we developed and validated a model to predict early extrahepatic recurrence (EHR) in patients undergoing local treatment for CRLM. Attempts with prediction models to accurately predict oncological outcomes, such as recurrence or OS, and thus improve the selection of CRLM patients has not yet led to a breakthrough

due to moderate discriminative power and thereby limited clinical utility. We aimed to develop a prediction model which includes novel prognostic factors, such as primary tumour location ('sidedness') and *RAS* and *BRAF*<sup>V600E</sup> mutational status, in addition to classic prognostic factors (e.g. number and size of metastases and serum carcinoembryonic antigen (CEA)). To increase the prognostic strength of our prediction model, we did not simplify the model by categorizing continuous variables and we selected an endpoint which might help guide clinical decision-making. Early recurrences within six months and EHR are independently associated with poor OS. We hypothesized that prediction of early EHR as combined endpoint would give clinically relevant information to patients and clinicians and could help to guide clinical decisions in patients with technically resectable CRLM.

The proposed endpoint of early EHR ( $\leq 6$  months) was analyzed by landmark analysis to avoid immortal time bias. Landmark analysis showed a median OS for patients with EHR, liver-only recurrence and no recurrence within six months of 19.5, 30.7 months and was not reached, respectively. We developed a Cox proportional hazards model for extrahepatic recurrence-free survival in our cohort of 1077 patients with local hepatic treatment of CRLM from the Netherlands Cancer Registry. The final prediction model after backward selection included sidedness, tumour and nodal-stage, *RAS/BRAF*<sup>V600E</sup> mutational status, number and size of liver metastases, excluding pre-operative serum CEA, disease-free interval between primary tumour and CRLM and neoadjuvant systemic therapy. We assessed the performance of the prediction model based on its calibration and discrimination. Harrell's C-index was 0.661 [95% C.I. 0.632; 0.689] and the optimism-adjusted C-index was 0.641 [95% C.I. 0.612; 0.669]. The model performance was confirmed using internal-external cross-validation. Based on quartiles of predicted EHR risk, the very high risk patients had a 6-month EHR rate of 32% versus 20%, 15% and 6% in high, moderate and low risk patients, respectively. Similarly, the recurrence-free survival differed between the patients based on their risk group. The median OS in very high risk patients was 40.4 months versus 46.7, 51.7 and not reached in high/moderate/low risk patients, respectively. Model-guided local treatment of CRLM achieved a net benefit at 6-month EHR thresholds of 0-40%.

Early EHR is a valuable and informative alternative endpoint for prediction models. The EHR prediction model, including *RAS/BRAF*<sup>V600E</sup> mutational status, showed robust performance in discriminating between patients based on EHR probability, reflected in differing EHRFS, RFS and OS. After further external validation, the EHR

prediction model might offer guidance in clinical decision-making in patients with resectable CRLM. Possible applications of the model include identifying low risk patients who might benefit from local treatment of CRLM. The EHR prediction model may also aid clinical decision-making by identifying moderate/high-risk patients for early EHR who may benefit from perioperative systemic treatment. A treatment strategy for these patients may be to initiate long-lasting systemic treatment, and upon sustained response, perform local treatment of CRLM. Once externally validated, the prediction model will lend well for studies examining the optimal treatment by stratifying patients who are at moderate/high risk for early EHR.

## Chapter 8

Trifluridine/tipiracil (FTD/TPI) is a type of chemotherapy which has been approved for patients with refractory mCRC. However, not all patients benefit from FTD/TPI, resulting in unnecessary toxicity and treatment costs. The Colon Life nomogram is a prediction model which has been developed to predict the 12-week death probability in refractory mCRC patients, including patients receiving FTD/TPI. The nomogram includes the following four predictors: performance score, primary tumour resection status, lactate dehydrogenase level and the presence of peritoneal metastases prior to starting FTD/TPI treatment. The Colon Life nomogram may aid in identifying patients with a poor prognosis, and subsequently to avoid exposing patients to FTD/TPI who will die within 12 weeks of starting treatment. The Colon Life nomogram should be externally validated in mCRC patients who receive FTD/TPI in daily practice, the results of which are presented in **Chapter 8**.

The Colon Life nomogram was validated in 150 patients who participated in the QUALITAS observational study, a substudy of the Prospective Dutch CRC cohort (PLCRC). The QUALITAS study examined the quality of life of patients who were treated with FTD/TPI in daily practice between 2016-2019. Model performance was assessed in terms of discrimination, calibration, and clinical usefulness. In addition, the added prognostic value of two known prognostic variables in FTD/TPI patients (quality of life and TK1 expression) was analyzed. Of the 150 patients, 25 (16.7%) died within 12 weeks of receiving FTD/TPI. The Colon Life nomogram was poorly calibrated, with a calibration slope of 0.41 (which ideally should be 1), indicating evident overestimation of the observed mortality. The discrimination was adequate, with a C-statistic of 0.63 (95% C.I. 0.56-0.70). Decision curve analysis showed the

nomogram did not yield a positive net benefit at clinically meaningful thresholds for predicted 12-week mortality. In an exploratory analysis, the addition of quality of life at start of treatment to the Colon Life nomogram improved the discrimination, however, results should be validated in an external population. TK1 tissue expression was prognostic for progression-free survival, but not OS in QUALITAS patients. We demonstrated evident miscalibration of the Colon Life nomogram upon external validation, which hampers its use in clinical practice. We recommend conducting a study with a sufficiently large sample size to update the Colon Life nomogram or to develop a new model including QoL.



# **GENERAL DISCUSSION & FUTURE PERSPECTIVES**

## **CHAPTER 10**

Despite advancements in treatment, the majority of mCRC patients have not benefitted from improved survival<sup>1</sup>. Personalized treatment improves survival by tailoring care to ensure optimal treatment efficacy, reduced side effects and costs. This thesis focusses on three main strategies to enable personalized treatment for patients with metastatic colorectal cancer (mCRC). The first strategy (I) described are patient-derived organoids (PDOs). PDOs can act as a predictive biomarker and preclinical model to identify optimized treatment strategies for CRC subtypes. Second (II), prediction models can be used to predict prognosis for patients who receive treatment and thus identify patients who may benefit more or may not benefit from a treatment. Third (III), population-based patient registries can be used to identify obstacles in treatment strategies for specific types of mCRC patients, assess and improve the performance of existing biomarkers and prediction models.

## **I - USING PDOs TO IMPROVE TREATMENT FOR mCRC PATIENTS**

PDOs provide a promising preclinical models which can be used to improve the treatment for mCRC patients. The various applications of PDOs include exploring disease mechanisms, identifying new treatment opportunities for specific groups of patients and acting as a predictive biomarker. The benefits of using PDOs as a preclinical model have been extensively described and include long-term stability in culture, applicability for high throughput drug screens, heterogeneity and enhanced capacity to capture tumour characteristics<sup>2-5</sup>.

### **Modeling CRC using PDOs**

PDOs can be used to model disease, including tumour biology, novel drug targets and drug sensitivity. This is especially beneficial for research regarding a rare subtype of mCRC for which data from clinical trials are lacking. For example, mCRC patients with a deficient mismatch repair (dMMR) tumour, which is observed in 5% of mCRC patients<sup>6</sup>. As reported in **Chapter 2**, we lack evidence for which type of chemotherapy and targeted treatment is beneficial in mCRC patients with dMMR tumours. In **Chapter 3**, we used a panel of PDOs derived from patients with dMMR CRC tumours and proficient mismatch repair (pMMR) CRC tumours to assess the tumour cell-intrinsic treatment sensitivity in dMMR versus pMMR PDOs. Moreover, we were able to compare drug responses between PDOs based on *BRAF* and *KRAS* mutational

status. In our panel of PDOs, there was no significant difference in sensitivity to commonly used chemotherapeutic agents and targeted treatments. Thus, although patients with dMMR mCRC have a poor prognosis, we do not see a decreased tumour-cell intrinsic sensitivity to treatment. Potentially, the poor prognosis may be caused by differences in the microenvironments of these patients<sup>7</sup>. Addressing this requires the generation of PDO co-culture models in which these distinct microenvironments are modeled, for instance by co-culturing PDOs with specific immune cell types, or by transplanting PDOs into mice with a human immune system<sup>8,9</sup>. Since the intrinsic tumour cells are still sensitive to treatment, we hypothesize that these patients may be rendered sensitive to treatments by targeting the microenvironment.

Our research would not have been possible without the availability of PDOs from biobanks. Fortunately, PDOs derived from mCRC patients are increasingly available for translational research through biobanks. Ideally, biobank PDOs are characterized on a molecular and genetic (DNA and RNA sequencing) basis and coupled to the patient's clinical data. Preferably, the biobank database includes relevant tumour and patient factors which have been shown to influence prognosis and relevant treatment data. A published consensus statement on essential patient characteristics in systemic treatment trials for mCRC may serve as a basis for which variables to collect<sup>10</sup>. A fully characterized biobank allows researchers to analyze a specific cohort of PDOs, e.g. PDOs from patients with a right-sided primary tumour location and a *BRAF*-V600E mutation who were resistant for adjuvant treatment, which closely reflects the patient population for the clinical problem. Thus, hopefully, resulting in more efficient translation of results to the clinic.

## Identifying new treatment opportunities for mCRC patients using PDOs

PDOs are amenable to high throughput drug screens and thus allow for efficient screening of different treatments. An interesting application is the use of PDOs to screen combinations of treatments to identify new treatment regimens for a given subtype of CRC patients, such as *RAS*-mutant CRC<sup>11</sup>. A recent study demonstrated that synergistic interactions for treatments are most frequently seen for agents with weak single-agent effectivity and which are separated by 1-2 nodes in a protein-protein interaction network<sup>12</sup>. These findings may serve as a basis for identifying novel treatment combinations for mCRC patients.

A circular approach for using PDOs in studies may offer advantages for research in drug discovery and identifying novel biomarkers. In this approach, PDOs can be used to identify novel treatment targets for a given subtype of CRC. When validating these results in a clinical trial, PDOs may be established from patients receiving the given treatment. These PDOs, which would be characterized for treatment response, can then further guide research to identify novel biomarkers for treatment response or identify strategies to overcome treatment resistance.

## **Organoid drug screens as a predictive biomarker for mCRC patients**

Through individualized tumour response testing on PDOs, PDOs may act as a predictive biomarker for treatment response in cancer patients. The review in **Chapter 4** showed that the clinical validity of using PDOs as a predictive biomarker is increasingly being confirmed. PDO-based drug screens are able to discriminate clinical response with a pooled sensitivity and specificity of 0.81 (95% C.I. 0.69-0.89) and 0.74 (95% C.I. 0.64-0.82), respectively, across different tumour types and treatments. Novel studies since the publication of our review have further increased the evidence for clinical validity in CRC<sup>13-17</sup>. One topic of debate includes whether PDOs can predict treatment response for oxaliplatin-based regimens, as contrasting results have been shown<sup>17-20</sup>. Fortunately, novel studies confirm that PDOs can also predict response for oxaliplatin-based regimens<sup>13-15,17</sup>. One of these studies included a blinded study design in stage IV CRC patients prior to starting neoadjuvant or post-operative chemotherapy with oxaliplatin or irinotecan-based chemotherapy regimens<sup>13</sup>. Cut-off values for organoid response to predict clinical response (complete or partial response at 6 month follow-up) were determined in a pilot with 21 patients<sup>13</sup>. Subsequently, in a blinded study in 45 patients, PDO drug screens were found to have a sensitivity of 63.3% (19/30 PDOs), specificity of 94.1% (32/34 PDOs) and accuracy of 79.7% (51/64 PDOs) in predicting clinical response<sup>13</sup>. Another notable novel study includes Park *et al.*, which reported a machine learning algorithm to predict clinical response to radiation (tumour regression grade; TRG) in 19 patients with rectal cancer based on the survival fraction of PDOs after radiation<sup>16</sup>. The prediction model for good response (TRG = 0) had an AUC of 0.918 and accuracy of 81.5% while the model for poor response (TRG = 3) had an AUC of 0.971 and accuracy of 92.1%<sup>16</sup>. Ideally future studies will validate the results in larger cohorts of patients to confirm that PDOs can predict response for all

treatment types and disease stages (e.g. adjuvant, first line treatment and later line treatment).

**Chapter 5** of this thesis presented novel results for PDOs as a predictive biomarker in mCRC patients. Our results demonstrate that PDOs can predict treatment response in patients receiving 5-fluorouracil monotherapy, irinotecan-based regimens and panitumumab targeted treatment. PDOs were best able to predict treatment response when examining the change in tumour size of the target lesions in response scans ('percentage size change') as patient response and the area under the curve of the growth rate-derived drug response curve ( $GR_{AUC}$ ). Interestingly, organoids had a differential response to panitumumab based on the *RAS/BRAF* mutational status and primary tumour location, which has not previously been shown.

An important aspect which has not been proven is the clinical utility in using PDOs to tailor treatment for patients. Studies need to assess if patients who receive PDO-guided treatment benefit from this treatment through increased survival or quality of life compared to patients receiving standard-of-care treatment and if this is cost-effective. Ideally, an intervention trial randomizes patients between PDO-guided treatment and standard-of-care treatment via the oncologist. A few studies have reported PDO-guided treatment for patients with refractory mCRC<sup>20,21</sup>, however the clinical benefit in patients has not been proven. In a single-arm prospective trial in refractory mCRC patients, 6 of the 61 included patients were able to receive PDO-guided treatment and these patients did not have a clinical response to the treatment received<sup>21</sup>. The low number of patients receiving PDO-guided treatment were largely a result of the PDO establishment rate (54% in patients who underwent a biopsy) and later line setting, with patients experiencing clinical deterioration while PDOs were being established and screened. The clinical utility may be higher in first line mCRC patients. In randomized clinical trials, oxaliplatin and irinotecan-based combination regimens show similar effectiveness as first line treatment<sup>22</sup>. Nonetheless, PDOs show varying sensitivity to these treatments, suggesting that treatments can be tailored by PDO drug screening to further improve response rate and survival in patients. PDO-guided treatment can be used to avoid overtreatment by selecting patients for 5-fluorouracil monotherapy if sufficiently sensitive in organoid screens, rather than combination treatment, or to identify a more effective combination treatment by comparing organoid sensitivity for irinotecan versus oxaliplatin regimens.

The optimal setting for using PDOs to guide treatment in the clinic is still unknown. This may vary depending on the PDO establishment rate and time period to obtaining drug screen results. Additionally, it depends on the ability of PDOs to guide treatment for all treatment types, or only a select number of treatments. Theoretically, implementing PDO-guided treatment for patients earlier in the disease, e.g. to guide adjuvant treatment (5-fluorouracil versus oxaliplatin-doublet adjuvant treatment) or to guide first line treatment, will result in more patients able to benefit from personalized treatment and hopefully benefit from increased survival and reduced side effects from ineffective treatment or overtreatment. Using PDO-guided treatment may be more cost effective earlier in the disease, since patients may gain more quality-adjusted life years (QALY), with less side effects due to tailored treatments, compared to patients with later-stage disease<sup>23</sup>.

### **Future perspectives for research regarding improving CRC treatment using PDOs**

Many improvements and discoveries have been made regarding the use of PDOs to predict treatment response in the past decade since the pioneering publications, including the first CRC PDO in 2011<sup>2</sup>, the first prospective CRC biobank in 2015<sup>3</sup> and the first study indicating that PDOs drug screen results are associated with clinical response in 2018<sup>24</sup>. However, we are currently unable to use PDO-guided treatment in the clinic. Which obstacles should be overcome to ensure that using PDOs to guide treatment is meaningful and has added value over standard treatment, resulting in more effective personalized treatment?

Various aspects for the analytical validity, clinical validity and clinical utility of PDO as predictive biomarker need to be assessed. The establishment of PDOs should be optimized, since the pooled establishment rate is 68.5% per sample (95% C.I. 56.5-78.5; **Chapter 4**). As well, the time to establish sufficient PDOs for screens should be minimized, since this can be quite prolonged, depending on the origin<sup>25</sup>. Ideally, a consensus will be established for standardized protocols for the establishment and culturing of PDOs. A recent study established a pan-cancer platform to establish and screen PDOs from over 1,200 cancer patients with different types of tumours<sup>25</sup>. Innovations which decrease the number of PDOs required for reliable drug screen results will help minimize the time needed to obtain screen results and avoid treatment delays for patients. Novel techniques which incorporate PDO establishment and

screening in chips or imaging-based read-outs may guide these developments<sup>25–29</sup>. An exciting new application includes culturing PDOs from liquid biopsies with circulating tumour cells, which was recently done with a success rate of 87.8% in pancreatic cancer patients<sup>30</sup>. This would enable longitudinal culturing of PDOs in a minimally invasive manner and decrease the risk for patients by avoiding biopsy-associated complications.

To decrease variability in drug screening results, synthetic extracellular matrices should be used, rather than animal-based matrices, since animal-based matrices are biologically variable and contain animal-derived growth factors<sup>31–33</sup>. Drug screening and analysis should be optimized and standardized to increase the sensitivity and specificity of predicting clinical response and to allow comparison among studies. Drug screen methods still vary widely and it is currently unclear which method results in the strongest associations with clinical response. Some important aspects which may influence this include, the read-out used (e.g. metabolic assay versus more direct cell death measurements), the quantification method (e.g. growth rate metrics versus viability) and the screening combination treatment set-up (e.g. a consistent ratio between drugs versus titrating one drug relative to a fixed concentration of the second drug). Studies which systematically compare different methods and examine the effect on the association with clinical response will help provide clarity. Fortunately, PDO sensitivity and growth does not seem affected by passage number of the PDO, location of the PDOs within an extracellular matrix droplet or seeding density<sup>34</sup>. Imaging-based read-outs at the organoid-level offer several benefits over well-based read-outs, including shorter time to detect treatment effects, longitudinal measurement of response and heterogeneity in response<sup>34–36</sup>.

For drug sensitivity studies, a key issue in the use of PDOs as a model is the lack of tumour microenvironment (TME) components. The TME has a crucial role in modulating treatment response in patients<sup>37</sup>. Certain treatment types target the ECM, including immunotherapy and angiogenesis modulators (e.g. bevacizumab). As such, PDOs are currently unable to predict treatment response for treatments which target the TME. These issues may be overcome by using complex PDO co-cultures involving components of the TME or by transplanting PDOs into mice with a human immune system to model the immune system<sup>8,9,38,39</sup>.

## II - TAILORING TREATMENT WITH PREDICTION MODELS

Correctly estimating prognosis is an important element in clinical decision-making. However this is difficult, with both patients and physicians overestimating survival chances<sup>40</sup>. Prediction models can help in improved clinical decision-making by providing better estimates of prognosis for patients based on individual patient's characteristics. Prediction models relate multiple predictors for an individual to the probability or risk for a given outcome (e.g. recurrence of disease or survival)<sup>41</sup>. In this manner, a prediction model can estimate prognosis for patients who receive a given treatment and thus be used in clinical decision-making. This thesis examines using prediction models to estimate prognosis for mCRC patients who receive one of two types of treatments: i) local ablation or resection of colorectal liver metastases in mCRC patients (**Chapters 6 and 7**) and ii) trifluridine/tipiracil treatment for refractory mCRC (**Chapter 8**).

For patients with CRC metastases confined to the liver (CRLM), local treatment of CRLM through resection or tumour ablation offers the only chance for long-term survival compared to palliative systemic treatment. Five-year survival rates of up to 55% have been reported for local treatment of CRLM<sup>42–44</sup>. Through improved surgical techniques, more lenient resection criteria and optimization of induction systemic therapy have increased the number of patients with CRLM that are considered technically resectable<sup>45,46</sup>. Unfortunately, up to 75% of patients relapse after liver resection<sup>47–49</sup> and a subgroup of patients have no long-term survival benefit. This underscores the urgent need to improve risk-stratification prior to surgery<sup>50</sup>. Being able to stratify a patient's prognosis prior to CRLM treatment would aid in clinical decision-making and may help assess which patients may benefit from peri-operative systemic treatment. Another application of a prediction model in this patient group includes stratifying patients for clinical trials, which may include comparing local treatment combined with systemic treatment versus local treatment alone or comparing local treatment versus palliative systemic treatment.

**Chapter 6** evaluated the generalizability and clinical validity of two prediction models, the widely used Fong score<sup>51</sup> and the more recent GAME score<sup>52</sup>, in a nationwide population-based cohort of patients with CRC and liver metastases. Existing models predicting prognosis of patients after resection of colorectal liver metastases were

developed in highly specialized centers and thus may not function in the general population. The prediction models were validated in two important subgroups: in patients with and without perioperative systemic therapy and in patients  $\leq 70$  and  $> 70$  years. Internationally, there are different guidelines regarding the use of perioperative systemic treatment<sup>53–55</sup>, reinforcing the need to examine the performance in this subgroup. In our population-based cohort, both prediction models showed predictive ability, with a concordance index (C-index) of 0.596 (95% C.I. 0.572–0.621) for the GAME score and 0.577 (95% C.I. 0.554–0.601) for the Fong score. Results were comparable in subgroups. A C-index of 0.6 means that for a pair of patients with and without the outcome, the prediction model correctly identifies the patient with the outcome in 60%<sup>56</sup>. Accordingly, 25% of patients identified as “high-risk” according to the GAME score did achieve long-term survival which exceeded five years and this rate was even higher in the Fong high-risk group. Thus, the prediction models were not able to sufficiently discriminate preoperative risk for clinical decision-making and require improvement. Our results reinforce the need to externally validate existing prediction models in real-world populations in order to identify obstacles prior to implementing the models in clinical practice.

Consequently, we developed a novel prediction model for patients undergoing local treatment of CRLM using our nationwide population-based cohort of patients with CRC and liver metastases (**Chapter 7**). Our model predicts the risk of an early extrahepatic recurrence (EHR) or death, i.e. within six months after local treatment, based on the following pre-operative variables: location of the primary tumour (‘sidedness’), T-stage, N-stage, *RAS* mutational status, *BRAF* mutational status, number of liver metastases and size of the largest liver metastasis. The end-point early EHR was chosen based on expert opinion, since patients who experience an early EHR have a poor prognosis after local treatment and are often not eligible for repeat resections, in contrast to patients with liver-only recurrences<sup>50,57–61</sup>. Thus, the added value of local treatment may not be justified in patients with an early EHR after local treatment of CRLM. This was reflected in the OS of this subgroup of patients. Our model improves on existing prediction models in the methods used to develop the model (avoiding cut-offs of continuous variables, modeling non-linear variables using splines, adequately handling missing data using multiple imputation), the large, high-quality data of population-based patients and by incorporating relevant, modern prognostic factors (including sidedness and *BRAF/RAS* mutational status).

In our cohort of patients undergoing CRLM local treatment, *BRAF* mutations are rare (2.7%). However these patients have a lower EHR-free survival (11.4 months) compared to patients with a *RAS* mutation (18.5 months) or wildtype tumour (28.2 months). Our EHR prediction model discriminated between patients based on EHR rates, reflected in differing EHR-free survival, recurrence-free survival and overall survival. The full model specifications have been published to enable external validation of our prediction model, which we were unable to do. To our knowledge, this is the first model predicting early EHR in patients after local treatment of CRLM. The EHR prediction model can be used to confirm that local treatment should be pursued in low-risk patients, but did not show benefit in avoiding local treatment in high-risk patients. The EHR prediction model may also aid clinical decision-making by identifying high-risk patients for early EHR who may benefit from peri-operative systemic treatment. A treatment strategy for high-risk patients may be to initiate long-lasting systemic treatment, and upon sustained response, perform local treatment of CRLM. External validation studies confirming the validity of our model in other patient populations receiving local treatment for CRLM and clinical impact studies confirming clinical utility in informing clinical decision-making using the prediction model are needed before the model can be implemented.

**Chapter 8** externally validated an existing prediction model (the Colon Life nomogram<sup>62</sup>) in 150 QUALITAS<sup>63</sup> patients receiving trifluridine/tipiracil treatment in daily practice for refractory mCRC. Trifluridine/tipiracil treatment improves OS in patients with refractory mCRC from 5.3 months (placebo) to 7.1 months and progression-free survival from 1.7 months (placebo) to 2.0 months<sup>64</sup>. There are currently no valid biomarkers or prediction models implemented in clinical practice which may identify patients who have a better prognosis during treatment. The Colon Life nomogram is a model which predicts 12-week mortality for patients with refractory mCRC receiving treatment<sup>62</sup>. The Colon Life nomogram can help guide clinical decision-making, since in patients with a high risk of 12-week mortality, trifluridine/tipiracil treatment may not be appropriate and result in unnecessary toxicity for patients. The model is however not widely used. Based on its discriminatory ability in external validation studies, the Colon Life nomogram was found to be “an accurate tool for estimating life expectancy and the risk of death” (C-statistic ranged from 0.657 to 0.788), however the calibration plots indicated that the model is not well calibrated in the external validation cohorts<sup>65–67</sup>. Our study results confirmed the

model's miscalibration with evident overestimation of 12-week mortality (calibration intercept of -1.06 and slope of 0.41) in QUALITAS patients. Decision curve analysis can be used to examine the net benefit when using model-informed clinical decisions versus treating all patients or none<sup>68,69</sup>. We showed that treatment decisions using the nomogram did not yield a net benefit at clinically meaningful thresholds. In addition to a model's discriminatory ability, its calibration is an important performance measure which should be confirmed in external validation studies<sup>70</sup>. The miscalibration of the Colon Life nomogram hampers its use in clinical practice.

In **Chapter 8**, we also examined the additional prognostic value of quality of life (QoL) at start of treatment to the Colon Life prediction model<sup>63</sup>. A low QoL at start of treatment was independently associated with shorter OS and PFS<sup>3</sup>. In our study, addition of QoL to the model improved the discrimination of the model (C-statistic) to 0.85 (0.81-0.89). Compared to the physician-reported performance score, which is already incorporated in the Colon Life nomogram, QoL likely better reflects patient functioning<sup>71,72</sup>. With the increasing use of patient-reported outcome measurements (PROMs), incorporating QoL in a prediction model is feasible in daily practice. We recommend conducting a study with sufficient power, e.g. through the use of individual patient data, to update the Colon Life model or to develop a new prediction model for patients receiving trifluridine/tipiracil treatment.

Prediction model research lacks impact studies for externally validated models<sup>56</sup>. Impact studies examine the effect of using a prognostic model on clinical decision-making and patient outcome. For our prediction models the following potential impact studies can be envisioned. For the novel CRLM model predicting early EHR, a prospective trial might examine if the implementation of the prediction model estimates for early EHR in a tumour decision board affect the clinical decision-making of surgeons and oncologists compared to patients where the prognostic results were not examined using the prediction model. As well, a study could compare the quality of life and survival in patients who receive counseling with and without the prognostic model. An interesting patient category would be elderly patients, who might be offered a beneficial local treatment and would otherwise not have received this advice. An extension might be a trial examining the benefit of systemic treatment (e.g. peri-operative or adjuvant) to improve prognosis for high risk patients. Similarly, a novel/updated model predicting 12-week mortality in patients receiving trifluridine/tipiracil treatment could be examined in an impact study. How does the use of the

prognostic model impact shared decision-making by the patient and oncologist and affect quality of life and survival of the patient? Additionally, it may be beneficial to examine potential differences in the patient's own expectations for prognosis (with and without the validation model) versus predicted prognosis and observed prognosis, since patients often overestimate their survival chances<sup>40</sup>. Impact studies are needed to assess the added value of using a prognostic model in the clinic and thus, hopefully, improve treatment for mCRC patients. Realizing impact studies with sufficient power requires funding, knowledge and adequate patient inclusions, which can only be realized through close collaboration of all parties involved (researchers, clinicians and potentially pharmaceutical companies). Depending on the patient population, international collaboration may be needed to obtain the required patient inclusions.

### III - USING REAL-WORLD DATA REGISTRIES TO ADDRESS TREATMENT CHALLENGES

An overarching theme in this thesis is the use of real-world registry data (**Chapters 1, 6, 7 and 8**). Real-world data include electronic health records, billing/insurance databases and disease registries, thus acquiring data outside of clinical trials<sup>73</sup>. Studies using real-world data are seen as an important complement to trial data, due to the diverse non-selected population, allowing an assessment of treatment effectiveness, toxicity, rare events and long-term outcomes<sup>73,74</sup>. Various examples of using real-world data to address clinical problems were explored in this manuscript, including examining the survival of patients with a rare subtype of CRC (deficient mismatch repair; **Chapter 1**), which provided data to interpret single-arm treatment trials. As well, a real-world cohort of patients with local treatment for CRLM was used to externally validate existing prediction models (**Chapter 6**) and develop a new prediction model (**Chapter 7**), allowing us to evaluate the generalizability of existing prediction models in this population and to create an improved prediction model which has optimal performance in real-world patients. Additionally, a national data registry was used to collect clinical data for QUALITAS trial patients, in which we externally validated an existing prediction model and examined the additional prognostic value of two new variables (QoL and TK1 expression; **Chapter 8**).

The research in this thesis would not have been possible without high-quality data collected and shared for research by the Netherlands Cancer Registry (NCR) and the Prospective Dutch CRC cohort (PLCRC), a nationwide observational cohort study

linked to the NCR. Using data from national registries better reflects the “every day” patient and thus may result in more efficient translation of research results to a larger group of patients<sup>74</sup>. For rare disease subtypes of groups of patients, national registries gives access to larger patient numbers rather than slow accrual via trials. In order to enable innovative research, registries should ideally collaborate with researchers and clinicians to ensure that the data which is collected contains all relevant characteristics. Challenges include potential confounders, missing data and limitations in the type of research which can be performed. Data collection for registries can be facilitated by using standardized reporting and integrating all relevant platforms in electronic patient files. Hopefully, the use of real-world data will decrease health care costs, increase efficiency in translating clinical trials results to every day patients and help in validating/identifying new prognostic and predictive biomarkers and prediction models. This can be facilitated if real-world data sources proactively collaborate with researchers and clinicians to assess if and how new prognostic variables or outcomes need to be collected.

The PLCRC cohort serves as a research infrastructure which collects extensive longitudinal clinical data, biospecimens intertwined in routine clinical care (e.g. blood, tumour tissue) and patient-reported outcomes in patients with stage I-IV CRC from primary diagnosis until death<sup>75</sup>. Two research projects in this thesis were embedded in PLCRC, including the OPTIC clinical trial (**Chapter 5**) and the QUALITAS trial (**Chapter 8**). The added advantage of the PLCRC is that it allows researchers to access more extensive clinical data and analyze additional biomarkers through the collected biospecimens. Since novel molecular markers are being discovered, it allows researchers to use real-world patients while still analysing new markers which otherwise would not be known from electronic patient files.

## CONCLUDING REMARKS

The manner in which cancer is treated will be drastically improved when clinicians are able to offer all patients personalized treatment. Despite innovations in treatments and biomarkers, the majority of mCRC patients are still unable to receive personalized treatment. The research in this thesis highlighted strategies to facilitate personalized treatment in mCRC patients: examining treatment sensitivity in a subgroup of mCRC patients with dMMR tumours, individualized tumour-response testing using PDOs, informing clinical decision-making using prediction models and using real-world

data to facilitate these strategies. Fortunately, immunotherapy has vastly improved survival for patients with dMMR mCRC. Further improvements can be made by understanding the reason why these patients have a poor prognosis while receiving chemotherapy and targeted treatment and identifying strategies to overcome this. Our results revealed that we lack clinical evidence for a differential sensitivity for a given type of chemotherapy or targeted treatment in these patients. Moreover, *in vitro* studies revealed that dMMR PDOs have similar sensitivity to treatment compared to pMMR PDOs, suggesting that differences in the tumour microenvironment may offer insight into overcoming treatment resistance in these patients. Patient-derived organoids may revolutionize personalized treatment by enabling a tailored treatment strategy for a much larger proportion of oncology patients. We examined the available evidence, provided concrete recommendations for future research and confirmed that PDOs can be used to predict treatment response in mCRC patients for standard-of-care treatments. Studies optimizing the feasibility of using PDOs to guide treatment, confirming results in specific patient cohorts and assessing the clinical utility, are needed. Highly anticipated are the results of studies examining if patients benefit from receiving PDO-guided treatment. These studies can evaluate cost effectiveness of PDO-guided treatment and will help examine if using PDOs to guide clinical decisions is meaningful. Lastly, prediction models can tailor treatment by selecting patients for treatment based on their predicted prognosis. In daily practice patients, external validation of existing prediction models for patients receiving local treatment for CRLM and for patients receiving trifluridine/tipiracil revealed that the models need to be improved. We developed a novel prediction model for early EHR in patients receiving local treatment for CRLM, which may aid in optimizing treatment strategies for these patients. We revealed that adding quality of life to an existing prediction model for trifluridine/tipiracil patients increases model performance. Our results need to be validated in other patient populations and assessed in clinical impact studies to assess that the use of the prediction models to guide clinical decision-making is beneficial. For clinicians and researchers, this is an exciting opportunity to initiate new research projects to confirm the clinical impact of using these strategies to personalize treatment for mCRC patients, ultimately bringing us one step closer to being able to provide personalized treatment for all patients with mCRC.

## REFERENCES

1. Hamers P.A.H., Elferink M.A.G., Stellato R.K., Punt C.J.A., May A.M., Koopman M., et al. Informing metastatic colorectal cancer patients by quantifying multiple scenarios for survival time based on real-life data. *Int J Cancer*. (2021);**148**(2):296–306.
2. Sato T., Stange D.E., Ferrante M., Vries R.G.J., Van Es J.H., Van Den Brink S., et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology*. (2011);**141**(5):1762–72.
3. Van De Wetering M., Francies H.E., Francis J.M., Bounova G., Iorio F., Pronk A., et al. Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell*. (2015);**161**(4):933–45.
4. Pampaloni F., Reynaud E.G., Stelzer E.H.K. The third dimension bridges the gap between cell culture and live tissue. *Nat Rev Mol Cell Biol*. (2007);**8**(10):839–45.
5. Li M., Izpisua Belmonte J.C. Organoids - Preclinical models of human disease. *NEJM*. (2019);**380**(6):569–79.
6. Koopman M., Kortman G.A.M., Mekenkamp L., Ligtenberg M.J.L., Hoogerbrugge N., Antonini N.F., et al. Deficient mismatch repair system in patients with sporadic advanced colorectal cancer. *Br J Cancer*. (2009);**100**(2):266–73.
7. McMillin D.W., Negri J.M., Mitsiades C.S. The role of tumour–stromal interactions in modifying drug response: challenges and opportunities. *Nat Rev Drug Discov*. (2013);**12**(March):217–28.
8. Dijkstra K.K., Cattaneo C.M., Weeber F., Chalabi M., van de Haar J., Fanchi L.F., et al. Generation of Tumor-Reactive T Cells by Co-culture of Peripheral Blood Lymphocytes and Tumor Organoids. *Cell*. (2018);**174**(6):1586–1598.e12.
9. Neal J.T., Li X., Zhu J., Zheng G.X.Y., Davis M.M., Kuo C.J., et al. Organoid Modeling of the Tumor Immune Resource Organoid Modeling of the Tumor Immune Microenvironment. *Cell*. (2018);**175**(7):1972–1988.e16.
10. Goey K.K.H., Sørbye H., Glimelius B., Arnold D., Berlin J.D., Adams R.A., et al. Consensus statement on essential patient characteristics in systemic treatment trials for metastatic colorectal cancer : Supported by the ARCAD Group. *Eur J Cancer*. (2018);**100**:35–45.
11. Verissimo C.S., Overmeer R.M., Ponsioen B., Drost J., Mertens S., Verlaan-Klink I., et al. Targeting mutant RAS in patient-derived colorectal cancer organoids by combinatorial drug screening. *Elife*. (2016);**5**(NOVEMBER2016):1–26.
12. Jaaks P., Coker E.A., Vis D.J., Edwards O., Carpenter E.F., Leto S.M., et al. Effective drug combinations in breast, colon and pancreatic cancer cells. *Nature*. (2022);**603**:166–173.
13. Wang T., Pan W., Zheng H., Zheng H., Wang Z., Li J.J., et al. Accuracy of Using a Patient-Derived Tumor Organoid Culture Model to Predict the Response to Chemotherapy Regimens in Stage IV Colorectal Cancer: A Blinded Study. *Dis Colon Rectum*. (2021);833–50.
14. Cho Y.W., Min D.W., Kim H.P., An Y., Kim S., Youk J., et al. Patient-derived organoids as a preclinical platform for precision medicine in colorectal cancer. *Mol Oncol*. (2021);1–17.
15. Yao L., Zao X.-L., Pan X.-F., Zhang H.-G., Wang F.-J., Qiao P.-F. Application of tumoroids derived from advanced colorectal cancer patients to predict individual response to chemotherapy. *J Chemother*. (2022);**14**:1–13.

16. Park M., Kwon J., Kong J., Moon S.M., Cho S., Yang K.Y., et al. A Patient-Derived Organoid-Based Radiosensitivity Model for the Prediction of Radiation Responses in Patients with Rectal Cancer. *Cancers (Basel)*. (2021);**13**(15):3760.
17. Geevimaan K., Guo J., Shen C., Jiang J., Fann C.S.J., Hwang M., et al. Patient-Derived Organoid Serves as a Platform for Personalized Chemotherapy in Advanced Colorectal Cancer Patients. *Front Oncol*. (2022);**12**(June):1–15.
18. Ganesh K., Wu C., O'Rourke K.P., Szeglin B.C., Zheng Y., Sauvé C.-E.G., et al. A rectal cancer organoid platform to study individual responses to chemoradiation. *Nat Med*. (2019);**25**(10):1607–14.
19. Ooft S.N., Weeber F., Dijkstra K.K., McLean C.M., Kaing S., van Werkhoven E., et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Sci Transl Med*. (2019);**11**(513):1–10.
20. Narasimhan V., Wright J.A., Churchill M., Wang T., Rosati R., Lannagan T.R., et al. Medium-throughput drug screening of patient-derived organoids from colorectal peritoneal metastases to direct personalized therapy. *CCR*. (2020);**26**:3662–70.
21. Ooft S.N., Weeber F., Schipper L., Dijkstra K.K., McLean C.M., Kaing S., et al. Prospective experimental treatment of colorectal cancer patients based on organoid drug responses. *ESMO Open*. (2021);**6**(3):100103.
22. Kawai S., Takeshima N., Hayasaka Y., Notsu A., Yamazaki M., Kawabata T., et al. Comparison of irinotecan and oxaliplatin as the first-line therapies for metastatic colorectal cancer: a meta-analysis. *BMC Cancer*. (2021);**21**(1):1–11.
23. Dvortsin E., Gout-Zwart J., Eijssen E.L.M., Van Brussel J., Postma M.J. Comparative cost-effectiveness of drugs in early versus late stages of cancer; Review of the literature and a case study in breast cancer. *PLoS One*. (2016);**11**(1):1–12.
24. Vlachogiannis G., Hedayat S., Vatsiou A., Jamin Y., Fernández-mateos J., Khan K., et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science*. (2018);**359**(6378):920–6.
25. Larsen B.M., Kannan M., Langer L.F., Khan A.A., Igartua C., Salahudeen A.A., et al. A pan-cancer organoid platform for precision medicine II A pan-cancer organoid platform for precision medicine. *Cell Rep*. (2021);**36**(4):109429.
26. Du Y., Li X., Niu Q., Mo X., Qui M., Ma T., et al. Development of a miniaturized 3D organoid culture platform for ultra-high-throughput screening. *J Mol Cell Biol*. (2020);**12**(8):630–43.
27. Kasendra M., Tovaglieri A., Sontheimer-Phelps A., Jalili-Firoozinezhad S., Bein A., Chalkiadaki A., et al. Development of a primary human Small Intestine-on-a-Chip using biopsy-derived organoids. *Sci Rep*. (2018);**8**(1):1–14.
28. Renner H., Grabos M., Becker K.J., Kagermeier T.E., Wu J., Otto M., et al. A fully automated high-throughput workflow for 3D-based chemical screening in human midbrain organoids. *Elife*. (2020);**9**:1–39.
29. Pinho D., Santos D., Vila A., Carvalho S. Establishment of Colorectal Cancer Organoids in Microfluidic-Based System. *Micromachines*. (2021);**12**(497):1–12.
30. Wu Y.-H., Hung Y.-P., Chiu N.-C., Lee R.-C., Li C.-P., Chao Y., et al. Correlation between drug sensitivity profiles of circulating tumour cell-derived organoids and clinical treatment response in patients with pancreatic ductal adenocarcinoma. *Eur J Cancer*. (2022);**166**:208–18.

31. Vukicevic S., Kleinman H.K., Luyten F.P., Roberts A.B., Roche N.S., Reddi A.H. Identification of multiple active growth factors in basement membrane matrigel suggests caution in interpretation of cellular activity related to extracellular matrix components. *Exp Cell Res.* (1992);**202**(1):1–8.
32. Hernandez-Gordillo V., Kassir T., Lampejo A., Choi G.H., Gamboa M.E., Gnecco J.S., et al. Fully synthetic matrices for in vitro culture of primary human intestinal enteroids and endometrial organoids. *Biomaterials.* (2020);**254**(May):120125.
33. Ng S., Tan W.J., Pek M.M.X., Tan M.H., Kurisawa M. Mechanically and chemically defined hydrogel matrices for patient-derived colorectal tumor organoid culture. *Biomaterials.* (2019);**219**(July):119400.
34. DeStefanis R.A., Olson A.M., DeZeeuw A.K., Gillette A.A., Sha G.C., Johnson K.A., et al. Impact of baseline culture conditions of mouse-derived cancer organoids when determining therapeutic response and tumor heterogeneity. *Sci Rep.* (2022);**12**:5205.
35. Pasch C.A., Favreau P.F., Yueh A.E., Babiarz C.P., Gillette A.A., Sharick J.T., et al. Patient-derived cancer organoid cultures to predict sensitivity to chemotherapy and radiation. *Clin Cancer Res.* (2019);**25**(17):5376–87.
36. Sharick J.T., Walsh C.M., Sprackling C.M., Pasch C.A., Pham D.L., Esbona K., et al. Metabolic Heterogeneity in Patient Tumor-Derived Organoids by Primary Site and Drug Treatment. *Front Oncol.* (2020);**10**(May):1–17.
37. Son B., Lee S., Youn H., Kim E., Kim W., Youn B. The role of tumor microenvironment in therapeutic resistance. *Oncotarget.* (2017);**8**(3):3933–45.
38. Votanopoulos K.I., Forsythe S., Sivakumar H., Mazzocchi A., Aleman J., Miller L., et al. Model of Patient-Specific Immune-Enhanced Organoids for Immunotherapy Screening: Feasibility Study. *Ann Surg Oncol.* (2019);**27**(6):1956–67.
39. Chalabi M., Fanchi L.F., Dijkstra K.K., Van den Berg J.G., Aalbers A.G., Sikorska K., et al. Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat Med.* (2020);**26**(April).
40. Young A.L., Lee E., Absolom K., Baxter H., Christophi C., Lodge J.P.A., et al. Expectations of outcomes in patients with colorectal cancer. *BJS Open.* (2018);**2**(5):285–92.
41. Moons K.G.M., Altman D.G., Reitsma J.B., Ioannidis J.P.A., Macaskill P., Steyerberg E.W., et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med.* (2015);**162**(1):W1–73.
42. Rees M., Tekkis P.P., Welsh F.K., O'Rourke T., John T.G. Evaluation of long-term survival after hepatic resection for metastatic colorectal cancer: a multifactorial model of 929 patients. *Ann Surg.* (2008);**247**(1):125–35.
43. Kanas G.P., Taylor A., Primrose J.N., Langeberg W.J., Kelsh M.A., Mowat F.S., et al. Survival after liver resection in metastatic colorectal cancer: review and meta-analysis of prognostic factors. *Clin Epidemiol.* (2012);**4**:283–301.
44. Andreou A., Aloia T.A., Brouquet A., Dickson P. V., Zimmiti G., Maru D.M., et al. Margin status remains an important determinant of survival after surgical resection of colorectal liver metastases in the era of modern chemotherapy. *Ann Surg.* (2013);**257**(6):1079–88.
45. Kopetz S., Chang G.J., Overman M.J., Eng C., Sargent D.J., Larson D.W., et al. Improved survival in metastatic colorectal cancer is associated with adoption of hepatic resection and improved chemotherapy. *J Clin Oncol.* (2009);**27**(22):3677–83.

46. Bolhuis K., Kos M., van Oijen M.G.H., Swijnenburg R.J., Punt C.J.A. Conversion strategies with chemotherapy plus targeted agents for colorectal cancer liver-only metastases: A systematic review. *Eur J Cancer*. (2020);**141**:225–38.
47. Nordlinger B., Sorbye H., Glimelius B., Poston G.J., Schlag P.M., Rougier P., et al. Perioperative chemotherapy with FOLFOX4 and surgery versus surgery alone for resectable liver metastases from colorectal cancer (EORTC Intergroup trial 40983): a randomised controlled trial. *Lancet*. (2008);**371**(9617):1007–16.
48. de Jong M.C., Pulitano C., Ribero D., Strub J., Mentha G., Schulick R.D., et al. Rates and patterns of recurrence following curative intent surgery for colorectal liver metastasis: an international multi-institutional analysis of 1669 patients. *Ann Surg*. (2009);**250**(3):440–8.
49. Jones R.P., Jackson R., Dunne D.F., Malik H.Z., Fenwick S.W., Poston G.J., et al. Systematic review and meta-analysis of follow-up after hepatectomy for colorectal liver metastases. *Br J Surg*. (2012);**99**(4):477–86.
50. Imai K., Allard M.A., Benitez C.C., Vibert E., Sa Cunha A., Cherqui D., et al. Early Recurrence After Hepatectomy for Colorectal Liver Metastases: What Optimal Definition and What Predictive Factors? *Oncologist*. (2016);**21**(7):887–94.
51. Fong Y., Fortner J., Sun R.L., Brennan M.F., Blumgart L.H. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. *Ann Surg*. (1999);**230**(3):309–21.
52. Margonis G.A., Sasaki K., Gholami S., Kim Y., Andreatos N., Rezaee N., et al. Genetic And Morphological Evaluation (GAME) score for patients with colorectal liver metastases. *Br J Surg*. (2018);**105**(9):1210–20.
53. National Comprehensive Cancer Network. Colon Cancer (version 2.2021) [Internet]. Available from: [https://www.nccn.org/professionals/physician\\_gls/pdf/colon.pdf](https://www.nccn.org/professionals/physician_gls/pdf/colon.pdf)
54. Nederlandse Vereniging voor Heelkunde. Colorectaal Carcinoom (CRC) (version 2019) [Internet]. Available from: [file:///H:/Downloads/Colorectaal\\_carcinoom\\_CRC\\_\(1\).pdf](file:///H:/Downloads/Colorectaal_carcinoom_CRC_(1).pdf)
55. Van Cutsem E., Cervantes A., Adam R., Sobrero A., Van Krieken J.H., Aderka D., et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol*. (2016);**27**(8):1386–422.
56. Hoesseini A., Van Leeuwen N., Sewnaik A., Steyerberg E.W., Baatenburg De Jong R.J., Lingsma H.F., et al. Key Aspects of Prognostic Model Development and Interpretation from a Clinical Perspective. *JAMA Otolaryngol - Head Neck Surg*. (2022);**148**(2):180–6.
57. Stewart C.L., Warner S., Ito K., Raoof M., Wu G.X., Kessler J., et al. Cytoreduction for colorectal metastases: liver, lung, peritoneum, lymph nodes, bone, brain. When does it palliate, prolong survival, and potentially cure? *Curr Probl Surg*. (2018);**55**(9):330–79.
58. Kanemitsu Y., Shimizu Y., Mizusawa J., Inaba Y., Hamaguchi T., Shida D., et al. Hepatectomy Followed by mFOLFOX6 Versus Hepatectomy Alone for Liver-Only Metastatic Colorectal Cancer (JCOG0603): A Phase II or III Randomized Controlled Trial. *J Clin Oncol*. (2021);**39**(34):3789–99.
59. Okuno M., Goumard C., Kopetz S., Vega E.A., Joechle K., Mizuno T., et al. RAS Mutation is Associated with Unsalvageable Recurrence Following Hepatectomy for Colorectal Cancer Liver Metastases. *Ann Surg Oncol*. (2018);**25**(8):2457–66.
60. Nanji S., Tsang M.E., Wei X., Booth C.M. Outcomes after repeat hepatic resection for recurrent metastatic colorectal cancer: A population-based study. *Am J Surg*. (2017);**213**(6):1053–9.

61. Matsuoka H., Morise Z., Tanaka C., Hayashi T., Ikeda Y., Maeda K., et al. Repeat hepatectomy with systemic chemotherapy might improve survival of recurrent liver metastasis from colorectal cancer-a retrospective observational study. *World J Surg Oncol.* (2019);**17**(1):33.
62. Pietrantonio F., Miceli R., Rimassa L., Lonardi S., Aprile G., Mennitto A., et al. Estimating 12-week death probability in patients with refractory metastatic colorectal cancer: The Colon Life nomogram. *Ann Oncol.* (2017);**28**(3):555–61.
63. Hamers P.A.H., Vink G.R., Elferink M.A.G., Stellato R.K., Dijksterhuis W.P.M., Punt C.J.A., et al. Quality of life and survival of metastatic colorectal cancer patients treated with trifluridine-tipiracil (QUALITAS). *Clin Colorectal Cancer.* (2022);**21**(2):154–66.
64. Mayer R.J., Van Cutsem E., Falcone A., Yoshino T., Garcia-Carbonero R., Mizunuma N., et al. Randomized Trial of TAS-102 for Refractory Metastatic Colorectal Cancer. *NEJM.* (2015);**372**(20):1909–19.
65. Cremolini C., Rossini D., Martinelli E., Pietrantonio F., Lonardi S., Noventa S., et al. Trifluridine/Tipiracil (TAS-102) in Refractory Metastatic Colorectal Cancer: A Multicenter Register in the Frame of the Italian Compassionate Use Program. *Oncologist.* (2018);**23**(10):1178–87.
66. Pietrantonio F., Fucà G., Manca P., Pagani F., Raimondi A., Prisciandaro M., et al. Validation of the Colon Life nomogram in patients with refractory metastatic colorectal cancer enrolled in the RECOURSE trial. *Tumori.* (2021);**107**(4):353–9.
67. Fernández Montes A., Carmona-Bayonas A., Jimenez-Fonseca P., Vázquez Rivera F., Martínez Lago N., Covela Rúa M., et al. Prediction of survival in patients with advanced, refractory colorectal cancer in treatment with trifluridine/tipiracil: real-world vs clinical trial data. *Sci Rep.* (2021);**11**(1):1–10.
68. Vickers A.J., Elkin E.B. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak.* (2006);**26**(6):565–74.
69. Rousson V., Zumbo T. Decision curve analysis revisited: Overall net benefit, relationships to ROC curve analysis, and application to case-control studies. *BMC Med Inform Decis Mak.* (2011);**11**(45):1–9.
70. Van Calster B., McLernon D.J., Van Smeden M., Wynants L., Steyerberg E.W., Bossuyt P., et al. Calibration: The Achilles heel of predictive analytics. *BMC Med.* (2019);**17**(1):1–7.
71. Gotay C.C., Kawamoto C.T., Bottomley A., Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J Clin Oncol.* (2008);**26**(8):1355–63.
72. Mol L., Ottevanger P.B., Koopman M., Punt C.J.A. The prognostic value of WHO performance status in relation to quality of life in advanced colorectal cancer patients. *Eur J Cancer.* (2016);**66**:138–43.
73. Boyle J.M., Hegarty G., Frampton C., Harvey-jones E., Dodkins J., Beyer K., et al. Real-world outcomes associated with new cancer medicines approved by the Food and Drug Administration and European Medicines Agency: A retrospective cohort study. *Eur J Cancer.* (2021);**155**:136–44.
74. Templeton A.J., Booth C.M., Tannock I.F. Informing Patients About Expected Outcomes: The Efficacy-Effectiveness Gap. *J Clin Oncol.* (2020);**38**(15):1651–4.
75. Derksen J.W.G., Vink G.R., Elferink M.A.G., Roodhart J.M.L., Verkooijen H.M., van Grevenstein W.M.U., et al. The Prospective Dutch Colorectal Cancer (PLCRC) cohort: real-world data facilitating research and clinical care. *Sci Rep.* (2021);**11**(1):1–12.



**SUMMARY IN DUTCH  
(NEDERLANDSE SAMENVATTING)**

**APPENDICES**

## **ACHTERGROND GEPERSONALISEERDE BEHANDELING BIJ UITGEZAAIDE DIKKEDARMKANKER**

In Nederland krijgen jaarlijks 13.000 mensen de diagnose dikkedarmkanker. Van deze patiënten heeft ongeveer 20-25% op het moment van diagnose reeds uitzaaiingen elders in het lichaam. Daarnaast ontwikkelt ongeveer 20% op een later moment uitzaaiingen. Patiënten met uitgezaaide dikkedarmkanker kunnen behandeld worden met systemische of lokale behandelingen. Voorbeelden van systemische behandelingen zijn chemotherapie, doelgerichte behandeling of immunotherapie. Doelgerichte behandeling grijpt aan op een kenmerkende eigenschap van een kankercel, waardoor het effect meer gericht is op specifiek de kankercellen. Voorbeelden van lokale behandelingen zijn het operatief verwijderen of bestralen van uitzaaiingen. In grote lijnen zijn de behandelopties voor patiënten met uitgezaaide dikkedarmkanker opgedeeld in twee groepen, afhankelijk van de uitgebreidheid van ziekte. Voor patiënten waarbij uitzaaiingen (operatief) verwijderd kunnen worden, heeft de behandeling als doel om te genezen. Bij patiënten waarbij de ziekte niet meer verwijderd kan worden, is de behandeling gericht op het verbeteren van de levensduur en levenskwaliteit.

Idealiter zou een oncoloog vooraf kunnen voorspellen welk behandeltype het meest effectief zal zijn voor een patiënt om hierdoor de kwaliteit van leven en overleving te verbeteren. Het toespitsen van een behandeling op een specifieke patiënt heet ‘personalized treatment’ (gepersonaliseerde behandeling). Voor veel patiënten met uitgezaaide dikkedarmkanker is gepersonaliseerde behandeling vooralsnog niet haalbaar, omdat er geen goede voorspellers zijn voor effect van de verschillende behandeltypes. Het gevolg is dat veel patiënten ineffectieve behandeling krijgen met potentieel schadelijke bijwerkingen en hoge zorgkosten zonder dat zij daar baat bij hebben gehad. Dit proefschrift richt zich op het verbeteren van de behandeling van patiënten met uitgezaaide dikkedarmkanker door middel van gepersonaliseerde behandeling.

## DEEL I – OPTIMALISEREN VAN DE BEHANDELING VOOR PATIËNTEN MET MISMATCH-REPARATIEDEFICIËNTE (dMMR) DIKKEDARMKANKER

De eerste methode voor gepersonaliseerde behandeling die in het proefschrift geanalyseerd wordt, bestaat uit het gebruik van specifieke subtypen die effect van behandeling kunnen beïnvloeden. Dikkedarmkanker kent veel subtypen, die ontstaan door verscheidene complexe cascades van moleculaire veranderingen. Bepaalde moleculaire veranderingen beïnvloeden het ziekteverloop en de gevoeligheid voor behandeling. Deze tumoreigenschappen kunnen fungeren als voorspellende biomarkers om patiënten te identificeren die meer kans hebben om al dan niet te reageren op een bepaalde behandeling en dus gepersonaliseerde behandeling te sturen.

Een belangrijk subtype van dikkedarmkanker wordt gevormd door mismatch-reparatiedeficiënte (dMMR) tumoren. Dit subtype wordt in 5% van patiënten met uitgezaaide dikkedarmkanker vastgesteld en wordt gekenmerkt door een niet functioneel DNA-herstelmechanisme, waardoor er veel mutaties ontstaan in de tumor. De andere 95% van de patiënten heeft een mismatch-reparatieproficiënte (pMMR) tumor. **Deel I** van dit proefschrift richt zich op hoe chemotherapie en doelgerichte behandeling voor patiënten met uitgezaaide dikkedarmkanker van het dMMR subtype geoptimaliseerd kan worden.

Patiënten met het dMMR subtype hebben in vergelijking met het pMMR subtype een goede prognose in vroeg stadium dikkedarmkanker, echter een slechtere overleving wanneer de dikkedarmkanker eenmaal is uitgezaaid. In een studie uit 2017 bleek dat uitgezaaide dMMR dikkedarmkanker zeer gevoelig is voor immunotherapie, een relatief nieuwe soort systemische behandeling die het afweersysteem van een patiënt activeert om de kanker te bestrijden. Er is in literatuur weinig data beschikbaar over of immunotherapie overleving van patiënten met uitgezaaide dMMR dikkedarmkanker verbetert vergeleken met reguliere behandeling (chemotherapie en doelgerichte therapie, zonder immunotherapie).

Het doel van **hoofdstuk 1** is om inzicht te krijgen in de overleving van patiënten met uitgezaaide dMMR dikkedarmkanker gedurende systemische behandeling met chemotherapie en doelgerichte behandeling die in deze tijd de standaard was. In 281 Nederlandse patiënten met uitgezaaide dMMR dikkedarmkanker werd de

overleving geanalyseerd vanaf diagnose van uitgezaaide dikkedarmkanker en vanaf start van de behandeling. Deze groep patiënten heeft in deze tijdsperiode nog geen immuuntherapie gekregen maar chemotherapie en eventuele doelgerichte therapie. De mediane algehele overleving vanaf de diagnose uitgezaaide dikkedarmkanker was 16,0 maanden voor patiënten die anti-tumor behandeling kregen, vergeleken met 2,5 maanden voor patiënten die geen anti-tumor behandeling kregen. De overleving van de groep patiënten met dMMR dikkedarmkanker die wel anti-tumor behandeling kregen was bijna 8 maanden korter dan de groep met de veel meer voorkomende pMMR uitgezaaide dikkedarmkanker. Voor dMMR patiënten was de overleving 12,8 maanden vanaf de start van eerstelijnsbehandeling en 6,2 maanden vanaf de start van tweedelijns behandeling. Van de patiënten die tweedelijnsbehandeling kregen was slechts 17.2% een jaar na start van de tweedelijnsbehandeling nog in leven, wat veel lager is dan in de recente studies met immunotherapie gezien is waarbij na tweedelijnsbehandeling na een jaar tot wel 85% van de patiënten nog in leven is.. Concluderend, de prognose van uitgezaaide dMMR dikkedarmkankerpatiënten in het hier gepresenteerde cohort is somber na behandeling met chemotherapie en doelgerichte therapie. Dit steunt de zeer goede resultaten die gezien worden in de studies met immuuntherapie.

Chemotherapie en doelgerichte therapie blijven een belangrijk onderdeel van de behandeling voor patiënten met uitgezaaide dMMR dikkedarmkanker, bijvoorbeeld wanneer immunotherapie niet meer effectief blijkt. In **hoofdstuk 2** wordt aan de hand van een analyse van 11 beschikbare studies onderzocht of er bewijs is dat patiënten met uitgezaaide dMMR dikkedarmkanker beter reageren op een bepaald type chemotherapie of doelgerichte behandeling. Gezien het geringe aantal studies en het beperkte aantal patiënten met uitgezaaide dMMR dikkedarmkanker in deze studies, kunnen geen harde conclusies worden getrokken over welke chemotherapie of doelgerichte behandeling meer effectief is voor deze patiëntengroep. Er waren drie trends in de resultaten van de studies. Ten eerste, patiënten met uitgezaaide dMMR dikkedarmkanker lijken geen betere uitkomsten te hebben als ze twee soorten chemotherapie krijgen (5-fluorouracil + irinotecan) ten opzichte van enkele chemotherapie (5-fluorouracil). Ten tweede, oxaliplatin-bevattende chemotherapie lijkt meer effectief in deze patiënten vergeleken met andere typen chemotherapie. Ten derde, de toevoeging van een bepaald type doelgerichte behandeling (bevacizumab) aan de behandeling is meer effectief dan het toevoegen van een ander type doelgerichte behandeling (cetuximab). Onderzoek naar de effectiviteit van chemotherapie en

doelgerichte behandeling bij patiënten met uitgezaaide dMMR dikkedarmkanker zou bespoedigd kunnen worden door grootschalige meta-analyses van eenduidig gerapporteerde patiëntdata. Tevens zou het vaker testen en registreren van dMMR status helpen om een groot aantal patiënten te kunnen meenemen in de analyse en dus een beter beeld te geven van de algehele populatie.

Naast klinische studies, zouden preklinische studies extra inzicht kunnen leveren in welk type behandeling effectiever is bij patiënten met dMMR dikkedarmkanker. Echter, de preklinische studies geven geen duidelijkheid over welke behandeling meer effectief is en, staan zelfs in contrast met de beschreven trends in de klinische studies. Een oorzaak hiervan kan zijn dat de preklinische studies zijn verricht met traditionele modellen die de tumoren van de patiënt niet representatief weergeven. Meer geavanceerde modellen in het laboratorium, zoals organoïden, zouden meer inzicht kunnen bieden. Tumor organoïden zijn driedimensionale mini-tumoren die worden gemaakt uit tumormateriaal van patiënten en kunnen worden gebruikt voor preklinisch onderzoek. Organoïden geven beter de eigenschappen van de tumor van een patiënt weer, omdat ze in tegenstelling tot traditionele modellen de driedimensionale structuur en heterogeniteit in celtypes behouden en langdurig stabiel zijn.

In **hoofdstuk 3** wordt gekeken of organoïden van patiënten met dMMR dikkedarmkanker intrinsiek minder gevoelig zijn voor chemotherapie en doelgerichte behandeling vergeleken met organoïden van patiënten met pMMR dikkedarmkanker. De effectiviteit van drie types chemotherapie (5-fluorouracil, oxaliplatin, irinotecan) en drie types doelgerichte behandelingen (binimetinib, encorafenib en cetuximab) is onderzocht in 12 organoïden (6 dMMR en 6 pMMR). Deze vergelijking toont geen significant verschil in gevoeligheid voor de geteste chemotherapieën en doelgerichte behandelingen in dMMR versus pMMR dikkedarmkanker organoïden. Aangezien organoïden de kankercel en diens intrinsieke gevoeligheid modelleren, suggereren deze resultaten dat de slechte prognose (overleving) van patiënten met uitgezaaide dMMR dikkedarmkanker niet verklaard kan worden door een intrinsieke verminderde kankercel gevoeligheid voor deze behandelingen. Mogelijk dat het verschil in effectiviteit van deze behandelingen bij patiënten met en zonder dMMR kan worden verklaard door verschillen in de omgeving van de tumoren, zoals het aantal en het type afweercellen. Daarnaast tonen organoïden met bepaalde mutaties (*BRAF* en *KRAS* mutatie) meer resistentie voor 5-fluorouracil en oxaliplatin ten opzichte van organoïden zonder mutatie. Verder waren organoïden met een *BRAF* of *KRAS* mutatie

meer gevoelig voor bepaalde doelgerichte behandelingen. Hieruit blijkt dat mutaties van een kankercel effect lijken te hebben op de gevoeligheid voor chemotherapie en doelgerichte behandeling, zowel in organoïden als in de patiënt. Ons onderzoek toont de relevantie van preklinisch studies met organoïden die representatief zijn voor een bepaald dikkedarmkanker subtype voor een beter inzicht in de gevoeligheid voor behandeling van dit subtype.

## DEEL II - ORGANOÏDEN ALS VOORSPELLENDE BIOMARKER VOOR KANKERBEHANDELING

Ondanks toenemende kennis op het gebied van gepersonaliseerde behandeling op basis van genetische eigenschappen, is het nog niet mogelijk om op basis van deze eigenschappen de effectiviteit van behandeling betrouwbaar te voorspellen. Daarnaast zijn er geen effectieve manieren om het effect te voorspellen voor veel gebruikte behandelingen, zoals chemotherapie. Het resultaat van een 'one size fits all' behandeling is dat veel patiënten worden blootgesteld aan behandelingen die niet effectief blijken, met als consequentie onnodige bijwerkingen en vermindering van de kwaliteit van leven van deze patiënten. Ook indien er wel voorafgaand aan doelgerichte behandeling bepaalde mutaties waarop deze aangrijpen worden geïdentificeerd, is de respons op behandeling laag. Er is dringend behoefte aan een betere afstemming van de behandeling voor individuele patiënten met uitgezaaide dikkedarmkanker. Een nieuw preklinisch model, de organoïde, kan worden gebruikt om de effectiviteit van behandeling te voorspellen. Hiervoor worden verschillende behandelingen in het laboratorium getest op de organoïde van een patiënt om te analyseren voor welke type behandeling de organoïde gevoelig is. Organoïden kunnen ook worden gebruikt om nieuwe behandelingen voor individuele patiënten en nieuwe behandelcombinaties voor dikkedarmkanker subtypes te identificeren. Op deze manier kunnen de organoïden gebruikt worden om te kijken welk type behandeling het meest effectief blijkt voor de individuele patiënt. **Deel II** van dit proefschrift bekijkt hoe organoïden gebruikt kunnen worden om effectiviteit van tumorgerichte behandeling te voorspellen voor het personaliseren van de behandeling.

Het gebruik van organoïden om de gevoeligheid voor behandeling te voorspellen is een nieuwe methode, waarvan vooralsnog onbekend is in hoeverre dit goed overeenkomt met de gevoeligheid van de patiënt voor diezelfde behandeling. In **hoofdstuk 4** zijn 17 studies bekeken die deze voorspellende waarde onderzoeken. De 17 studies tonen

dat de gevoeligheid van organoïden overeenkomt met de gevoeligheid van patiënten voor dezelfde behandeling met een sensitiviteit van 81% en specificiteit van 74% voor verschillende kanker- en behandeltypes. Het meeste beschikbare bewijs is gevonden voor dikkedarmkanker. Daarnaast tonen kleinere studies veelbelovende resultaten voor andere tumortypes. Voordat organoïden gebruikt kunnen worden om de behandelgevoeligheid van een patiënt te voorspellen, zouden de resultaten bevestigd moeten worden in andere studies om robuustheid en reproduceerbaarheid te tonen. Er zijn een aantal uitdagingen in het implementeren van de experimenten met organoïden om behandelkeuzes te sturen. Een eis hiervoor is dat het opkweken van organoïden en testen voor behandelgevoeligheid voor een meerderheid van de patiënten mogelijk is zonder vertraging van de behandeling. Tot slot zijn er studies nodig om aan te tonen dat organoïde-gestuurde behandeling kosteneffectief is en dat het een verbetering in de overleving en/of kwaliteit van leven van patiënten oplevert. Concluderend, organoïden zijn een veelbelovende voorspeller voor de effectiviteit van behandeling in patiënten met (dikkedarm)kanker en kunnen potentieel behandeling sturen.

Hoewel beschikbare studies aantonen dat voor patiënten met dikkedarmkanker de respons op specifieke types behandeling voorspeld kan worden met organoïden, is het vooralsnog onduidelijk of dit standhoudt voor alle behandeltypes, behandellijnen en klinische uitkomstmaten. Doordat de uitkomstmaten van organoïde respons en van patiëntrespons sterk variëren tussen studies is het niet duidelijk welke methode het beste werkt en klinisch het meest relevant is. **Hoofdstuk 5** onderzoekt of organoïden gebruikt kunnen worden om de respons op verschillende reguliere dikkedarmkanker behandelingen te voorspellen en welke uitkomstmaten van organoïde respons en patiëntrespons goed overeenkomen. Hiervoor zijn 19 organoïden afkomstig van patiënten met dikkedarmkanker in het laboratorium blootgesteld aan chemotherapie (5-fluorouracil, irinotecan en combinatie irinotecan + 5-fluorouracil) en doelgerichte behandelingen (panitumumab). Per type behandeling is op verschillende doseringen de organoïde respons (remming van groei) gemeten en gekwantificeerd met drie uitkomstmaten. Van deze uitkomstmaten kwam de oppervlakte van de dosis-responscurve het beste overeen met de respons van de patiënt (de mate van groeiafname van uitzaaiingen) gedurende dezelfde behandeling voor alle behandeltypes. Het is mogelijk om aan de hand van de organoïde respons de patiëntrespons op chemotherapie te voorspellen in 13 van de 16 patiënten met een sensitiviteit van 78% en specificiteit van 86%. De studieresultaten tonen aan dat voor de geteste behandeltypes de organoïde

respons overeenkomt met de respons van de patiënt voor dezelfde behandeling.

## **DEEL III - PREDICTIEMODELLEN VOOR PATIËNTEN MET UITGEZAAIDE DIKKEDARMKANKER**

Een derde methode om gepersonaliseerde behandeling voor patiënten met uitgezaaide dikkedarmkanker mogelijk te maken, is het gebruik van klinische predictiemodellen om de prognose van patiënten in te schatten. Dit is het onderwerp van **deel III** van dit proefschrift. Predictiemodellen kunnen helpen om een inschatting te maken van het risico op een bepaalde gebeurtenis, zoals terugkeer van ziekte na het operatief verwijderen van leveruitzaaiingen. Hoe goed een klinische predictiemodel in staat is het risico voor individuele patiënten in de algemene bevolking ('real life') te voorspellen, kan worden beoordeeld door een model dat is ontwikkeld in een bepaald patiëntencohort extern te valideren in een ander patiëntencohort. Bij externe validatie wordt gekeken of de modelinschattingen overeenkomen met de daadwerkelijke prognose van patiënten en of er goed onderscheid gemaakt kan worden tussen patiënten met een hoog versus laag risico voor een bepaalde gebeurtenis, om uiteindelijk klinische beslissingen te kunnen ondersteunen. Patiëntencohorten met 'real life' patiënten, zoals het Prospectieve Landelijk Colorectaalcarcinoom cohort (PLCRC), en nationale registraties, zoals de Nederlandse Kankerregistratie (NKR), zijn bij uitstek geschikt om de prestaties van predictiemodellen te beoordelen en om modellen te verbeteren.

**Hoofdstuk 6 en 7** beschrijven het gebruik van klinische predictiemodellen om de prognose in te schatten van patiënten met dikkedarmkanker die lokale behandeling krijgen voor leveruitzaaiingen. Voor patiënten met uitzaaiingen in de lever (en niet elders in het lichaam) kan lokale behandeling van de uitzaaiingen, door operatieve verwijdering of andere technieken, resulteren in een langdurige overleving met een 5-jaars overlevingskans tot 55%. Door geoptimaliseerde chirurgische technieken en behandelopties, komen meer patiënten in aanmerking voor lokale behandeling. Niet alle patiënten hebben daar echter baat bij. Terugkeer van ziekte na chirurgische resectie van leveruitzaaiingen komt voor bij tot 75% van de patiënten en een deel van de patiënten ervaart geen overlevingswinst op de lange termijn. Om te zorgen dat patiënten niet onnodig blootgesteld worden aan een ingrijpende operatie waar ze weinig of geen baat bij hebben, is het belangrijk om vooraf patiënten te identificeren die baat hebben bij lokale behandeling. Predictiemodellen kunnen helpen om de

prognose van patiënten in te schatten voorafgaand aan de operatie en kunnen worden gebruikt in de besluitvorming rondom een lokale behandeling.

Meerdere predictiemodellen zijn ontwikkeld om de prognose na lokale behandeling van leveruitzaaiingen te voorspellen voor dikkedarmkankerpatiënten. Twee predictiemodellen worden frequent beschreven, de Fong score en de Genetic and Morphological Evaluation (GAME) score. De Fong score bevat traditionele risicofactoren, zoals de aanwezigheid van lymfklieruitzaaiingen bij diagnose en de grootte en het aantal van leveruitzaaiingen. De GAME score bevat daarnaast meer recent ontwikkelde risicofactoren zoals de aanwezigheid van een *KRAS* mutatie. De predictiemodellen zijn ontwikkeld voor patiënten uit gespecialiseerde behandelcentra, waardoor het onduidelijk is of de modellen ook van toepassing zijn op 'real life' patiënten die in aanmerking komen voor lokale behandeling voor leveruitzaaiingen.

Externe validatie voor de Fong en GAME predictiemodellen wordt onderzocht in **hoofdstuk 6** in een cohort patiënten uit de NKR die lokale behandeling voor leveruitzaaiingen hebben ondergaan. Bij 1105 patiënten, gediagnosticeerd met naar de lever uitgezaaid dikkedarmkanker in 2015 of 2016, is de mediane ziektevrije overleving 10,1 maanden en de mediane algehele overleving 51,3 maanden na de lokale behandeling van de leveruitzaaiingen. De predictiemodellen resulteren in een goed onderscheid in de overleving van patiënten met een laag versus hoog risico. Bijvoorbeeld de algehele overleving voor patiënten met een hoog risico volgens de GAME score is 32,4 maanden vergeleken met 68,1 maanden voor patiënten met een laag risico volgens de GAME score. Echter, over het geheel genomen was het onderscheidend vermogen van de predictiemodellen zwak. In 25% van de patiënten waarbij er volgens de GAME score een hoog risico was op een slechte uitkomst, was de algehele overleving na behandeling langer dan 5 jaar. Hieruit volgt dat de Fong en GAME predictiemodellen de prognose van 'real life' patiënten die lokale behandeling van leveruitzaaiingen ondergaan suboptimaal voorspellen en niet in staat zijn patiënten te identificeren die niet in aanmerking zouden moeten komen voor lokale behandeling van leveruitzaaiingen.

Er zijn momenteel nog geen predictiemodellen die kunnen voorzien in de noodzaak om te voorspellen of een patiënt baat gaat hebben van een lokale behandeling voor leveruitzaaiingen. Daarom richt **hoofdstuk 7** zich op het ontwikkelen van een nieuw predictiemodel dat de kans voorspelt op een vroeg extrahepatisch recidief. Een vroeg

extrahepatisch recidief is een uitzaaiing die buiten de lever is gelokaliseerd en die binnen 6 maanden na de lokale behandeling van leveruitzaaiingen wordt vastgesteld. Patiënten met een vroeg extrahepatisch recidief zijn niet genezen door de lokale behandeling en komen in het algemeen niet meer in aanmerking voor een nieuwe lokale behandeling en hebben dus een slechtere overleving. Het gebruik van deze uitkomst in plaats van algemene overleving biedt als voordeel dat het een beter beeld geeft van het effect van de lokale behandeling omdat het minder beïnvloed wordt door eventuele andere behandelingen die patiënten nadien krijgen. Het ontwikkelde predictiemodel bevat naast de traditionele risicofactoren ook nieuwe risicofactoren bevat, zoals de locatie van de primaire tumor bij diagnose en de aanwezigheid van een *RAS* of *BRAF* mutatie.

Het predictiemodel voor extrahepatisch recidief is ontwikkeld in een cohort van 1097 patiënten uit de NKR die een lokale behandeling voor leveruitzaaiingen hebben gehad (identiek aan Hoofdstuk 6). Zes maanden na lokale leverbehandeling was de overleving 19,5 maanden voor patiënten met een extrahepatisch recidief vergeleken met 30,7 maanden voor patiënten met een recidief beperkt tot de lever en > 5 jaar voor patiënten zonder terugkeer van ziekte. Het predictiemodel maakt goed onderscheid tussen patiënten die wel of niet een vroeg extrahepatisch recidief krijgen. De kans op extrahepatisch recidief na 6 maanden is 32%, 20%, 15% en 6% in respectievelijk patiënten met een zeer hoog, hoog, matig en laag risico volgens het model. Ook de algehele overleving van patiënten verschilt tussen de risicogroepen op basis van het predictiemodel. De mediane overleving bij patiënten met een zeer hoog risico is 40,4 maanden, versus 46,7, 51,7 en >60 maanden bij patiënten met respectievelijk een hoog, matig en laag risico. Het predictiemodel heeft potentie om besluitvorming omtrent lokale behandelingen te informeren en zal extern gevalideerd moeten worden voordat het in de praktijk gebruikt kan worden.

Indien blijkt dat het predictiemodel ook in andere patiëntcohorten met een lokale behandeling voor leveruitzaaiingen het risico op een vroeg extrahepatisch recidief goed voorspelt, dan zijn studies nodig om te kijken of het gebruik van het predictiemodel leidt tot betere beslissingen of uitkomsten in deze patiënten. De score zou bijvoorbeeld gebruikt kunnen worden in een klinische trial om te beoordelen of het selecteren van hoog risico patiënten voor een vroeg extrahepatisch recidief voor chemotherapie en doelgerichte behandeling rondom de lokale behandeling tot een betere overleving leidt in vergelijking met alleen lokale behandeling. Dergelijke 'impact studies' zijn

nodig voordat het predictiemodel gebruikt kan worden om behandelbeslissingen te beïnvloeden.

**Hoofdstuk 8** beschrijft resultaten voor externe validatie van een predictiemodel, de Colon Life nomogram, voor het inschatten van de prognose van patiënten die behandeld worden voor refractaire ziekte. Van refractaire ziekte is sprake indien reguliere chemotherapie en doelgerichte behandeling niet effectief blijken. Deze patiënten hebben een slechte overleving en reageren vaak niet meer goed op behandeling. Trifluridine/tipiracil (FTD/TPI) is een type chemotherapie dat beschikbaar is voor patiënten met refractaire, uitgezaaide dikkedarmkanker. Niet alle patiënten hebben echter baat bij FTD/TPI en ervaren wel onnodige toxiciteit en behandelingskosten. Het Colon Life nomogram is een predictiemodel dat is ontwikkeld om de 12-weeks overlijdenskans te voorspellen voor refractaire patiënten met dikkedarmkanker. Het Colon Life nomogram kan helpen bij het identificeren van patiënten met een slechte prognose om te voorkomen dat patiënten aan FTD/TPI worden blootgesteld indien ze (desondanks) op basis van het predictiemodel binnen 12 weken zullen overlijden. Externe validatie van de Colon Life nomogram in een ‘real life’ patiëntenpopulatie is nodig voordat het gebruikt kan worden om klinische besluitvorming te ondersteunen.

Het Colon Life nomogram is gevalideerd in 150 patiënten die deelnamen aan de QUALITAS studie, een deelstudie van het PLCRC waarin de kwaliteit van leven van patiënten die FTD/TPI kregen werd onderzocht. Deze patiënten zijn tussen 2016 en 2019 behandeld met FTD/TPI in de dagelijkse praktijk. Van de 150 patiënten overleden er 25 (16,7%) binnen 12 weken na start van FTD/TPI. Het Colon Life nomogram was slecht gekalibreerd, wat een sterke overschatting veroorzaakt van de mortaliteit waargenomen in het cohort. Analyse toonde een negatief netto voordeel (schade) van het theoretisch gebruik van de Colon Life nomogram om FTD/TPI te sturen. Vanwege de miskalibratie van het Colon Life nomogram mag het niet in de klinische praktijk worden gebruikt. Wij adviseren om een studie uit te voeren met een voldoende grote steekproef om de kalibratie te actualiseren of een nieuw model te ontwikkelen om de prognose te schatten voor patiënten die een FTD/TPI behandeling krijgen.

A

## CONCLUDERENDE OPMERKINGEN

De manier waarop kanker wordt behandeld zal drastisch verbeteren wanneer klinici in staat zijn om alle patiënten een gepersonaliseerde behandeling aan te bieden. Ondanks

innovaties in behandelingen en voorspellende biomarkers is het voor de meerderheid van de uitgezaaide dikkedarmkankerpatiënten nog steeds niet mogelijk om behandeling te personaliseren. Het onderzoek in dit proefschrift belicht strategieën om gepersonaliseerde behandeling bij uitgezaaide dikkedarmkankerpatiënten mogelijk te maken: het onderzoeken van de gevoeligheid voor reguliere systemische behandeling in patiënten met uitgezaaide dMMR dikkedarmkanker, het gebruik van organoïden om de respons op chemotherapie en doelgerichte behandeling te voorspellen, en het gebruik van predictiemodellen om klinische besluitvorming te kunnen informeren.

Gelukkig heeft immunotherapie de overleving van dMMR dikkedarmkankerpatiënten enorm verbeterd. Verdere verbeteringen kunnen worden bereikt door te begrijpen waarom deze patiënten weinig baat hebben bij chemotherapie en doelgerichte behandeling en door strategieën te identificeren om dit te overkomen. Onze resultaten tonen aan dat er onvoldoende klinisch bewijs is dat een bepaald type chemotherapie of doelgerichte behandeling meer effectief is bij deze patiënten. Bovendien toonde ons onderzoek aan dat dMMR organoïden een vergelijkbare gevoeligheid hebben voor behandeling als pMMR organoïden. Dit suggereert dat de behandelresistentie niet intrinsiek is voor de kankercellen maar dat omgevingsfactoren een rol spelen en inzicht kunnen geven over hoe de behandelingsresistentie kan worden overwonnen.

Organoïden kunnen een revolutie teweegbrengen in de gepersonaliseerde behandeling door een op maat gemaakte behandelingsstrategie te bieden voor een veel groter deel van de patiënten met kanker. Wij hebben eerdere studies geëvalueerd, concrete aanbevelingen voor toekomstig onderzoek gedaan en bevestigd dat organoïden kunnen worden gebruikt om behandelrespons te voorspellen bij patiënten met uitgezaaide dikkedarmkanker. Studies zijn nodig die de haalbaarheid van het gebruik van organoïden voor het sturen van behandeling optimaliseren, voorspellende waarde voor gevoeligheid bevestigen in patiënten en het klinisch nut voor levensverlenging of verbetering van kwaliteit van leven en kosteneffectiviteit beoordelen. Er wordt veel verwacht van de resultaten van studies die het voordeel onderzoeken voor patiënten die een organoïde-begeleide behandeling krijgen en die de kosteneffectiviteit evalueren, wat zal helpen evalueren of het gebruik van organoïden om klinische beslissingen te sturen zinvol is.

Ten slotte kunnen predictiemodellen voor klinische risicoscores bijdragen aan een behandeling op maat mogelijk maken door patiënten voor behandeling te selecteren

op basis van een verwachte uitkomst. Uit onze externe validaties bleek dat bestaande predictiemodellen voor patiënten die lokale behandeling voor levermetastasen of die trifluridine/tipiracil kregen verbeterd moeten worden. Wij ontwikkelden een nieuw predictiemodel gericht op het voorspellen van een vroeg extrahepatisch recidief bij patiënten die een lokale behandeling voor levermetastasen ondergaan. Dit model dient gevalideerd te worden in andere patiënten en beoordeeld te worden in ‘impact studies’ om te beoordelen of het gebruik van de predictiemodellen als leidraad voor klinische besluitvorming waardevol is.

Voor klinici en onderzoekers is dit een opwindende tijd om nieuwe onderzoeksprojecten te starten om de klinische impact te bevestigen van het gebruik van deze strategieën voor het personaliseren van de behandeling voor patiënten met uitgezaaide dikkedarmkanker, wat ons uiteindelijk een stap dichterbij het kunnen bieden van een gepersonaliseerde behandeling voor alle patiënten met uitgezaaide dikkedarmkanker.



## **AUTHORS & AFFILIATIONS**

## **APPENDICES**

**S.D. (Sandra) Bakker**

Department of Medical Oncology, Zaans Medisch Centrum, Zaandam, the Netherlands

**S.F. (Sylvia) Boj**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**K. (Karen) Bolhuis**

Department of Medical Oncology, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, the Netherlands

**M.J.G. (Marinde) Bond**

Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**I.H.M. (Inne) Borel Rinkes**

Department of Surgery, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**M. (Manon) Braat**

Department of Cancer and Imaging, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**A. (Anneta) Brousali**

Utrecht Platform for Organoid Technology (UPORT), University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**G. (Geert) Cirkel**

Department of Medical Oncology, Meander Medical Center, Amersfoort, the Netherlands

**G.-J. (Geert-Jan) Creemers**

Department of Medical Oncology, Catharina Ziekenhuis, Eindhoven, the Netherlands

**J.W.B. (Jan Willem) de Groot**

Department of Medical Oncology, Isala Ziekenhuis, Zwolle, the Netherlands

**G.J. (Gerty) de Klerk**

Department of Medical Oncology, Spaarne Gasthuis, Hoofddorp, the Netherlands

**W.P.M. (Willemieke) Dijksterhuis**

Department of Medical Oncology, Amsterdam University Medical Center, University of Amsterdam, Amsterdam;  
Department of Research and Development, Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, the Netherlands

**M. (Mayke) Doorn**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**M.A.G. (Marloes) Elferink**

Department of Research and Development, Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, the Netherlands

**S.G. (Sjoerd) Elias**

Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht;  
Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**R.J.A. (Remond) Fijneman**

Department of Pathology, The Netherlands Cancer Institute, Amsterdam, the Netherlands

**B.C.M. (Brigitte) Haberkorn**

Department of Medical Oncology, Maasstad Ziekenhuis, Rotterdam, the Netherlands

**J. (Jeroen) Hagendoorn**

Department of Surgery, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**P.A.H. (Patricia) Hamers**

Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**A.W. (Annebeth) Haringhuizen**

Department of Medical Oncology, Ziekenhuis Gelderse Vallei, Ede, the Netherlands

**C. (Celia) Higuera Barón**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**R. (Ronald) Hoekstra**

Department of Medical Oncology, Ziekenhuisgroep Twente, Almelo, the Netherlands

**M. (Maarten) Huismans**

Department of Molecular Cancer Research, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**J.C.B. (Cornelis) Hunting**

Department of Medical Oncology, St. Antonius Ziekenhuis, Nieuwegein, the Netherlands

**E.D. (Emile) Kerver**

Department of Medical Oncology, Onze Lieve Vrouwe Gasthuis, Amsterdam, the Netherlands

**E. (Esmee) Koedoot**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**M. (Miriam) Koopman**

Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**R. (Ricardo) Korporaal**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**O. (Onno) Kranenburg**

Department of Surgery, University Medical Center Utrecht, Utrecht University, Utrecht; Utrecht Platform for Organoid Technology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**E. (Emre) Küçükköse**

Laboratory Translational Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**M.M. (Miangela) Laclé**

Department of Pathology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**J. (Jamila) Laoukili**

Laboratory Translational Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**R.A. (Renee) Lunenberg**

Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**D. (Danielle) Mathijssen-van Stein**

Department of Medical Oncology, Franciscus Gasthuis & Vlietland, Schiedam, the Netherlands

**A.M. (Anne) May**

Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**L. (Linda) Mol**

Department of Research and Development, Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, the Netherlands

**J. (Jasper) Mullenders**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**R. (Roshni) Nair**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**R. (René) Overmeer**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**K.-C. (Katerina-Chitsa) Pitsa**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**M.B. (Marco) Polée**

Department of Medical Oncology, Medisch Centrum Leeuwarden, Leeuwarden, the Netherlands

**J.F.M. (Johannes) Pruijt**

Department of Medical Oncology, Jeroen Bosch Ziekenhuis, 's-Hertogenbosch, the Netherlands

**C.J.A. (Cornelis) Punt**

Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**D.A.E. (Danielle) Raats**

Laboratory Translational Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**S. (Sandra) Radema**

Department of Medical Oncology, Radboudumc, Nijmegen, the Netherlands

**R.C. (Ronald) Rietbroek**

Department of Medical Oncology, Rode Kruis Ziekenhuis, Beverwijk, the Netherlands

**C.M. (Celine) Roelse**

Laboratory Translational Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**J.M.L. (Jeanine) Roodhart**

Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**J. (Jorieke) Salij**

Utrecht Platform for Organoid Technology (UPORT), University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**L.H.J. (Lieke) Simkens**

Department of Medical Oncology, Maxima Medisch Centrum, Eindhoven, the Netherlands

**L.P. (Lidwien) Smabers**

Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**R.-J. (Rutger-Jan) Swijnenburg**

Department of Surgery, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, the Netherlands

**B.C. (Bea) Tanis**

Department of Medical Oncology, Groene Hart Ziekenhuis, Gouda, the Netherlands

**E. (Emma) Teal**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**D. (Daan) ten Bokkel Huinink**

Department of Medical Oncology, Diaconessenhuis, Utrecht, the Netherlands

**M.L.R. (Manuel) Tjin-A-Ton**

Department of Medical Oncology, Ziekenhuis Rivierenland, Tiel, the Netherlands

**C.S. (Cathrien) Tromp-van Driel**

Department of Medical Oncology, Gelre Ziekenhuis, Apeldoorn, the Netherlands

**M.M. (Monique) Troost**

Department of Medical Oncology, Bravis Ziekenhuis Bergen op Zoom, Roosendaal, the Netherlands

**L. (Liselot) Valkenburg**

Department of Medical Oncology, Maastricht University Medical Center, Maastricht, the Netherlands

**A.J. (Agnes) van de Wouw**

Department of Medical Oncology, VieCuri Medical Centrum, Venlo, the Netherlands

**F.W.P.J. (Franchette) van den Berkmortel**

Department of Medical Oncology, Zuyderland Medisch Centrum Heerlen, Sittard-Geleen, the Netherlands

**A.J.M. (Anke) van der Pas**

Department of Medical Oncology, LangeLand Ziekenhuis, Zoetermeer, the Netherlands

**A.M.T. (Ankie) van der Velden**

Department of Medical Oncology, Tergooi, Hilversum, the Netherlands

**M.A. (Marjan) van Dijk**

Department of Medical Oncology, ZorgSaam Ziekenhuis, Terneuzen, the Netherlands

**J.M. (Joyce) van Dodewaard-de Jong**

Department of Medical Oncology, Meander Medisch Centrum, Amersfoort, the Netherlands

**E.B. (Edith) van Druten**

Department of Medical Oncology, Reinier de Graaf Gasthuis, Delft, the Netherlands

**E. (Edwin) van Oosten**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**S.J. (Susanne) van Schelven**

Laboratory Translational Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**M. (Maarten) van Smeden**

Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

**P. (Patricia) van Ufford-Mannesse**

Department of Medical Oncology, HagaZiekenhuis, den Haag, the Netherlands

**T. (Theo) van Voorthuizen**

Department of Medical Oncology, Ziekenhuis Rijnstate, Arnhem, the Netherlands

**I. (Inez) van Weersch**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**G.J. (Gerrit Jan) Veldhuis**

Department of Medical Oncology, Antonius Ziekenhuis Sneek, Sneek, the Netherlands

**H.M.W. (Henk) Verheul**

Department of Medical Oncology, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, the Netherlands

**C. (Carla) Verissimo**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**J.H.M.J. (Hanneke) Vestjens**

Department of Medical Oncology, VieCuri Medical Centrum, Venlo, the Netherlands

**J. (Jeroen) Vincent**

Department of Medical Oncology, Elkerliek Ziekenhuis, Helmond, the Netherlands

**G.R. (Geraldine) Vink**

Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht;  
Department of Research and Development, Netherlands Comprehensive Cancer Organisation  
(IKNL), Utrecht, the Netherlands

**R.G.J. (Robert) Vries**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands

**H. (Hamdy) Warda**

Foundation Hubrecht Organoid Technology (HUB), Utrecht, the Netherlands



REVIEW COMMITTEE

APPENDICES

**Dr. M. Chalabi**

Medical Oncologist at the Department of Gastrointestinal Oncology, Antoni van Leeuwenhoek Ziekenhuis, Amsterdam, the Netherlands.

**Prof. dr. C.H. van Gils**

Professor of Clinical Epidemiology of Cancer at the Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands.

**Prof. dr. G.J.P.L. Kops**

Professor of Molecular Tumor Cell Biology at the Department of Molecular Cancer Research and the Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands.

**Prof. dr. M.J.L. Ligtenberg**

Professor of Molecular Tumor Genetics at the Faculty of Medical Sciences, Radboudumc, Radboud University, Nijmegen, the Netherlands.

**Prof. dr. P.O. Witteveen**

Professor of Medical Oncology at the Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands.





## LIST OF PUBLICATIONS

## APPENDICES

K. Bolhuis\*, **G.E. Wensink\***, M.A.G. Elferink, M.J.G. Bond, W.P.M. Dijksterhuis, R.J.A. Fijneman, O.W. Kranenburg, I.H.M. Borel Rinkes, M. Koopman, R.-J. Swijnenburg, G.R. Vink, J. Hagendoorn, C.J.A. Punt, S.G. Elias<sup>^</sup>, J.M.L. Roodhart<sup>^</sup>. External Validation of Two Established Clinical Risk Scores Predicting Outcome after Local Treatment of Colorectal Liver Metastases in a Nationwide Cohort. *Cancers*, 14(10), 2356 (2022). \*Shared first author. <sup>^</sup>Shared last author.

E. Küçükköse\*, **G.E. Wensink\***, C.M. Roelse, S.J. van Schelven, D.A. Raats, S.F. Boj, M. Koopman, J. Laoukili, J.M.L. Roodhart<sup>^</sup>, O. Kranenburg<sup>^</sup>. Mismatch Repair Status in Patient-Derived Colorectal Cancer Organoids Does Not Affect Intrinsic Tumor Cell Sensitivity to Systemic Therapy. *Cancers*, 13(21), 5434 (2021). \*Shared first author. <sup>^</sup>Shared last author.

**G.E. Wensink**, S.G. Elias, J. Mullenders, M. Koopman, S.F. Boj, O.W. Kranenburg, J.M.L. Roodhart. Patient-derived organoids as a predictive biomarker for treatment response in cancer patients. *npj Precis. Onc.* 5, 30 (2021).

**G.E. Wensink**, M.J.G. Bond, E. Küçükköse , M. Koopman, O.W. Kranenburg<sup>^</sup>, J.M.L. Roodhart<sup>^</sup>. A review of the sensitivity of metastatic colorectal cancer patients with deficient mismatch repair to standard-of-care chemotherapy and monoclonal antibodies, with recommendations for future research. *Cancer Treatment Reviews* 95, 102174 (2021). <sup>^</sup>Shared last author.

**G.E. Wensink**, M.A.G. Elferink, A.M. May, L. Mol, P.A.H. Hamers, S.D. Bakker, G.-J. Creemers, J.W.B. de Groot, G.J. de Klerk, B.C.M. Haberkorn, A.W. Haringhuizen, R. Hoekstra, J.C.B. Hunting, E.D. Kerver, D. Mathijssen-van Stein, M.B. Polée, J.F.M. Pruijt, P. Quarles van Ufford-Mannesse, S. Radema, R.C. Rietbroek, L.H.J. Simkens, B.C. Tanis, D. ten Bokkel Huinink, M.L.R. Tjin-A-Ton, C.S. Tromp-van Driel, M.M. Troost, A.J. van de Wouw, F.W.P.J. van den Berkmortel, A.J.M. van der Pas, A.M.T. van der Velden, M.A. van Dijk, J.M. van Dodewaard-de Jong, E.B. van Druten, T. van Voorthuizen, G.J. Veldhuis, H.M.W. Verheul, J.H.M.J. Vestjens, J. Vincent, O.W. Kranenburg, C.J.A. Punt, G.R. Vink, J.M.L. Roodhart, M. Koopman. Survival of patients with deficient mismatch repair metastatic colorectal cancer in the pre-immunotherapy era. *Br J Cancer* 124, 399–406 (2021).

**G.E. Wensink**, A.F. Schoffelen, H.A. Tempelman, M.B. Rookmaaker, A.I.M. Hoepelman, R.E. Barth. Albuminuria is associated with traditional cardiovascular

risk factors and viral load in HIV-infected patients in rural South Africa. *PLoS ONE*. 10(8):e0136529 (2015).

C.H.P. Cremers, I.C. van der Schaaf, **G.E. Wensink**, J. Greving, G.J.E. Rinkel, B. K. Velthuis, M.D.I. Vergouwen. CT-perfusion and delayed cerebral ischemia in aneurysmal subarachnoid hemorrhage: a systematic review and meta-analysis. *The Journal of Cerebral Blood Flow & Metabolism* 34(2):200-207 (2014).

### *Oral presentation*

Abstract 4467: **G.E. Wensink**, M.A.G. Elferink, A.M. May, L. Mol, P.A.H. Hamers, C.J.A. Punt, G.R. Vink, J.M.L. Roodhart, M. Koopman. Survival of patients with microsatellite instable (MSI) metastatic colorectal cancer (mCRC) upon systemic non-immunotherapy. *AACR Annual Meeting*. March 29-April 3, 2019; Atlanta, GA. Published July 2019.



# **ACKNOWLEDGEMENTS**

## **(DANKWOORD)**

# **APPENDICES**

Met veel trots rond ik mijn promotie af. Graag deel ik dit moment met degenen die geholpen hebben om dit te realiseren. Als iemand me vooraf had verteld dat mijn promotie gedurende een pandemie zou plaatsvinden, waarin onderzoek maandenlang stil gelegd zou worden en bezoek aan familie en vrienden bemoeilijkt zou zijn, dan was ik waarschijnlijk gillend weggerend. Het spreekt dus boekdelen over de hulp en steun die ik heb gekregen dat ik desondanks dit alles met een glimlach terugkijk op mijn onderzoeksperiode. Dit proefschrift is het resultaat van de samenwerking en steun van velen. Ik ben hier enorm dankbaar voor! Graag zou ik een aantal personen in het bijzonder bedanken.

Allereerst wil ik de **patiënten en medische specialisten** bedanken die direct of indirect deel hebben genomen aan onderzoek dat dit proefschrift heeft mogelijk gemaakt. Jullie bereidwilligheid om deel te nemen aan onderzoek betekent zo veel voor toekomstige patiënten en onderzoekers. Dank hiervoor!

De ontwikkelingen die ik op persoonlijk en wetenschappelijk vlak tijdens mijn promotietraject heb doorgemaakt zijn mede te danken aan de begeleiding van mijn (co)promotoren.

**Prof. dr. M. Koopman**, beste Miriam, dank voor deze mooie kans waarin je vanaf het begin vertrouwen hebt gehad. Het is knap om te zien hoe jij naast de kliniek zo veel onderzoek, samenwerkingen en projecten hebt lopen om de oncologische zorg te verbeteren. Jouw visie, kennis en enthousiasme zijn inspirerend voor mij.

**Prof. dr. O. Kranenburg**, beste Onno, vanaf het begin voelde ik me welkom in het ‘Laboratory of Translational Oncology’. Dank voor de kennis en overlegmomenten en dat je mij toegang hebt gegeven tot de translationele onderzoekswereld.

**Dr. J.M.L. Roodhart**, beste Jeanine, je bent de oprichter, drijfveer en aanmoedigster geweest gedurende mijn promotie. Dank voor je oneindige begeleiding, enthousiasme, kennis en vertrouwen. In vier jaar heb je een eigen onderzoeksrichting opgezet, waar meerdere promovendi nu deel van uitmaken. Ik vond het een eer om jouw eerste ‘eigen’ promovenda te zijn!

Alle coauteurs, dank voor jullie samenwerking en bijdragen aan de manuscripten. In het bijzonder wil ik **prof. dr. Anne May**, **dr. Geraldine Vink**, **prof. dr. Kees Punt** en **dr. Sjoerd Elias** bedanken. **Anne**, dank voor jouw kennis en inspanningen voor meerdere projecten. **Geraldine**, dank voor jouw kritische blik en inbreng. Het is knap

om te zien hoe jij samen met Miriam PLCRC door Nederland uitbreidt en ontwikkelt. Hierdoor waren veel onderzoeksvragen haalbaar, dank hiervoor! **Kees**, ik heb jouw kritische blik en toevoegingen aan de manuscripten zeer gewaardeerd. **Sjoerd**, ik wil je enorm bedanken voor het delen van jouw kennis. Het was maar een tip van de sluier, maar ik heb er veel van geleerd. Ondanks jouw overvolle agenda maakte je tijd vrij om de (voor mij) ingewikkelde onderwerpen uit te leggen en mij op weg te helpen.

Leden van de beoordelings- en verdedigingscommissie, **Dr. M. Chalabi**, **Prof. dr. C.H. van Gils**, **Prof. dr. G.J.P.L. Kops**, **Prof. dr. M.J.L. Ligtenberg**, **Dr. H.J.G. Snippert** en **Prof. dr. P.O. Witteveen**, dank voor uw tijd en het kritisch lezen van mijn proefschrift. Beste **Els**, dankzij jouw aanbeveling is dit promotietraject mogelijk geworden. Dank hiervoor!

Graag zou ik de **Nederlandse Kankerregistratie (NKR)** en in het bijzonder **Marloes Elferink** en **Marianne van der Mark** willen bedanken voor hun inzet, kennis en bijdrage aan de vele NKR-gebaseerde onderzoeksprojecten in dit proefschrift.

Alle leden van het OPTIC team, de lokale hoofdonderzoekers (**dr. Geert Cirkel**, **dr. Frank Jeurissen**, **dr. Liselot Valkenburg** en **dr. Henk Verheul**) en de onderzoeksmedewerkers (**Els**, **Karlijn**, **Noortje**, **Preeti**, **Nicol**, **Rutger** en **Mirte**), bedankt voor jullie voortdurende inzet! Ik heb ervan genoten om de OPTIC studie samen met jullie te mogen uitvoeren. Ondanks de toegenomen drukte door corona waren jullie onvermoeibaar in jullie inzet.

Beste **Roelien** en **Ellis** van het trialbureau bij de afdeling medische oncologie, dank voor jullie geduld, kennis en steun. Jullie hebben mij wegwijs gemaakt in de wereld van klinische studies, METC en amendementen! Beste **oncologie studie verpleegkundigen**, dank voor jullie inzet en hulp bij de vele OPTIC studie inclusies.

Stafleden en fellows van de afdeling medische oncologie van het UMC Utrecht, dank voor jullie betrokkenheid bij de OPTIC studie.

Dear Hubrecht Organoid Technology colleagues, including **Rob**, **Sylvia**, **Carla**, **Jasper**, **Esmee**, **Renee**, **Begoña**, **Emma**, **Celia**, **Katerina** and **Tiny**, thank you for welcoming me into the fascinating world of organoids and drug screening. A big thanks for your knowledge, hard work, vision and enthusiasm during the OPTIC study. I am proud to see what we have established in the past years.

Beste **Karen Bolhuis**, dank voor de fijne samenwerking aan de CRLM projecten. Het is leuk om te zien hoe jij jouw carrière voortzet in het Antoni van Leeuwenhoek. We zullen elkaar vast vaker tegenkomen.

To all students who helped out with research projects, including **Itziar Aranguren Abeigón, Renate Bezemer, Marinde Bond, Renee Lunenberg, Ayla Pieterse, Celine Roelse** and **Lidwien Smabers**, thank you for your enthusiasm, dedication and hard work. Lidwien, het is mooi om te zien hoe jij nu vol enthousiasme jouw PhD begonnen bent. De OPTIC studie is in goede handen en belooft mooi te worden!

Gedurende mijn promotie had ik het geluk om onderdeel te mogen zijn van twee onderzoeksgroepen – het ‘Laboratory of Translational Oncology’ en de ‘Koopman groep’. Ik wil mijn collega’s bedanken voor alle gezellige momenten samen. **Alec, André, Annet, Bas, Daan, Danielle, Emma, Emre, Esther, Inge, Jamila, Kyah, Layla, Liza, Lotte, Nicola, Niek, Paulien, Sinisha, Susanne, Suzanna & Ahmed, Dave, Frederieke, Guus, Ingrid, Jeroen, Karel, Karlijn, Koen, Lidwien, Maarten, Marinde, Mira, Patricia, Sietske, Suzanna**; enorm bedankt! Ik heb genoten van de koffie momenten, pannenkoek lunches, hardloop sessies, vele taarten en uiteraard borrels. Lidwien, Maarten en Emre, mede dankzij jullie is de OPTIC studie een succes!. Liza, Lotte and Suzanna, I’ve enjoyed having you as roomies! Patricia, we begonnen als borrel buddies bij de interne opleiding. Dank voor de gezellige momenten en samenwerking!

Lieve vrienden, dank voor alle gezellige momenten samen. Lieve **Heleen, Ruben, Nina** en **Annemarijn**, we zijn al ruim 12 jaar vrienden. Dank voor alle steun, gezelligheid en dat jullie onderdeel zijn van mijn leven. De wandel en eet/gebak momentjes zijn me dierbaar! Lieve **Judith, Sanne, Tom** en **Vince**, onze fietsritjes, etentjes en vakanties blijven een favoriet. Dank voor de gezelligheid en steun de afgelopen jaren – dat we veel mooie momenten samen mogen beleven. Ik ben extra dankbaar dat jullie mij uit de wind houden zodat ik de illusie heb enigszins te kunnen wielrennen.

Lieve **moeder** en **vader**, dankzij jullie heb ik de vrijheid en de kansen gehad om mijn hart te volgen. Dank dat jullie altijd voor ons klaar staan. Jullie lieten je zelfs niet door corona ontmoedigen, en kwamen ons gedurende de kraamtijd bezoeken om onze vriezer vol te proppen met verse maaltijden. Lieve **Cox, Chris** en **Frits**, ik geniet van de unieke momenten dat we samen zijn en elkaar weer zien. Ik kijk uit naar de vakanties in Canada om jullie gezinnen te bezoeken. Cox en Chris, ik bewonder

wat jullie bereikt hebben en ben blij dat jullie naast mij als paranimf deze dag samen vieren. Lieve schoonouders, **Elly, Evert, Lien Ay** en **Willem**, dank dat jullie mij zo warm onthaald hebben in jullie gezinnen. Ik vind het mooi om te zien hoe jullie van Quinn genieten en ben dankbaar voor de logeerpartijen en oppasmomenten.

Lieve **Onno**, de afgelopen jaren hebben we veel feestelijke momenten samen gehad – samenwonen, ons eerste eigen huis (aan het water!), trouwen en natuurlijk de komst van Quinn. Samen met jou is het leuker, gezelliger en zijn er ongetwijfeld meer grappen. Met jou is het leven een feest. Nu moet ik alleen nog overtuigend genoeg zijn dat kanoreizen en wandelvakanties mooie vakantieplannen zijn! Lieve **Quinn**, blijf met jouw lach en vrolijke blik de wereld bewonderen. Wij genieten enorm van jou!

Ik wens u heel veel leesplezier toe,

Emerens



**ABOUT THE AUTHOR**

**APPENDICES**



Emerens Wensink was born in 1986 in Lelystad and grew up with her family in Zeewolde. In 1994, she emigrated to Canada along with her parents, two sisters and brother, where her parents established a dairy farm in Innerkip, Ontario. On the farm, she enjoyed adventuring with her siblings and dogs, building forts, sledding and baking. She has fond memories of the family's annual camping and canoeing trips in Canada.

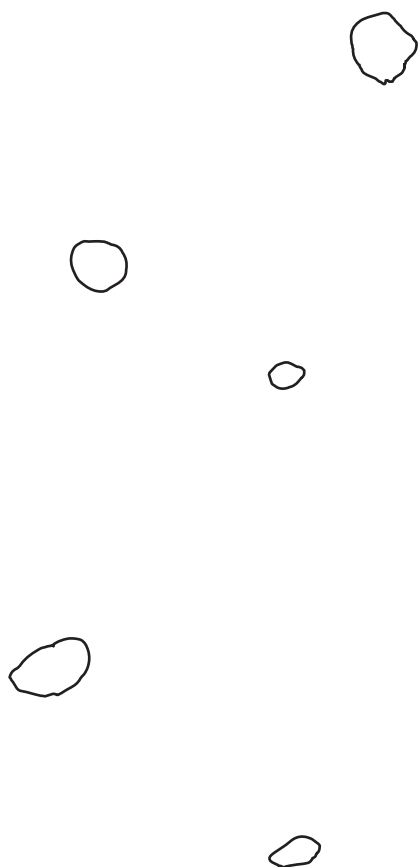
In 2008, Emerens obtained a Honours Bachelor of Science degree in Biochemistry at the Queen's University in Kingston, Ontario. Afterwards, she worked for a year as a research trainee in the lab of Dr. Ramesh Pillai at the European Molecular Biology Laboratory in Grenoble, France; spending her free time hiking and skiing in the mountains.

She moved to Utrecht in 2010, to start the Selective Utrecht Medical Master (SUMMA), a combined Master's degree for medicine and clinical researcher. As part of her Master's research project, Emerens spent 6 months at the Ndlovu Medical Centre in Elandsdoorn, South Africa. This gave her valuable insight into medical care in South Africa.

In 2014, Emerens worked as a physician for the Internal Medicine Department at the Diakonessenhuis in Utrecht. Subsequently, she started her internal medicine residency at the University Medical Center Utrecht (UMC Utrecht) in 2015, working at the Diakonessenhuis Utrecht and the UMC Utrecht. In 2018 she started her PhD research in medical oncology at the UMC Utrecht with Prof. Miriam Koopman and Prof. Onno Kranenburg as promotor and Dr. Jeanine Roodhart as co-promotor. She will continue her internal medicine residency in 2022, with the aim of becoming a medical oncologist. She hopes to continue combining clinical research while working as a medical oncologist.

Emerens currently lives in Vleuten with her husband Onno de Haan and son Quinn. In her spare time she enjoys cooking up a storm, baking, hiking and road biking with family and friends.





TOWARDS PERSONALIZED  
TREATMENT FOR PATIENTS WITH  
METASTATIC COLORECTAL CANCER  
© Emerens Wensink, 2022