



Choosing the right strategy based on individualized treatment effect predictions: combination versus sequential chemotherapy in patients with metastatic colorectal cancer

Johannes J. M. Kwakman^a, Rob C. M. van Kruijsdijk^b, Sjoerd G. Elias^c, Matthew T. Seymour^d, Angela M. Meade^e, Frank L. J. Visseren^f, Cornelis J. A. Punt^g  and Miriam Koopman^h 

^aDepartment of Medical Oncology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; ^bDepartment of Internal Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; ^cDepartment of Epidemiology, Julius Center for Health Sciences and Primary Care University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; ^dDepartment of Medical Oncology, The Leeds Teaching Hospitals, University of Leeds, Leeds, UK; ^eClinical Trials Unit at UCL, Institute of Clinical Trials and Methodology, London, UK; ^fDepartment of Vascular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; ^gDepartment of Medical Oncology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; ^hDepartment of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

ABSTRACT

Background: Translating results from randomized trials to individual patients is challenging, since treatment effects may vary due to heterogeneous prognostic characteristics. We aimed to demonstrate model development for individualized treatment effect predictions in cancer patients. We used data from two randomized trials that investigated sequential versus combination chemotherapy in unresectable metastatic colorectal cancer (mCRC) patients.

Material and methods: We used data from 803 patients included in CAIRO for prediction model development and internal validation, and data from 1423 patients included in FOCUS for external validation. A Weibull model with pre-specified patient and tumour characteristics was developed for a prediction of gain in median overall survival (OS) by upfront combination versus sequential chemotherapy. Decision curve analysis with net benefit was used. A nomogram was built using logistic regression for estimating the probability of receiving second-line treatment after the first-line monochemotherapy.

Results: Median-predicted gain in OS for the combination versus sequential chemotherapy was 2.3 months (IQR: -1.1 to 3.7 months). A predicted gain in favour of sequential chemotherapy was found in 231 patients (29%) and a predicted gain of >3 months for combination chemotherapy in 294 patients (37%). Patients with benefit from sequential chemotherapy had metachronous metastatic disease and a left-sided primary tumour. Decision curve analyses showed improvement in a net benefit for treating all patients according to prediction-based treatment compared to treating all patients with combination chemotherapy. Multiple characteristics were identified as prognostic variables which identify patients at risk of never receiving second-line treatment if treated with initial monochemotherapy. External validation showed good calibration with moderate discrimination in both models (C-index 0.66 and 0.65, respectively).

Conclusions: We successfully developed individualized prediction models including prognostic characteristics derived from randomized trials to estimate treatment effects in mCRC patients. In times where the heterogeneity of CRC becomes increasingly evident, such tools are an important step towards personalized treatment.



ARTICLE HISTORY


Received 30 September 2018
Accepted 27 December 2018

Introduction

In recent years, a better understanding of prognostic and predictive patient and tumour characteristics has significantly influenced the selection of cancer treatments for individual patients. Together with a growing number of effective and target-specific drugs, cancer treatment becomes increasingly personalized. Also, there is currently more focus on

treatment strategies rather than isolated treatment regimens. Examples in metastatic colorectal cancer (mCRC) are the use of sequential versus combination chemotherapy [1,2], and the use of maintenance treatment with a reintroduction of initial treatment upon progression [3,4]. For clinicians, it is challenging to predict the treatment effects of such strategies in an individual patient, with the availability of only the average treatment effects as observed in randomized

CONTACT Johannes J. M. Kwakman  j.j.kwakman@amc.uva.nl  Department of Medical Oncology, Amsterdam University Medical Center, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

 Supplemental data for this article can be accessed [here](#).

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

clinical trials. With a growing understanding of patient and tumour heterogeneity, the development of individualized prediction models to estimate absolute treatment effects may be an important step towards personalized treatment. Data from randomized phase III trials can be used to develop multivariable prediction models that help to identify which patients benefit from a specific treatment.

Such endeavours have been successfully undertaken in vascular medicine and lung cancer [5,6]. For this purpose, we aimed to demonstrate the development of individualized prediction models that estimate the optimal treatment strategy in patients with mCRC. As to the use of chemotherapy, doublet or triplet regimens result in higher response rates compared to monotherapy and are therefore preferred in patients with potentially resectable metastases, symptomatic disease, and/or aggressive tumours such as those harbouring a *BRAF*^{V600E}-mutation [7]. In other situations, upfront treatment with fluoropyrimidine monotherapy is considered a valid alternative. The CAIRO trial demonstrated that the strategy of sequential capecitabine, irinotecan and oxaliplatin did not compromise patients' survival or quality of life compared to upfront doublet chemotherapy [1]. This finding was confirmed in the FOCUS trial [2]. However, in the CAIRO trial, only 67% of patients in the sequential treatment arm received second-line treatment. Therefore patients who are treated with the first-line monotherapy are at risk not to complete the full strategy, and thus not to benefit from all available treatment options.

With the use of methodological frameworks [8,9], we demonstrate the development of a prediction model with patient and tumour characteristics for the individualized prediction of survival time for two treatment strategies: upfront combination versus sequential chemotherapy in patients with asymptomatic and unresectable mCRC. For patients without a clear predicted survival benefit for combination or sequential chemotherapy, we built a model to estimate the probability of receiving second-line treatment in patients exposed to upfront monotherapy in order to further guide clinical decision making. We aim to assess if individualized treatment effect predictions can assist in the realization of personalized treatment in mCRC.

Material and methods

Patients

A complete description of the methods is provided in [Supplementary Methods](#). In short, CAIRO data were used for the development of the models. In the CAIRO trial, 803 patients with mCRC not amenable for curative surgery were randomized to receive either (a) first-line treatment with capecitabine monotherapy, second-line treatment with irinotecan and third-line treatment with capecitabine plus oxaliplatin (CAPOX) or (b) first-line treatment with capecitabine plus irinotecan (CAPIRI) and second-line treatment with CAPOX. Both arms were used for the model predicting survival times. For the development of a model predicting the probability of receiving second-line treatment after the first-line monotherapy, only patients in arm A with a

complete follow-up – i.e., until death or exposure to second-line treatment – were included.

Data of the FOCUS trial were used for the external validation of the models. In FOCUS, 2135 patients were randomized between (a) first-line treatment 5-fluorouracil (5-FU) and second-line irinotecan, (b) first-line treatment 5-FU and second-line 5-FU plus irinotecan (FOLFIRI) or 5-FU plus oxaliplatin (FOLFOX), or (c) upfront combination chemotherapy with FOLFIRI or FOLFOX. For the model predicting survival times, arm A (sequential chemotherapy) and arm C (combination chemotherapy) were included. For the model predicting the probability of receiving second-line treatment arm A and B were used.

Development of model estimating overall survival times

CAIRO data were used to build an accelerated failure time model with a Weibull distribution for prediction of gain in median overall survival (OS) for individual patients (i.e., time-point from which onwards it is more likely that patient is dead than alive). Pre-specified predictors of survival included sex, WHO performance status (PS) (0, 1, or 2), body mass index (BMI), number of metastatic sites (0, 1, 2 or ≥ 3), presentation of metastatic disease (synchronous or metachronous), resection of the primary tumour (yes or no), sidedness of the primary tumour (right colon until splenic flexure, or left colon/rectum from splenic flexure on), alkaline phosphatase (ALP), and white blood cell (WBC) count [10–13]. Missing values were imputed. Treatment arm was added to the model as a predictor for survival, and WHO performance score was added as treatment interaction since previous data indicated that patients with poor performance may benefit from intensified upfront therapy [2,14]. The presence of additional treatment interactions was tested [15]. Data on *(K)RAS/BRAF*^{V600E} and serum lactate dehydrogenase (LDH) were not included due to incompleteness.

Prior to obtaining predictions in the external validation set, model coefficients were penalized in order to obtain reliable estimates by adjusting for optimism. Model performance was measured using the C-index [16], and a calibration plot was constructed to evaluate how close the predictions were to the observed survival times.

Development of model estimating the probability of receiving second-line treatment

A step-by-step protocol [9] was followed for the development of this model: (1) potential prognostic variables were identified and missing data were imputed; (2) predictors were selected using logistic regression analysis with backward stepwise selection; (3) the model was subjected to 1000 bootstrap resamples for internal validation and appraised with Harrell's C-index [17]; (4) model coefficients were shrunk after which FOCUS data were used for external validation and a nomogram was constructed.

Model outcomes

The first model was used to predict median OS upon sequential treatment and combination treatment for every

Table 1. Patient characteristics.

	CAIRO <i>n</i> = 803	Missings, no. (%)	FOCUS <i>n</i> = 1423	Missings, no. (%)
Age, median (IQR), years	63 (56–69)	0	64 (56–69)	1 (0.1)
Sex (%)				
Male	507 (63)	0	981 (69)	0
Female	296 (37)		442 (31)	
WHO performance status, no. (%)				
0	501 (62)	0	589 (41)	0
1	268 (33)		713 (50)	
2	34 (4)		121 (9)	
Body mass index, median (IQR), kg/m ²	25.0 (22.7–27.4)	39 (5)	25.1 (22.7–28.3)	7 (0.5)
Number of metastatic sites, no. (%)				
0	0	10 (1)	30 (2)	0
1	354 (45)		582 (41)	
2	287 (36)		563 (40)	
≥3	152 (19)		248 (17)	
Presentation of metastases, no. (%)				
Synchronous	517 (64)	0	897 (65)	44 (3)
Metachronous	286 (36)		482 (35)	
Resection primary tumour, no. (%)	634 (79)	0	1071 (75)	0
Sidedness of primary tumour, no. (%)				
Left	536 (74)	76 (9)	677 (73)	499 (35)
Right	191 (26)		247 (27)	
White blood cell count, median (IQR), ×10 ⁹ /L	8.0 (6.6–10.0)	145 (18)	8.2 (6.7–10.1)	11 (0.8)
Alkaline phosphatase, median (IQR), U/L	114 (86–188)	147 (18)	129 (88–231)	7 (0.5)

Data are based on unimputed values. SI conversion factor: to convert alkaline phosphatase to microkatal per liter, multiply by 0.0167.

individual patient in the CAIRO trial. The predicted gain in median OS was subsequently calculated as the difference between these two survival estimates. The second model was used to predict the probability of receiving second-line treatment after the first-line monotherapy. Since exposure to all available drugs is associated with improved survival [18], this model would be particularly helpful for patients without a clear predicted survival benefit for either combination or sequential chemotherapy.

For the model predicting gain in median OS, decision curve analysis was used to determine whether treatment decisions based on the model predictions would result in better clinical outcomes than treating patients based on group level results (treating all or none with sequential chemotherapy) [19]. This method includes calculation of net benefit. A detailed description of net benefit calculations with an example is described in [Supplementary Methods](#). Positive net benefit indicates that the treatment strategy is superior to treating all patients with sequential chemotherapy, which is the reference (net benefit equals zero), whereas negative net benefit indicates the worse clinical outcome. In both CAIRO and FOCUS, the net benefit in OS of the following treatment strategies was compared: treat all patients with sequential chemotherapy, treat all patients with upfront combination chemotherapy, prediction-based treatment, and prediction-based treatment treating only those with a predicted treatment effect with $p < .05$. As the appropriate treatment threshold is subjective, we calculated the net benefit for thresholds ranging from 0 to 6 months gain in median OS. Analyses were performed using SPSS version 24 and R version 3.3.3.

Results

Prediction model for estimation of overall survival times

The baseline characteristics of eligible study patients included for the development of the model predicting OS

times are shown in [Table 1](#). One or more variables were missing in 293 CAIRO (29.8%) and 534 FOCUS (37.5%) patients, and mainly concerned WBC and ALP. Overall, the CAIRO population included more females, had a better WHO PS and lower ALP levels compared to the FOCUS population. All other characteristics were comparable. There were no major differences in baseline characteristics between study arms [1,2]. In CAIRO, updated results with a follow-up until June 2013 (median 16.6 months, range 0.3–115.0) and 785 deaths (98%) showed a median OS of 17.2 months in the combination arm and 16.1 months in the sequential arm (hazard ratio [HR]: 0.89, 95% CI: 0.78–1.03; $p = .12$). In FOCUS (follow-up until October 2006, median 14.5 months [0.0–65.3]), median OS with 1223 deaths (86%) was 15.9 months for combination treatment and 13.9 months for sequential treatment (HR: 0.88, 95% CI: 0.78–0.98; $p = .02$).

Model coefficients accompanied with p values and unpenalized HR with corresponding 95% CIs are shown in [Table 2](#). ALP was log-transformed to optimize model fit. Primary tumour location (p for interaction = .09) and synchronous metastatic disease (p for interaction = .02) were identified and added next to WHO PS as treatment interactions.

Calibration plots of predicted versus observed median OS in the derivation set showed good internal calibration, with a slight overestimation in patients with the highest predicted probabilities ([Supplementary Figure 1](#)). More overestimation was present in the calibration plot of the external dataset ([Supplementary Figure 2](#)). The C-index in the derivation and external validation set were 0.69 (95% CI: 0.67–0.72) and 0.66 (95% CI: 0.64–0.68), respectively.

The formula for the predicted gain in median OS for combination chemotherapy versus sequential chemotherapy is shown in [Supplementary Figure 3](#). A wide range of predicted gain was observed in CAIRO, with a median of 2.3 months (IQR: –1.1 to 3.7; [Supplementary Figure 4](#)). A comparable distribution was found in FOCUS ([Supplementary Figure 5](#)).

Table 2. Model coefficients derived from CAIRO.

Predictor	AFT coefficient (95% CI) ^{a,b}	p-value	Hazard ratio (95% CI) ^b
Sex (male vs. female)	0.07 (−0.04 to 0.18)	.15	0.90 (0.77 to 1.04)
WHO performance status ^c	−0.10 (−0.24 to 0.04)	.11	1.16 (0.96 to 1.40)
Body mass index (kg/m ²) ^c	0.01 (0.00 to 0.02)	.14	0.99 (0.97 to 1.00)
Number of metastatic site ^c	−0.22 (−0.29 to −0.15)	<.01	1.39 (1.26 to 1.53)
Presentation of metastatic disease (metachronous vs. synchronous)	−0.14 (−0.31 to 0.03)	.07	1.24 (0.98 to 1.55)
Resection primary tumour (no vs. yes)	0.28 (0.14 to 0.42)	<.01	0.66 (0.54 to 0.80)
Primary tumour location (right colon vs. left colon or rectum)	0.37 (0.21 to 0.54)	<.01	0.57 (0.46 to 0.71)
White blood cell count (×10 ⁹ /L) ^c	−0.02 (−0.04 to 0.00)	.04	1.03 (1.00 to 1.07)
Logarithm of ALP (U/L) ^c	−0.28 (−0.38 to −0.18)	<.01	1.53 (1.33 to 1.75)
Sequential versus combination chemotherapy ^c	0.07 (−0.19 to 0.33)	.57	0.90 (0.63 to 1.28)
Sequential versus Combination Chemotherapy × WHO Performance Status ^{c,d}	0.06 (−0.12 to 0.24)	.50	0.92 (0.72 to 1.17)
Sequential versus Combination Chemotherapy × Presentation of Metastatic Disease (metachronous vs. synchronous)	0.24 (0.02 to 0.47)	.02	0.69 (0.51 to 0.94)
Sequential versus Combination Chemotherapy × Primary Tumour Location ^c (right colon vs. left colon or rectum)	−0.19 (−0.42 to 0.05)	.09	1.32 (0.95 to 1.83)

AFT, accelerated failure time; CI, confidence interval; ALP, alkaline phosphatase.

SI conversion factor: to convert alkaline phosphatase to microkatal per liter, multiply by 0.0167.

^aA negative AFT coefficient indicates a negative effect on predicted survival time, whereas a positive AFT coefficient indicates a positive effect.

^bUniform shrinkage was applied to the AFT coefficients, but not the hazard ratios, because penalizing increases external validity of the model overall, yet leads to underestimation of the importance of the predictors.

^cValues per unit increase.

^dTo interpret the coefficients and hazard ratios of the main effects of WHO performance status, primary tumour location and combination chemotherapy, the interaction effects need to be taken into account. For example, the AFT coefficient of combination versus sequential chemotherapy for a patient with WHO performance status 2, synchronous disease, and right-sided primary tumour is: $1 \times 0.07 + 2 \times 0.06 + 1 \times 0.24 + 0 \times -0.19 = 0.43$. The hazard ratio is: $\exp(1 \times \ln(0.90) + 2 \times \ln(0.92) + 1 \times \ln(0.69) + 0 \times \ln(1.32)) = 0.53$.

In CAIRO, 294 patients (36.6%) had a predicted median gain of >3 months in favour of combination chemotherapy, 278 patients (34.6%) between 0 and 3 months in favour of combination chemotherapy, and 231 patients (28.8%) in favour of sequential chemotherapy. A difference with a *p* value <.05 in favour of sequential chemotherapy was observed in 142 patients (17.7%), compared to 150 patients (18.7%) for combination chemotherapy. Patients with a predicted benefit for combination chemotherapy had a worse WHO PS, more often synchronous metastatic disease and/or a right-sided primary tumour (Table 3). Patients with a predicted benefit for sequential chemotherapy had a left-sided primary tumour location, a primary tumour resection, and metachronous disease.

The decision curves are visualized in Figures 1 and 2. In CAIRO data (Figure 1), these curves show that prediction-based treatment is more favourable than treating all patients with combination treatment regardless of the treatment threshold. However, when the treatment threshold is >1.8 months, the net benefit for prediction-based treatment is inferior to treating all patients with sequential chemotherapy. Treating only patients with a predicted treatment effect with a *p* value <.05 results in a higher net benefit compared to treating all patients with prediction-based treatment regardless of statistical significance when the threshold is higher than approximately 1.5 months (Figure 1), but is associated with a negative net benefit when the threshold is >2.0 months. In FOCUS data, prediction-based treatment is more favourable than treating all patients with combination chemotherapy when the treatment threshold is >1.0 months, and the net benefit of prediction-based treatment becomes comparable to treating all patients with sequential chemotherapy when the treatment threshold is >2.2 months (Figure 2). Supplementary Table 1 shows the clinical implications of treating patients according to the decision curve analyses of CAIRO data.

Prediction model for estimating the probability of receiving second-line treatment

A total of 5 patients in CAIRO (2.4%) and 89 patients in FOCUS (6.3%) were excluded from the model for predicting the probability of receiving second-line treatment in patients treated with first-line monochemotherapy due to incomplete follow-up, resulting in a training set of 396 patients and an external validation set of 1333 patients. One or more variables were missing in 119 (30.1%) of CAIRO and 452 (33.9%) of FOCUS patients. In CAIRO, 267 (67%) patients in the sequential chemotherapy group received second-line treatment, compared to 796 (60%) in FOCUS. Exposure to second-line treatment was strongly associated with longer OS in the CAIRO population (median OS: 19.4 [95% CI: 18.0–20.9] versus 7.9 months [6.5–9.3], respectively; HR 0.53 [0.43–0.65]; *p* < .01).

Age, WHO PS, BMI, WBC, resection of primary tumour and primary tumour location were identified as predictive variables for receiving second-line treatment after the first-line monochemotherapy. All continuous variables were linearly associated with the endpoint and no interactions terms were identified. Model coefficients accompanied with *p*-values and unpenalized odds ratios with corresponding 95% CIs are shown in Table 4 and visualized in the nomogram (Supplementary Figure 6).

The median-predicted probability of receiving second-line chemotherapy after the first-line monochemotherapy in the overall population is 71% (range: 22–90%). In patients with a predicted survival gain between −3.0 and 3.0 months, the median-predicted probability is 70% (range: 26–87%). The internal and external calibration plots illustrated a good validation, with C-indices of 0.68 and 0.65, respectively (Supplementary Figures 7 and 8). The calculator is illustrated in Supplementary Figure 9.

Table 3. Baseline characteristics of CAIRO patients according to predicted overall survival by sequential chemotherapy versus combination chemotherapy.

	Predicted benefit in favour of sequential chemotherapy, <i>n</i> = 231 (29%)	Predicted benefit of >0 and ≤3 months in favour of combination chemotherapy, <i>n</i> = 278 (35%)	Predicted benefit of >3 months in favour of combination chemotherapy, <i>n</i> = 294 (37%)
Age, median (IQR), years	65 (58–71)	61 (54–68)	63 (56–69)
Male sex, no. (%)	144 (62)	194 (70)	169 (57)
WHO performance status, no. (%)			
0	164 (71)	175 (63)	162 (55)
1	60 (26)	92 (33)	116 (40)
2	7 (3)	11 (4)	16 (5)
BMI, median (IQR), kg/m ²	26.0 (23.9–28.4)	24.2 (22.1–26.7)	24.9 (23.0–27.7)
Number of metastatic sites, no. (%)			
1	95 (41)	86 (31)	179 (61)
2	89 (39)	111 (40)	90 (31)
≥3	47 (20)	81 (29)	25 (9)
Presentation of metastatic disease, no. (%)			
Synchronous	0	223 (80)	100
Metachronous	231 (100)	55 (20)	0
Primary tumour resected, no. (%)	231 (100)	161 (58)	241 (82)
Primary tumour location, no. (%)			
Left colon and rectum	231 (100)	206 (74)	153 (52)
Right colon	0	72 (26)	141 (48)
WBC, median (IQR), ×10 ⁹ /L	7.5 (6.3–8.9)	8.9 (7.5–11.3)	7.7 (6.4–9.5)
ALP, median (IQR), U/L	97 (82–131)	170 (111–297)	105 (81–152)

OS: overall survival; BMI: body mass index; WBC: white blood cell count; ALP: alkaline phosphatase. SI conversion factor: to convert alkaline phosphatase to microkatal per liter, multiply by 0.0167.

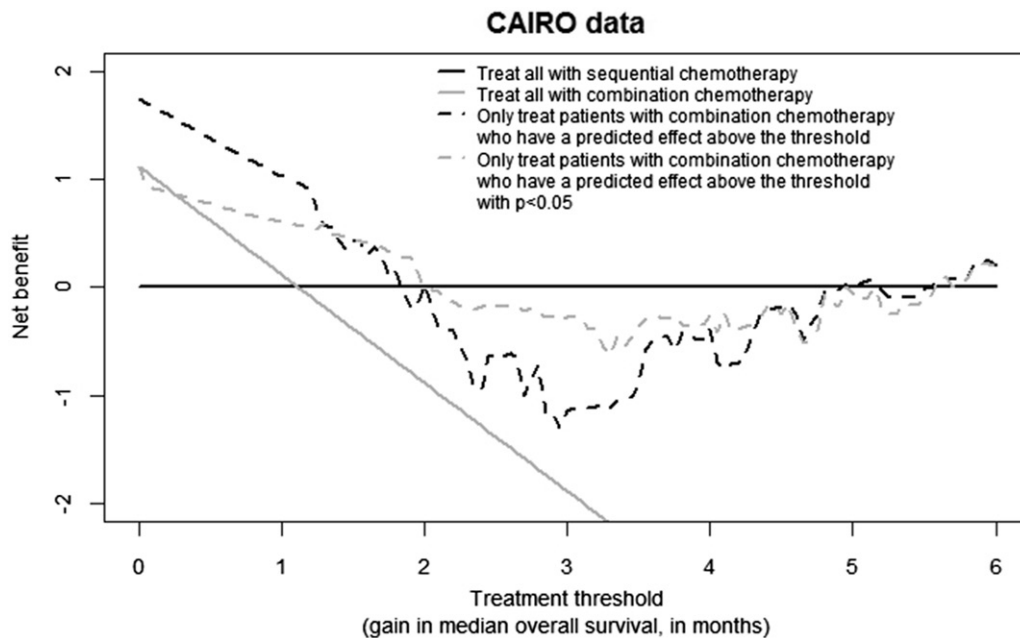


Figure 1. Decision curves for net benefit assessment regarding overall survival of various decision strategies as present in CAIRO data. Reading the net benefit plot starts with choosing a treatment-threshold, which is the gain in median overall survival at which one would opt for combination chemotherapy instead of sequential chemotherapy (i.e., from that point onwards, the benefits are considered to outweigh the harms, e.g., toxicity). Positive net benefit means that the treatment strategy led to a more favourable trade-off between benefits and harms.

Discussion

In this study, we demonstrate the development of two complementary individualized prediction models based on data from randomized trials. Our models show good graphical calibrations, proper internal and external validity, and substantial discriminative abilities, which are key aspects of such models [8]. Our results indicate that a substantial heterogeneity exists in survival times for upfront combination chemotherapy compared to sequential chemotherapy starting with single-agent fluoropyrimidine monotherapy in mCRC patients, and that treatment effects can be predicted using a combination of easily obtainable patient characteristics.

These models may contribute to the ultimate promise of personalized treatment in mCRC, where therapy can be accurately tailored for each individual patient.

Randomized phase III trials in mCRC generally represent a heterogeneous study population, which is evidenced by a large amount of clinical and molecular prognostic parameters that have been identified in recent years [20]. To date, our ability to predict the clinical treatment effects of specific therapeutic approaches in individual patients remains limited. The presented models represent evidence-based tools to guide treatment decisions in clinical practice. Our models provide individualized absolute treatment effects, include a

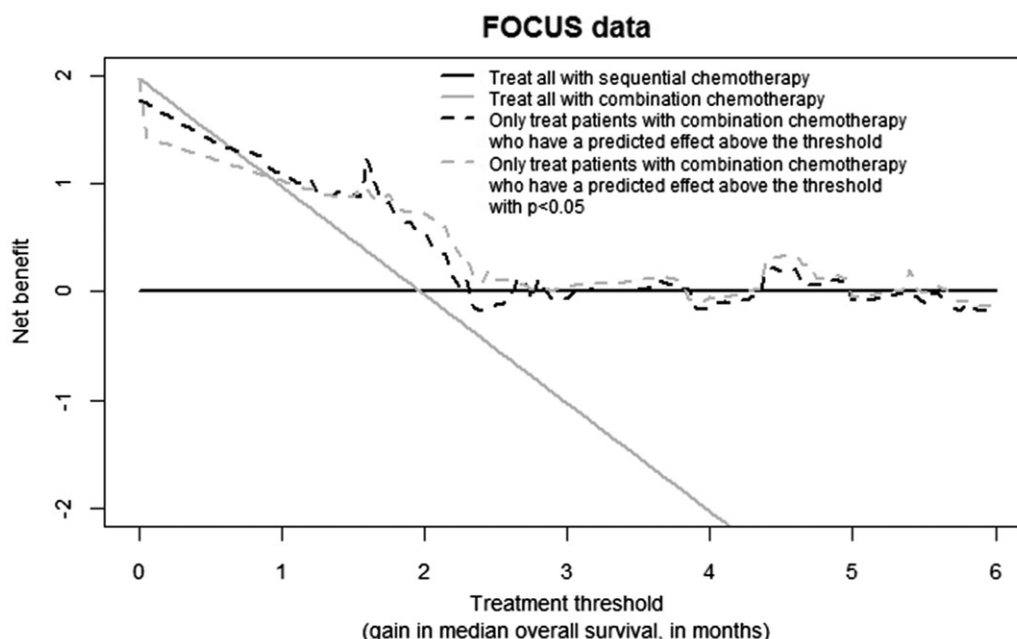


Figure 2. Decision curves for net benefit assessment regarding overall survival of various decision strategies as present in FOCUS data. Reading the net benefit plot starts with choosing a treatment-threshold, which is the gain in median overall survival at which one would opt for combination chemotherapy instead of sequential chemotherapy (i.e., from that point onwards, the benefits are considered to outweigh the harms, e.g., toxicity). Positive net benefit means that the treatment strategy led to a more favourable trade-off between benefits and harms.

Table 4. Variables associated with exposure to second-line treatment after first-line treatment with monotherapy.

	Distribution ^a	β^b	OR ^b	95% CI	<i>p</i> -value
Age, median, years ^c	64				
56		0.00	Ref.		
70		-0.23	0.75	0.54 to 1.05	.09
WHO performance status, no. (%)					
0	253 (64%)	0.00	Ref.		
1	125 (32%)	-0.23	0.75	0.46 to 1.22	.24
2	18 (5%)	-0.42	0.33	0.11 to 0.98	.05
BMI, median, kg/m ^{2c}	24.9				
22.8		0.00	Ref.		
27.0		0.22	1.28	0.97 to 1.68	.08
Resection primary tumour, no. (%)					
No	89 (23%)	0.00	Ref.		
Yes	307 (78%)	0.53	1.85	1.08 to 3.17	.02
Sidedness of primary tumour, no. (%)					
Right	101 (28%)	0.00	Ref.		
Left or rectum	259 (72%)	0.81	2.62	1.62 to 4.21	<.01
WBC, median, $\times 10^9/L^c$	8.1				
6.6		0.00	Ref.		
9.3		-0.32	0.74	0.55 to 0.99	.05

β : parameter estimate; OR: odds ratio; CI: confidence interval; WBC: white blood cell count; LDH: lactate dehydrogenase.

^aDistribution of variables is based on unimputed values.

^bPenalized maximum likelihood was applied to the AFT coefficients, but not the hazard ratios, because penalizing increases the external validity of the model overall, yet leads to underestimation of the importance of the predictors.

^cThe lowest and highest interquartiles are used as reference and comparator, respectively.

combination of patient and tumour characteristics, and allow for the evaluation of multiple potential treatment interactions. This systematic approach of model development and validation has primarily been used in cardiovascular diseases and has led to the implementation of calculators for individualized treatment effect predictions in clinical practice [5,21]. Van Kruijsdijk and colleagues demonstrated that the methodology is also suitable for survival time predictions in non-small-cell lung cancer patients [6]. This indicates that the design may be suitable for any well-specified clinical question in the presence of an extensive database, preferably randomized trial data.

Primary tumour location and synchronous metastatic disease were the main discriminatory predictors for a survival benefit from a combination or sequential chemotherapy. Previous subgroup analyses of chemotherapy trials suggested that upfront intensified treatment would be beneficial for patients with a poor PS [2,14]. Therefore, we added WHO PS as a treatment interaction, but the relative effect of this variable on treatment outcome was low. This supports that treatment outcomes are dependent upon multiple rather than a single characteristic.

We identified several different predictive factors for estimating the probability of receiving second-line treatment as

compared to predicting survival time. Also, a wide range of estimated probabilities was observed, indicating that the models can be used complementarily. For example, a 45-year-old male patient with metastases at three sites and an ALP level of 100 U/L has a limited predicted benefit for combination chemotherapy of 2.5 months and an estimated probability of receiving second-line treatment after monotherapy of 82% when he has a left-sided, unresected primary tumour, synchronous metastatic disease, WHO PS 0, WBC count $4.5 \times 10^9/L$, and BMI 28. The same patient has a predicted survival benefit of 2.3 months for combination chemotherapy and an estimated probability of receiving second-line treatment of 33% when he has right-sided and resected primary tumour, metachronous metastatic disease, WHO PS 2, WBC count $15.0 \times 10^9/L$, and BMI 18. In the latter situation, upfront combination chemotherapy may be more appropriate due to the low probability of receiving second-line treatment after the first-line monotherapy.

In order to objectify the benefit of prediction-based treatment for survival, we constructed decision curve analyses with a net benefit. Net benefit aims to determine whether predictions from a model can be used to apply the results of randomized trials to individual patients, as opposed to using group-level results [19]. Treatment thresholds are hereto implemented as a measure of weighing harms and benefits. Our results show that the net benefit for treating patients according to prediction-based treatment is superior to treating all patients with upfront combination chemotherapy regardless of the treatment threshold, but that the net benefit is lower compared to treating all patients with sequential chemotherapy when the threshold is >1.8 months. Importantly, setting a treatment threshold is subjective in nature for both physicians and patients and requires a careful balance between potential toxicities/effect on quality of life and expected survival gain. The value of statistical significance or confidence intervals in individualized prediction models is questionable since we are prone to select the treatment strategy that is likely to result in the best outcome, regardless of whether we believe it will be superior most of the time [22].

Our study has a few major limitations. Prognostic parameters as serum LDH values and *BRAF/RAS*-mutation status were not included due to missing data [10,23,24]. Since patients with *BRAF*^{V600E}-mutated tumours have a poor life expectancy and are less likely to be exposed to second-line treatment [23,25], triplet chemotherapy in combination with an anti-VEGF antibody is currently recommended as first-line treatment [7]. Also, after the publication of the CAIRO and FOCUS trials, the introduction of anti-VEGF antibodies, anti-EGFR monoclonal antibodies, regorafenib, and trifluridine/tipiracil have further improved the life expectancy of mCRC patients. Hence, the predicted survival times of our model are an underestimation of the current survival times. Nonetheless, chemotherapy used in CAIRO and FOCUS remains the backbone of first- and second-line treatments, and – with the addition of a targeted drug – are still valid first-line treatment options. When molecular characteristics are added as prognostic characteristics and the

discriminative abilities of our models are validated in a patient population treated according to the latest guidelines, we believe that our models can be implemented in daily practice.

In conclusion, we demonstrate that absolute treatment effects in mCRC can be estimated with systematically developed personalized prediction models derived from randomized trial data. With the use of readily available patient and tumour characteristics, the optimal treatment strategy for individual patients with mCRC can be selected. Such tools can be used to facilitate shared decision making and enable us to further tailor treatment decisions.

Disclosure statement

Prof. Dr. C.J.A.P. and Prof. Dr. M.K. received research grants from the Dutch Colorectal Cancer Group and Servier outside the submitted work; Dr. J.J.M.K. received research grants and lecture fees from Servier and Nordic Pharma outside the submitted work. All remaining authors have declared no conflicts of interest.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ORCID

Cornelis J. A. Punt  <http://orcid.org/0000-0003-0846-1445>
Miriam Koopman  <http://orcid.org/0000-0003-1550-1978>

References

- [1] Koopman M, Antonini NF, Douma J, et al. Sequential versus combination chemotherapy with capecitabine, irinotecan, and oxaliplatin in advanced colorectal cancer (CAIRO): a phase III randomised controlled trial. *Lancet*. 2007;370:135–142.
- [2] Seymour MT, Maughan TS, Ledermann JA, 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;370:143–152.
- [3] Simkens LH, Van Tinteren H, May A, et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet*. 2015;385:1843–1852.
- [4] Tournigand C, Cervantes A, Figuer A, et al. OPTIMOX1: a randomized study of FOLFOX4 or FOLFOX7 with oxaliplatin in a stop-and-go fashion in advanced colorectal cancer – a GERCOR study. *J Clin Oncol*. 2006;24:394–400.
- [5] Van Kruijsdijk RC, Visseren FL, Ridker PM, et al. Individualised prediction of alternate-day aspirin treatment effects on the combined risk of cancer, cardiovascular disease and gastrointestinal bleeding in healthy women. *Heart*. 2015;101:369–376.
- [6] Van Kruijsdijk RC, Visseren FL, Boni L, et al. Pemetrexed plus carboplatin versus pemetrexed in pretreated patients with advanced non-squamous non-small-cell lung cancer: treating the right patients based on individualized treatment effect prediction. *Ann Oncol*. 2016;27:1280–1286.
- [7] Van Cutsem E, Cervantes A, Adam R, et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol*. 2016;27:1386–1422.

- [8] Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925–1931.
- [9] Iasonos A, Schrag D, Raj GV, et al. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol*. 2008;26:1364–1370.
- [10] Köhne CH, Cunningham D, Di Costanzo F, et al. Clinical determinants of survival in patients with 5-fluorouracil-based treatment for metastatic colorectal cancer: results of a multivariate analysis of 3825 patients. *Ann Oncol*. 2002;13:308–317.
- [11] Renfro LA, Loupakis F, Adams RA, et al. Body mass index is prognostic in metastatic colorectal cancer: pooled analysis of patients from first-line clinical trials in the ARCAD database. *J Clin Oncol*. 2016;34:144–150.
- [12] Holch JW, Ricard I, Stintzing S, et al. 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.
- [13] Venderbosch S, de Wilt JH, Teerenstra S, et al. Prognostic value of resection of primary tumor in patients with stage IV colorectal cancer: retrospective analysis of two randomized studies and a review of the literature. *Ann Surg Oncol*. 2011;18:3252–3260.
- [14] Cremolini C, Loupakis F, Antoniotti C, 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;16:1306–1315.
- [15] Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer; 2009.
- [16] Gerds TA, Kattan MW, Schumacher M, et al. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. *Statist Med*. 2013;32:2173–2184.
- [17] Harrell FE. Regression modeling strategies with applications to linear models, logistic regression, and survival analysis. New York: Springer; 2001.
- [18] Grothey A, Sargent D, Goldberg RM, et al. Survival of patients with advanced colorectal cancer improves with the availability of fluorouracil-leucovorin, irinotecan, and oxaliplatin in the course of treatment. *J Clin Oncol*. 2004;22:1209–1214.
- [19] Vickers AJ, Kattan MW, Daniel S. Method for evaluating prediction models that apply the results of randomized trials to individual patients. *Trials*. 2007;8:14.
- [20] Punt CJ, Koopman M, Vermeulen L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat Rev Clin Oncol*. 2017;14:235–246.
- [21] Kaasenbrood L, Bhatt DL, Dorresteijn JAN, et al. Estimated life expectancy without recurrent cardiovascular events in patients with vascular disease: the SMART-REACH model. *J Am Heart Assoc*. 2018;7:e009217.
- [22] Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ*. 1999;18:341–364.
- [23] Tol J, Nagtegaal ID, Punt CJ. BRAF mutation in metastatic colorectal cancer. *N Engl J Med*. 2009;361:98–99.
- [24] Schwartzberg LS, Rivera F, Karthaus M, et al. PEAK: a randomized, multicenter phase II study of panitumumab plus modified fluorouracil, leucovorin, and oxaliplatin (mFOLFOX6) or bevacizumab plus mFOLFOX6 in patients with previously untreated, unresectable, wild-type KRAS exon 2 metastatic colorectal cancer. *J Clin Oncol*. 2014;32:2240–2247.
- [25] Seligmann J, Fisher D, Elliott F, et al. Exploring the poor outcomes of BRAF mutant (BRAF Mut) advanced colorectal cancer (aCRC): analysis from 2,530 patients (pts) in randomized clinical trials (RCTs). *J Clin Oncol*. 2015;33:abstr3509.