A nomogram to predict severe toxicity in DPYD wild-type patients treated with capecitabine-based anticancer regimens

Knikman, J.E.; Lopez-Yurda, M.; Meulendijks, D.; Deenen, M.J.; Schellens, J.H.M.; Beijnen, J.; ... ; Guchelaar, H.J.

A Nomogram to Predict Severe Toxicity in DPYD Wild-Type Patients Treated With Capecitabine-Based Anticancer Regimens

Jonathan E. Knikman[1,2,*], Marta Lopez-Yurda[3], Didier Meulendijks[1,4,5], Maarten J. Deenen[6,7], Jan H.M. Schellens[8], Jos Beijnen[2,9], Annemieke Cats[10] and Henk-Jan Guchelaar[7]

***DPYD*-guided dosing has improved the safety of fluoropyrimidine-based chemotherapy in recent years. However, severe toxicity remains in ~23% of patients not carrying *DPYD* variant alleles treated with capecitabine. Therefore, we developed a predictive model based on patient-related and treatment-related factors aimed at estimating the risk of developing severe capecitabine-related toxicity. The nomogram was developed using data from two large clinical trials (NCT00838370 and NCT02324452). Patients with cancer carrying a *DPYD* variant allele (*DPYD**2A, c.1236G>A, c.2846A>T, and c.1679T>G) were excluded. Univariable and multivariable logistic regression using predetermined predictors based on previous findings, including age, sex, body surface area, type of treatment regimen, and creatinine levels were used to develop the nomogram. The developed model was internally validated using bootstrap resampling and cross-validation. This model was not externally or clinically validated. A total of 2,147 *DPYD* wild-type patients with cancer treated with capecitabine-based chemotherapy regimens were included of which complete data of 1,745 patients were available and used for the development of the nomogram. Univariable and multivariable logistic regression showed that age, sex, and type of treatment regimen were strong predictors of severe capecitabine-related toxicity in *DPYD* wild-type patients. Internal validation demonstrated a concordance index of 0.68 which indicates a good discriminative ability for prediction of severe capecitabine-related toxicity. The developed nomogram includes readily available parameters and may be a helpful tool for clinicians to assess the risk of developing severe capecitabine-related toxicity in patients without known risk *DPYD* variant alleles treated with capecitabine-based anticancer regimens.**

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Dose-individualization of fluoropyrimidines based on *DPYD* genotype is now widely recommended in clinical practice guidelines. However, severe fluoropyrimidine-related toxicity remains present in ~23% of *DPYD* wild type patients. Previously, factors besides *DPYD* status, such as sex, age, body surface area, treatment schedule, and renal function, have been associated and could potentially predict with severe fluoropyrimidine-related toxicity.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ Can prediction tool based on patient-related and treatment-related factors accurately predict severe fluoropyrimidine-related toxicity in *DPYD* wildtype patients treated with capecitabine?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ In this study, we developed a prediction tool with good discriminative ability for prediction of severe fluoropyrimidine-related toxicity in *DPYD* wildtype patients treated with capecitabine. Age, sex, and type of treatment were strong predictors of severe fluoropyrimidine-related toxicity.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ Our developed prediction tool may be a helpful for clinicians to assess the risk of developing severe fluoropyrimidine-related toxicity in *DPYD* wildtype patients.

[1]Division of Pharmacology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; [2]Department of Pharmacy & Pharmacology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; [3]Biometrics Department, The Netherlands Cancer Institute, Amsterdam, The Netherlands; [4]Division of Medical Oncology, Department of Clinical Pharmacology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; [5]Late Development Oncology, AstraZeneca, Cambridge, UK; [6]Department of Clinical Pharmacy, Catharina Hospital, Eindhoven, The Netherlands; [7]Department of Clinical Pharmacy and Toxicology, Leiden University Medical Centre, Leiden, The Netherlands; [8]Department of Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands; [9]Division of Pharmacoepidemiology and Clinical Pharmacology, Department of Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands; [10]Division of Medical Oncology, Department of Gastrointestinal Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands. *Correspondence: Jonathan E. Knikman (j.knikman@nki.nl)

Received February 15, 2023; accepted October 25, 2023. doi:10.1002/cpt.3100

Capecitabine is an anticancer agent belonging to the group of fluoropyrimidines and is a pro-drug of 5-fluorouracil (5-FU) and is widely used in the treatment of various cancers.[1–3] Despite being used for over 2 decades, the efficacy of capecitabine is often negatively impacted by severe fluoropyrimidine-related toxicity, resulting in dose reductions, delays, treatment discontinuation, loss of quality of life, and, in some cases, even death.[4–6] Approximately 10%–30% of patients treated with capecitabine experience severe toxicity, which includes nausea, diarrhea, vomiting, mucositis, neutropenia, and hand-foot syndrome.[4,5] One of the main causes of these toxicities during treatment with fluoropyrimidines is a deficiency of the main catabolic enzyme dihydropyrimidine dehydrogenase (DPD). Genetic polymorphisms in the *DPYD* gene, which encode for the DPD enzyme, can reduce the metabolism of 5-FU into inactive metabolites, thereby affecting the risk of severe fluoropyrimidine-induced toxicity.[7,8] Hence, pre-therapeutic screening for *DPYD* variant alleles (*DPYD**2A, c.1236G>A, c.2846A>T, and c.1679T>G) and subsequent dose-individualization were studied and proved to reduce severe fluoropyrimidine-related toxicity in *DPYD* variant allele carriers.[8,9] As a result, *DPYD* genotyping is now widely recommended by several clinical guidelines and the European Medicines Agency (EMA) and used in several European countries.[10–12] Although *DPYD* genotype-guided dosing reduces the incidence of toxicity, nearly a quarter of the *DPYD* wild-type patients still experience severe fluoropyrimidine-related toxicity.[8,9] Besides *DPYD* genotyping, DPD phenotyping methods have been explored to further reduce the incidence of severe fluoropyrimidine-related toxicity.[13] However, these methods are also aimed toward detecting DPD deficiency and rarely take other factors into account which could influence the risk of developing severe toxicity emphasizing the need for dose-individualization strategies for patients with cancer without DPD deficiency. Previously, other factors besides DPD, such as sex, body composition, age, body surface area (BSA), type of capecitabine-based treatment regimen, and renal function, have been associated with the early onset of fluoropyrimidine-related toxicity.[14–16] It has been suggested that women have decreased 5-FU clearance and increased 5-FU exposure, and therefore are at increased risk of developing severe fluoropyrimidine-related toxicity.[13] Moreover, the possible relation between body composition and severe toxicity could possibly be explained by the relatively low proportion of lean body mass or muscle mass in women compared with men. Furthermore, a higher clearance of 5-FU and subsequently a lower risk of severe toxicity has been found in patients with higher BSA.[14] Interestingly, the association between renal function and severe fluoropyrimidine-related toxicity was unexpected as 5-FU is predominantly metabolized in the liver and tumor tissue.[17] However, pooled data from phase I studies showed that creatinine clearance significantly influences exposure to 5-FU. Indicating that renal function needs to be considered when dosing fluoropyrimidines, even though the exact mechanism by which renal function increases risk of severe fluoropyrimidine-related toxicity is unclear.[13,17] These patient- and treatment-related factors could potentially be used as a dose-individualization strategy for *DPYD* wild-type patients treated with fluoropyrimidines to reduce the remaining risk for severe toxicity. Therefore, we aimed to develop a prediction tool based on patient-related and treatment-related factors that could accurately predict severe toxicity in patients without known risk *DPYD* variant alleles treated with capecitabine.

METHODS

Patient population

Patients from two large multicenter clinical trials (Deenen *et al.*[8] (NCT00838370) and Henricks *et al.*[9] (NCT02324452)) including 1,463 and 913 patients with cancer, respectively. Only patients treated with capecitabine-based treatment regimens were included, due to the small number of patients treated with 5-FU in both trials.[8,9] The design and study population of both studies have previously been published.[8,9] Briefly, in Deenen *et al.*,[8] patients were prospectively screened for *DPYD**2A, and heterozygous *DPYD**2A variant carriers received a 50% fluoropyrimidine dose reduction. In addition, patients were also retrospectively screened for c.1236G>A, c.2846A>T, c.1679T>G, and c.1601G>A. In Henricks *et al.*,[9] upfront genotyping of 4 *DPYD* variant alleles was performed. *DPYD**2A and c.1679T>G variant allele carriers received a 50% fluoropyrimidine dose reduction, and c.1236G>A and c.2846A>T variant allele carriers a 25% fluoropyrimidine dose reduction. Patients carrying a *DPYD* variant allele (*DPYD**2A, c.1236G>A, c.2846A>T, c.1679T>G, and c.1601G>A) were excluded from the analysis, resulting in 1,302 and 845 patients, respectively (**Figure 1**). All toxicities were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) version 3.0 or 4.0 and from day 1 of treatment until the end of treatment with CTCAE grade ≥ 3 being considered as severe toxicity. Only toxicities scored for causality as possibly, probable, or definitely related to fluoropyrimidines were taken into account for fluoropyrimidine-related toxicity.

Prediction model construction and nomogram

The outcome of interest in this study was severe (grade ≥ 3) capecitabine-related toxicity during treatment with capecitabine-based regimens. Age, sex, BSA, treatment regimen (**Table S1**), and renal function were previously shown to be associated with capecitabine-related toxicity,

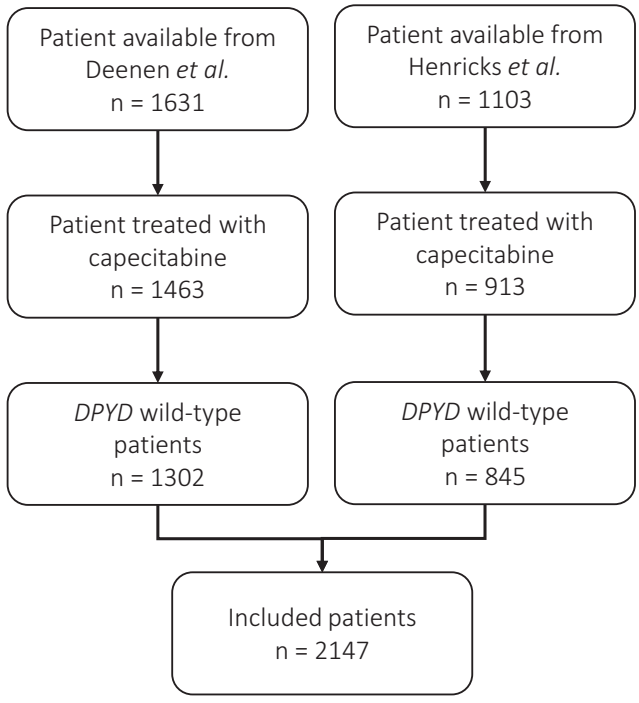

Figure 1 Flow diagram of patient inclusion.

and therefore included as covariates in the multivariable logistic regression model, regardless of their significance in the univariable logistic regression analysis.[13–15] However, due to the correlation between renal function (glomerular filtration rate (GFR)) and age, sex, and BSA (dependent on the formula used to calculate GFR) serum creatinine levels were used as a marker for renal function instead. A nomogram was constructed from this model to facilitate its interpretation in a visual way, by computing predicted capecitabine-related severe toxicity probabilities and mapping them into points on a scale from 0 to 100. For this purpose, the estimates of effect of the different covariates in the multivariable model were ranked, regardless of their statistical significance, by absolute value. The biggest effect was assigned 100 points on the scale, whereas the rest of covariates in the multivariable model were assigned a number of points proportional to their effect size.

Statistical analysis

Patient characteristics for continuous variables were summarized as mean (±standard deviation) or median (interquartile range), depending on their distribution. For categorical variables, frequency and percentage were presented. Categorical variables were compared using Pearson's chi-square test (Fisher's exact test in case of sparse data) and the Mann–Whitney *U* test was used to test differences in continuous variables.

Univariable and multivariable logistic regression models were used in the development of the prediction model for the nomogram. Correlations between variables were assessed using Pearson's and Spearman's correlation coefficients. The inclusion of interaction terms was explored by estimating pairwise interactions using a *P* value of 0.01 as cutoff for inclusion in the model, and restricted cubic splines were used to assess nonlinear relationships with the regression outcome. The discriminative power of the model was evaluated by calculating the area under the receiving operating characteristics curve (AUC), which corresponds with the concordance index, and its corresponding 95% confidence interval (CI). Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and prevalence were also calculated. CIs for predictive values were calculated according to Mercaldo *et al.*[18] Accuracy was evaluated with locally estimated scatterplot smoothing (LOESS)-based calibration curves and confidence bands (smoothing parameter 0.75) and the mean absolute error, which was calculated from the difference between the actual (observed) probability and the predicted probability of toxicity grade ≥3 with smoothing using the LOESS algorithm. The nomogram was internally validated using bootstrap resampling and leave-one-out cross-validation to provide an unbiased estimate of the model performance with the concordance index. The clinical utility of the prediction model in the nomogram was estimated by decision analysis curves,[19] based on the threshold probability (that is, the probability at which the harm of falsely declaring toxicity equals the harm of falsely declaring non-toxicity). Statistical analyses were performed using R statistical software (version 4.2.1).

RESULTS

Nomogram construction

A total of 2,147 wild-type patients (**Figure 1**) were included. An overview of patient characteristics of included patients from both studies is shown in **Table 1**. For 1,745 patients, all predefined predictors to be used in the nomogram were available for a complete case analysis. The prevalence of toxicity grade ≥3 among these patients was 20% (19% in the Deenen *et al.* study and 21% in Henricks *et al.*). Univariable and multivariable logistic regression results are displayed in **Table 2**. Age, sex, and type of treatment regimen were strong predictors of toxicity with increasing risk of severe toxicity with age (per 10 years an increase in odds ratio (OR) of 1.17, 95% CI: 1.04–1.32, $P = 0.01$) and male sex having a decreased risk of developing severe toxicity (OR: 0.68, 95% CI: 0.49–0.95, $P = 0.02$). Pairwise interactions between all predictors in the model where explored, but for all of them global *P* values were above 0.10 and thus not included.

Figure 2 shows the developed nomogram that can be used to predict the likelihood for a patient to develop severe capecitabine-related toxicity. For example, a female patient (17 points), aged 45 years (19 points), with BSA 2.7 (0 points), serum creatinine level of 100 µmol/L (15 points), and receiving capecitabine in combination with a platinum compound (34 points) would have a total of 85 points, which corresponds to a probability of severe toxicity of 20%. In order to obtain this, a vertical line can be drawn on **Figure 2** intersecting sex equal to female, to then obtain at which number of points (first segment in **Figure 2**) the vertical line intersects. After performing these steps for each of the patient characteristics, the cumulative number of points is calculated and marked on the "Total Points" segment in **Figure 2**. From there, a vertical line crossing this number of total points can be drawn to obtain where it crosses the "Probability of Toxicity" segment right below. This will yield the probability of severe toxicity for this patient. As a second example, a male patient (0 points), aged 65 years (33 points), with BSA 1.7 (6 points), having a serum creatinine level of 135 µmol/L (22 points), and receiving capecitabine in combination with 2 other anticancer agents (capecitabine – triplet, 85 points) would have a total of 146 points, which corresponds to a probability of severe toxicity of 51%.

To accompany **Figure 2**, a dynamic nomogram was created using the shiny package in R software (https://biometricsdept.shinyapps.io/dynamic_nomogram). It must be noted that the ranges of predictor values used in the nomogram displayed in **Figure 2**, as well as in the dynamic nomogram, correspond to ranges in the data used for building the prediction model (except for age, which, for display purposes, has been represented ranging from 18 to 90 years). Applying a prediction model to patients with characteristics outside these ranges may compromise model performance, because this involves extrapolation of data.

Nomogram performance

The model's discriminative ability, as measured by the concordance index, was 0.68 (95% CI: 0.64–0.71). See **Figure 3** for the corresponding receiver operating characteristic curve. This indicates that our model can discern a patient with severe toxicity from a patient without severe toxicity 68% of the time. To correct for overfitting, the bias-corrected concordance index was obtained to be 0.67 using bootstrapping with 1,000 repetitions, and 0.67 with 10-fold cross-validation. The model's predictive accuracy can be observed in the calibration curve (**Figure S1**). This figure displays the predicted probabilities for the nomogram vs. the actual probabilities, which would fall in a 45-degree line if the prediction model were perfectly accurate. Judging from this figure, the calibration curve stays close to the reference line, with slight underprediction or overprediction along the range of predicted values, and poorer precision with increasing predicted values as well as values close to zero. The mean absolute error was 0.006 and can thus be considered small (the smaller this value, the better the calibration, with a value of zero indicating perfect calibration). For obtaining measures

Table 1 Patient characteristics per study

	Original dataset			Subset of patients used in prediction model		
	Deenen *et al.*[8] (*N* = 1,302)	Henricks *et al.*[9] (*N* = 845)	Total (*N* = 2,147)	Deenen *et al.*[8] (*N* = 977)	Henricks *et al.*[9] (*N* = 768)	Total (*N* = 1,745)
Age (years)						
N	1,302	845	2,147	977	768	1,745
Median (range)	61 (21–89)	64 (56–71)	62.0 (54–69)	60.5 (21–89)	64.0 (19–89)	62.0 (19–89)
Sex						
Female	741 (56.9%)	390 (46.2%)	1,131 (52.7%)	578 (59.2%)	350 (45.6%)	928 (53.2%)
Male	561 (43.1%)	455 (53.8%)	1,016 (47.3%)	399 (40.8%)	418 (54.4%)	817 (46.8%)
BSA						
N	1,302	775	2,077	977	768	1,745
Missing	0	70	70	0	0	0
Median (range)	1.9 (1.1–2.7)	1.9 (1.3–2.7)	1.9 (1.1–2.7)	1.9 (1.1–2.7)	1.9 (1.3–2.7)	1.9 (1.1–2.7)
Primary tumor						
Breast cancer	318 (24.4%)	100 (11.8%)	418 (19.5%)	282 (28.9%)	87 (11.3%)	369 (21.1%)
Colorectal cancer	712 (54.7%)	601 (71.1%)	1,313 (61.2%)	484 (49.5%)	547 (71.2%)	1,031 (59.1%)
Gastric cancer	163 (12.5%)	50 (5.9%)	213 (9.9%)	127 (13.0%)	49 (6.4%)	176 (10.1%)
Other[a]	109 (8.4%)	94 (11.1%)	203 (9.5%)	84 (8.6%)	85 (11.1%)	169 (9.7%)
Type of regimen[b]						
Missing	1	0	1	0	0	0
Capecitabine – monotherapy	382 (29.4%)	171 (20.2%)	553 (25.8%)	300 (30.7%)	149 (19.4%)	449 (25.7%)
Capecitabine – platinum	345 (26.5%)	345 (40.8%)	690 (32.2%)	183 (18.7%)	312 (40.6%)	495 (28.4%)
Capecitabine – taxane	57 (4.4%)	1 (0.1%)	58 (2.7%)	57 (5.8%)	1 (0.1%)	58 (3.3%)
Capecitabine – triplet	105 (8.1%)	47 (5.6%)	152 (7.1%)	82 (8.4%)	45 (5.9%)	127 (7.3%)
Capecitabine – other	15 (1.2%)	35 (4.1%)	50 (2.3%)	12 (1.2%)	31 (4.0%)	43 (2.5%)
Capecitabine – radiotherapy	397 (30.5%)	246 (29.1%)	643 (30.0%)	343 (35.1%)	230 (29.9%)	573 (32.8%)
Creatinine (μmol/L)						
N	978	836	1,814	977	768	1,745
Missing	324	9	333	0	0	0
Median (range)	71 (35–354)	73 (34–213)	72 (34–354)	71.0 (35–354)	74.0 (34–213)	71.0 (34–354)

BSA, body surface area.
[a]Other tumor types included: head and neck cancer, anal cancer, vulvar cancer, urethral cancer, esophagogastric cancer, and several rare tumor types.
[b]Capecitabine – platinum includes combinations of capecitabine and cisplatin or oxaliplatin and monoclonal antibodies (bevacizumab, trastuzumab or panitumumab); capecitabine – taxane includes combinations of capecitabine and docetaxel or paclitaxel; capecitabine – triplet includes combinations of docetaxel and oxaliplatin, cisplatin and epirubicin, oxaliplatin and epirubicin, and doxorubicin and cyclophosphamide. Capecitabine – other includes combination with irinotecan, monoclonal antibodies (bevacizumab, trastuzumab, or panitumumab), temozolomide, and vinorelbine; capecitabine-radiotherapy includes combinations of capecitabine, radiotherapy, and mitomycin C.

of diagnostic accuracy, we contemplated different choices for a probability threshold. The prevalence in the data used to build the nomogram was 20% (351/1,745), which led to sensitivity 0.54, specificity 0.71, PPV 0.32, and NPV 0.85. However, this threshold did not necessarily minimize misclassification of patients, and we aimed at maximizing the PPV and, in a lesser degree, the NPV. As the dose of capecitabine can be rapidly escalated in patients misclassified as high risk, those experiencing severe toxicity may need to interrupt treatment or, in severe cases, require hospitalization. We therefore chose a threshold of 0.4 and we obtained a PPV of 0.49 (95% CI: 0.41–0.56), NPV of 0.83 (95% CI: 0.81–0.85), specificity of 0.94, and a sensitivity of 0.23. The relatively low value of the PPV is not surprising given that in our model prevalence is low (toxicity grade ≥ 3 occurs in 20% of patients), and it can be derived that the rarer the outcome, the higher the NPV and the lower the PPV.[20,21] We also attempted to evaluate the clinical utility of our model. The net benefit is calculated in true-positive units, as the proportion of true positives in the sample (benefit of adjusting the treatment due to predicted toxicity) minus the proportion of false positives in the sample (harm of adjusting treatment due to predicted toxicity) weighted by the odds of the threshold. The net benefit is calculated across all possible thresholds from 0 to 1 and is depicted for our prediction model as well as for default decisions of not adjusting treatment for anyone (net benefit zero) and adjusting treatment for all. Concerning our model, if the probability of severe toxicity is deemed high for a particular patient according to our chosen threshold of 0.4,

Table 2 Univariable and multivariable logistic regression results for probability of severe capecitabine-related toxicity

			Univariable logistic regression			Multivariable logistic regression		
Variable	Events	*N*	OR	95% CI	*P* value	OR	95% CI	*P* value
Sex								
Female	208	928	1 (ref.)			1 (ref.)		
Male	143	817	0.73	0.58–0.93	0.01	0.68	0.49–0.95	0.02
Age, per 10 years	351	1,745	1.05	0.94–1.17	0.36	1.17	1.04–1.32	0.01
BSA	351	1,745	0.68	0.4–1.17	0.17	0.87	0.44–1.72	0.69
Creatinine, per 10 µmol/L	351	1,745	1.00	0.95–1.06	0.87	1.05	0.99–1.12	0.13
Type of regimen								
Capecitabine – monotherapy	83	449	1 (ref.)			1 (ref.)		
Capecitabine – platinum	101	495	1.13	0.82–1.56	0.46	1.32	0.94–1.86	0.11
Capecitabine – taxane	31	58	5.06	2.87–8.94	<0.001	5.95	3.29–10.75	<0.001
Capecitabine – triplet	57	127	3.59	2.35–5.48	<0.001	4.27	2.74–6.66	<0.001
Capecitabine – other	15	43	2.36	1.21–4.62	0.01	2.47	1.25–4.89	0.01
Capecitabine – radiotherapy	64	573	0.55	0.39–0.79	<0.001	0.61	0.42–0.88	0.01

BSA, body surface area; CI, confidence interval; OR, odds ratio.

a dose reduction might be proposed, which in turn might lead to reduced efficacy. From **Figure 4**, it can be derived that with threshold probabilities between 20% and 50%, the net benefit of classifying patients at high risk of severe capecitabine-related toxicity would be higher than the default situations of assuming toxicity, and thus adjusting treatment, for all or none of the patients.

Additional analyses were performed to study the robustness of these results. A multivariable logistic model adjusted for study next to the predetermined predictors was also run to examine possible differences in severe toxicity between studies. No significant difference in severe fluoropyrimidine-related toxicity was found (hazard ratio: 1.22, 95% CI: 0.94–1.69, $P = 0.14$; results not shown) for Henricks *et al.*[9] vs. Deenen *et al.*[8] Furthermore, an additional analysis was performed in which the model was applied on patients treated with capecitabine monotherapy ($n = 449$ of whom 83 experienced severe toxicity) to assess the suitability of the model for prediction of severe toxicity, which could only be attributed to capecitabine. This resulted in a concordance-index of 0.58 (95% CI: 0.51–0.65).

The nomogram displayed in **Figure 2** was based on a complete-case analysis that omitted patients with missing creatinine and BSA. An analysis based on multiply imputed data was taken into consideration for tackling missing data. However, there were limited data available in both datasets that were correlated to the variables of interest in the model, or that helped to maintain the randomness of the missing process. For this reason, no appropriate auxiliary variables could be found for the imputation procedure and only a complete-case analysis was performed.

DISCUSSION

Over the last decade, safety of fluoropyrimidine-based treatment was greatly improved by *DPYD* genotype-guided dosing, significantly reducing the incidence of severe fluoropyrimidine-related toxicity.[8,9] However, despite the success of *DPYD*-guided dose-individualization severe toxicity remains in ~23% of patients without one of the 4 known risk *DPYD* variant alleles treated with fluoropyrimidines.[9] Our study aimed to develop a tool that could accurately predict severe fluoropyrimidine-related toxicity in *DPYD* wild-type patients treated with capecitabine-based chemotherapy regimens. This resulted in a nomogram including creatinine concentration, sex, age, type of treatment regimen, and BSA, which predicts the probability of developing severe capecitabine-related toxicity in patients treated with capecitabine-based treatment regimens. Our nomogram has a concordance-index of 0.67 after bias correction, which indicates a good discriminative ability of the model to predict severe capecitabine-related toxicity. This suggests that our model can relatively accurately predict the probability of severe fluoropyrimidine-related toxicity in wild-type patients treated with capecitabine-based treatment regimens and could also be easily used by clinicians in daily clinical practice because all required model parameters are readily available. However, this model has not yet been validated for clinical use.

The clinical validity of our model to predict severe toxicity was assessed by sensitivity, specificity, PPV, and NPV. The main aim of our model was to accurately predict severe capecitabine-related toxicity, therefore a high PPV is desired. However, possible misclassification of patients being at high risk of severe fluoropyrimidine-related toxicity is also not desirable, and therefore NPV values should not be too low either. A PPV of 49% and NPV of 83% were found in our study. This PPV could be interpreted as low. However, both PPV and NPV are relative to frequency of patients with severe toxicity. PPV can remain limited even though there is a high risk of severe fluoropyrimidine-related toxicity, if adverse events are rare. This was also the case in our study, with 20% of patients experiencing severe fluoropyrimidine-related toxicity. We therefore regarded our PPV and NPV as acceptable. PPV and NPV of *DPYD* testing ranges from 23.5% to 100% and 50.5% to 91.5%, respectively.[22–24] These results indicate that patients who carry a *DPYD* variant

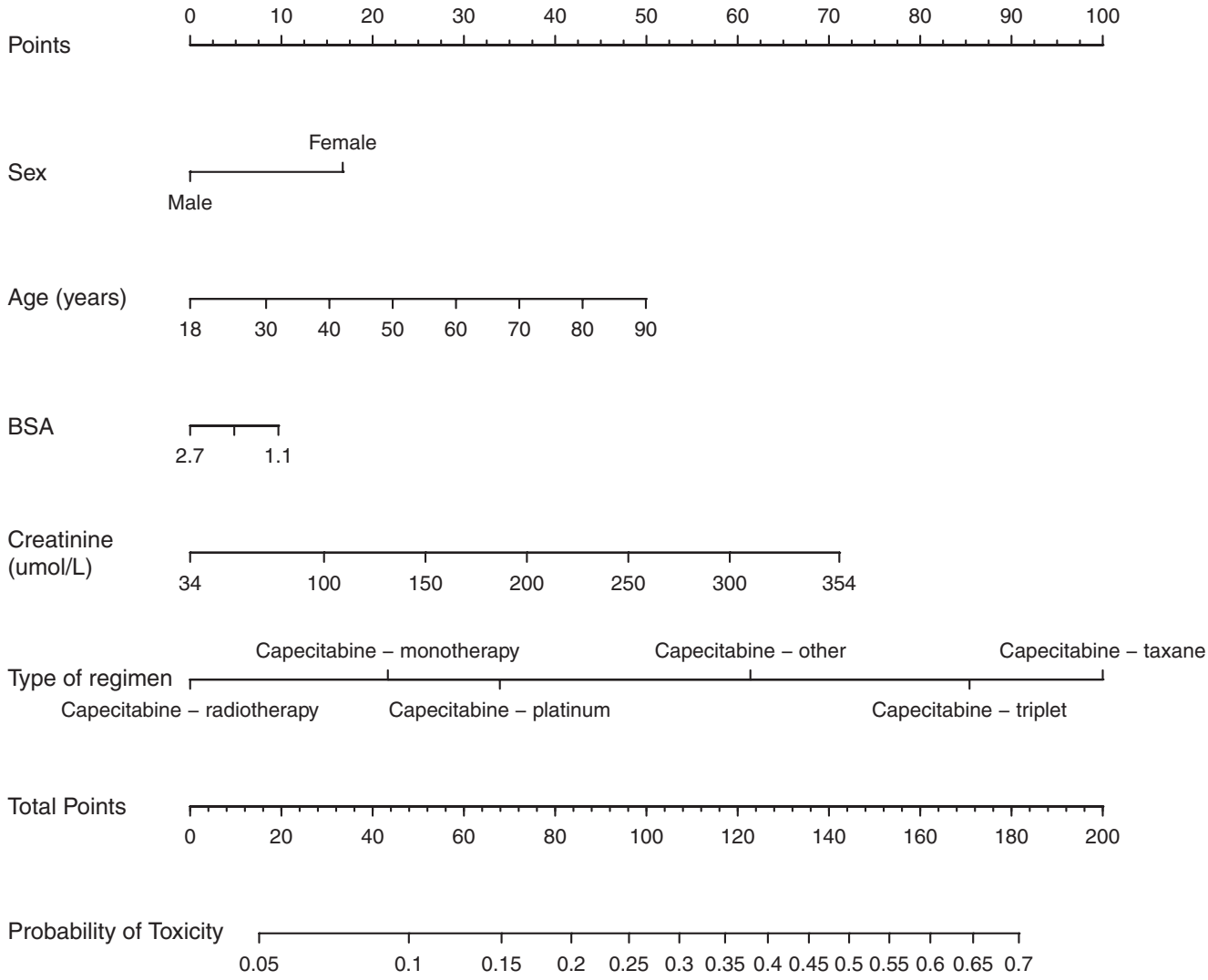

Figure 2 Nomogram to predict severe capecitabine-related toxicity using predetermined clinical predictors. BSA, body surface area; CAP, capecitabine.

allele have high risk of developing severe toxicity (high specificity). However, conversely noncarriers still develop severe toxicity which cannot be predicted by *DPYD* genetic testing. Additionally, due to the high prevalence of severe toxicity in *DPYD* variant carriers it was expected that the PPV would be relatively high. In our cohort, the relative prevalence of toxicity is significantly lower compared with the prevalence of toxicity in *DPYD* variant carriers and therefore a lower PPV was expected when using our model. These results indicate that our model could be complementary to *DPYD* genotyping and could further reduce the risk of severe toxicity in patients treated with fluoropyrimidines without a large risk of suboptimal treatment.

Ideally, this model would be used in a multi-parametric approach, as shown in **Figure 5**. Such a two-step decision tool could be used in patients who are first screened for *DPYD* variant alleles associated with severe fluoropyrimidine-related toxicity. Subsequently, if none of the four *DPYD* variants are present, our nomogram could be used to predict the probability of developing severe fluoropyrimidine-related toxicity. If, for example, the probability of severe toxicity exceeds 40% a dose reduction could be applied. However, it should be kept in mind that this model has not yet been externally validated and predicted probabilities of severe toxicity should be interpreted with the appropriate caution. As the optimal threshold has not been determined yet, up titration of individualized doses based in the subsequent cycles based on toxicity is recommended to ensure an adequate and safe dose for all patients.

A possible useful additional variable could be pretreatment uracil levels, as it has been shown to be associated with an increased risk of severe fluoropyrimidine-related toxicity.[25] However, due to critical pre-analytical factors, it is currently not yet deemed suitable to include uracil in the nomogram. It is therefore possible that our predictive accuracy may increase when using uracil levels as a predictor in our model. However, uracil is a complex biomarker influenced by multiple factors, including food intake, circadian rhythm, and instability at room temperature after blood sampling.[26–28] By

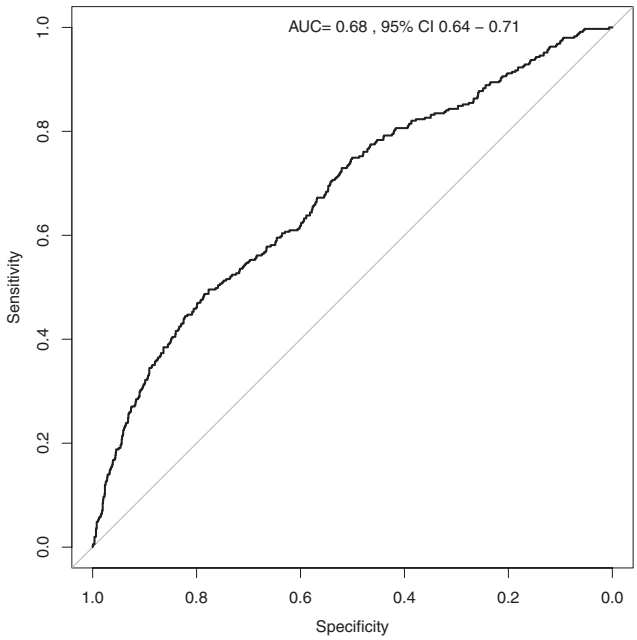

Figure 3 Receiver operating characteristic curve for probability of severe capecitabine-related toxicity. AUC, area under the curve; CI, confidence interval.

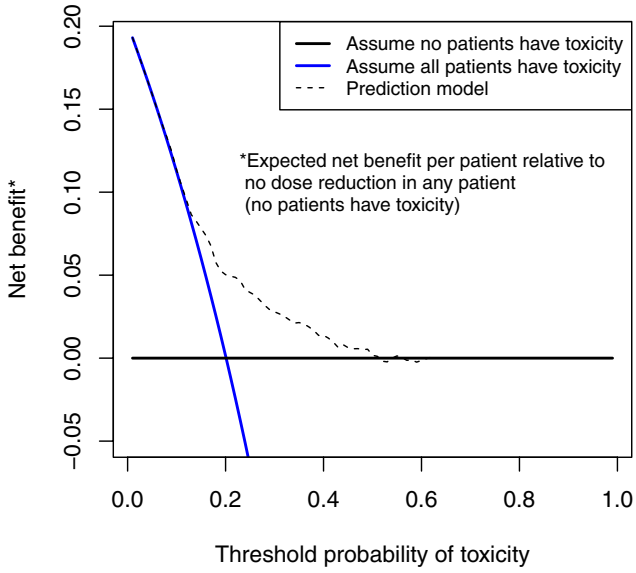

Figure 4 Decision curve for prediction model in nomogram. [Color figure can be viewed at wileyonlinelibrary.com]

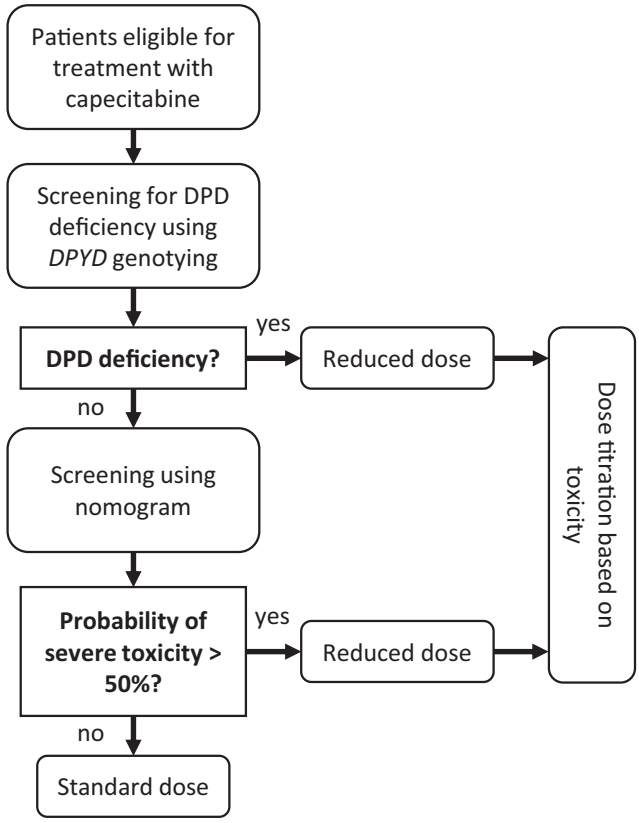

Figure 5 Example of possible approaches for dose-individualization in patients treated with capecitabine-based treatment regimens using a two-step dosing strategy including *DPYD*-guided dosing and our multi-parametric nomogram.

including uracil, our model would become more complex and difficult to use in clinical practice.

One of the main limitations of our nomogram is that it is only applicable to patients treated with capecitabine-based treatment regimens, as creatinine levels were missing for all patients treated with 5-FU in Deenen *et al.*[8] Moreover, even within the subgroup receiving capecitabine-based treatment regimens, serum creatinine levels were only available for patients from two participating hospitals in Deenen *et al.*[8] Lack of auxiliary data hampered the use of multiple imputation techniques to deal with missing creatinine for the remaining patients.[8] Due to exclusion of these patients for the complete-case analysis, selection bias may have been introduced, although there were no indications in our dataset that missing creatinine data were related to patient condition or particular patient characteristics. Furthermore, it could be questioned whether this model is best suited for specifically predicting capecitabine-related toxicity as multi-drug regimens are included in the model. A model specifically aimed toward capecitabine could be considered to isolate the toxicity as being described to capecitabine. When applying the model only on capecitabine monotherapy patients the concordance-index was 0.58 (95% CI: 0.51–0.65), indicating a substantially lower discriminative ability for prediction of severe fluoropyrimidine-related toxicity. This raises the question whether the toxicity can be fully attributed to treatment with capecitabine. However, only toxicity related to fluoropyrimidines has been considered. It is possible that the simultaneous use of other drugs in a multi-drug treatment regimen reduces the tolerability of the treatment with capecitabine. Alternatively, a novel algorithm based on capecitabine patients only could be developed. However, in clinical practice, patients are often treated with multi-drug regimens and therefore such a model may be of limited value. The multiple combination regimens, as shown in our study and collected from real-world data, underscore this heterogeneity. Furthermore, our model was only internally validated. To assess if the model accurately predicts toxicity in a clinical setting and to determine the

ideal threshold, a prospective validation in a large external cohort is required.

CONCLUSION

We developed a simple nomogram using easily measured or obtainable variables that can predict severe toxicity and may be useful in improving the safety of capecitabine-based treatment regimens in patients without the four known *DPYD* risk variant alleles. This nomogram requires further validation through external and prospective validation to ensure adequate prediction of toxicity. Nonetheless, our nomogram is a simple and easy tool for physicians to estimate the risk of severe capecitabine-related toxicity and to further personalize capecitabine treatment to reduce severe toxicity.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

ACKNOWLEDGMENTS

The authors thank our colleagues involved in the clinical trials of which the data were used to develop the prediction model; in particular all the co-investigators Marjolein K. Schechterberger, Johan L. Severns, Henk Boot, Paul H. Smits, Caroline Mandigers, Marcel Soesan, Geert-Jan Creemers, Arnold Baars, Vincent O. Dezentjé, Alexander L.T. Imholz, Frank J. F. Jeurissen, Johanna E. A. Portielje, Rob L. H. Jansen, Paul Hamberg, Albert J. ten Tije, Helga J. Droogendijk, Miriam Koopman, Peter Nieboer, Marlène H. W. van de Poel, Caroline M. P. W. Mandigers, Ron H. J. Mathijssen, Jesse J. Swen, and Hans Gelderblom.

FUNDING INFORMATION

No funding was received for this work.

CONFLICT OF INTEREST

J.H.M.S. and J.B. are (part-time) employees, stock- and patent holders of Modra Pharmaceuticals, a spin-out company developing oral taxane formulations. J.H.M.S. is also a part-time employee of Byondis bv and received consultancy fees from Debiopharm, all not related to the contents of the manuscript. D.M. is a current full-time employee and shareholder of AstraZeneca, not related to the contents of the manuscript. All other authors declared no competing interest for this work.

AUTHOR CONTRIBUTIONS

J.E.K. and M.L.-Y. wrote the manuscript. J.E.K., M.L.-Y., D.M., M.J.D., J.H.M.S., J.B., A.C., and H.-J.G. designed the research. J.E.K. and M.L.-Y. performed the research. J.E.K. and M.L.-Y. analyzed the data.

1. Blum, J.L. *et al*. Pooled analysis of individual patient data from capecitabine monotherapy clinical trials in locally advanced or metastatic breast cancer. *Breast Cancer Res. Treat.* **136**, 777–788 (2012).
2. Ocvirk, J., Reberšek, M., Škof, E., Hlebanja, Z. & Boc, M. Randomized prospective phase II study to compare the combination chemotherapy regimen epirubicin, cisplatin, and 5-fluorouracil with epirubicin, cisplatin, and capecitabine in patients with advanced or metastatic gastric cancer. *Am. J. Clin. Oncol.* **35**, 237–241 (2012).
3. Haller, D.G. *et al*. Capecitabine plus oxaliplatin compared with fluorouracil and folinic acid as adjuvant therapy for stage III colon cancer. *J. Clin. Oncol.* **29**, 1465–1471 (2011).
4. Twelves, C. *et al*. Capecitabine as adjuvant treatment for stage III colon cancer. *N. Engl. J. Med.* **352**, 2696–2704 (2005).
5. Rosmarin, D. *et al*. Genetic markers of toxicity from capecitabine and other 5-fluorouracil- based regimens investigation in the QUASAR2 study, systematic review and meta-analysis. *J Clin Oncol* **32**, 1031–1039 (2014).
6. Hoff, P.M., Ansari, R., Batist, G. *et al*. Comparison of oral capecitabine versus intravenous fluorouracil plus leucovorin as first-line treatment in 605 patients with metastatic colorectal cancer: results of a randomized phase III study. *J. Clin. Oncol.* **19**, 2282–2292 (2017).
7. Meulendijks, D. *et al*. Clinical relevance of DPYD variants c.1679T>G, c.1236G>A/HapB3, and c.1601G>A as predictors of severe fluoropyrimidine-associated toxicity: a systematic review and meta-analysis of individual patient data. *Lancet Oncol.* **16**, 1639–1650 (2015).
8. Deenen, M.J. *et al*. Upfront genotyping of DPYD*2A to individualize fluoropyrimidine therapy: a safety and cost analysis. *J. Clin. Oncol.* **34**, 227–234 (2016).
9. Henricks, L.M. *et al*. DPYD genotype-guided dose individualisation of fluoropyrimidine therapy in patients with cancer: a prospective safety analysis. *Lancet Oncol.* **19**, 1459–1467 (2018).
10. Amstutz, U. *et al*. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for Dihydropyrimidine dehydrogenase genotype and fluoropyrimidine dosing: 2017 update. *Clin. Pharmacol. Ther.* **103**, 210–216 (2018).
11. European Medicines Agency. EMA recommendations on DPD testing prior to treatment with fluorouracil, capecitabine, tegafur and flucytosine (2020); 31: 3.
12. With, M., Knikman, J., Man, F.M. *et al*. Dihydropyrimidine dehydrogenase phenotyping using pretreatment uracil: a note of caution based on a large prospective clinical study. *Clin. Pharmacol. Ther.* **112**, 62–68 (2022).
13. Knikman, J.E., Gelderblom, H., Beijnen, J.H., Cats, A., Guchelaar, H.J. & Henricks, L.M. Individualized dosing of fluoropyrimidine-based chemotherapy to prevent severe fluoropyrimidine-related toxicity: what are the options? *Clin. Pharmacol. Ther.* **0**, 1–14 (2020).
14. Meulendijks, D. *et al*. Renal function, body surface area, and age are associated with risk of early-onset fluoropyrimidine-associated toxicity in patients treated with capecitabine-based anticancer regimens in daily clinical care. *Eur. J. Cancer* **54**, 120–130 (2016).
15. Prado, C.M.M., Baracos, V.E., McCargar, L.J. *et al*. Body composition as an independent determinant of 5-fluorouracil-based chemotherapy toxicity. *Clin. Cancer Res.* **13**, 3264–3268 (2007).
16. Stein, B.N., Petrelli, N.J., Douglass, H.O., Driscoll, D.L., Arcangeli, G. & Meropol, N.J. Age and sex are independent predictors of 5-fluorouracil toxicity. Analysis of a large scale phase III trial. *Cancer* **75**, 11–17 (1995).
17. EMEA. Capecitabine Scientific Discussion (2005).
18. Mercaldo, N.D., Lau, K.F. & Zhou, X.H. Confidence intervals for predictive values with an emphasis to case-control studies. *Stat. Med.* **26**, 2170–2183 (2007).
19. Vickers, A.J. & Elkin, E.B. Decision curve analysis: a novel method for evaluating prediction models. *Med. Decis. Making* **26**, 565–574 (2006).
20. Altman, D.G. & Martin, B.J. Diagnostic tests 1: sensitivity and specificity. *BMJ* **308**, 6493 (1994).
21. Altman, D.G. & Bland, J.M. Diagnostic test 2: predictive values. *BMJ* **309**, 6947 (1994).
22. Innocenti, F., Mills, S.C., Sanoff, H., Ciccolini, J., Lenz, H.-J. & Milano, G. All you need to know about DPYD genetic testing for patients treated with fluorouracil and capecitabine: a practitioner-friendly guide. *JCO Oncol. Pract.* **16**, 793–798 (2020).
23. Lee, A.M., Shi, Q., Pavey, E. *et al*. DPYD variants as predictors of 5-fluorouracil toxicity in adjuvant colon cancer treatment (NCCTG N0147). *J. Natl. Cancer Inst.* **106**, 1–12 (2014).
24. Health, O. DPYD genotyping in patients who have planned cancer treatment with fluoropyrimidines: a health technology

assessment. *Ontario Health Technol. Assess. Ser.* **21**, 1–186 (2021).
25. Meulendijks, D. *et al*. Pretreatment serum uracil concentration as a predictor of severe and fatal fluoropyrimidine-associated toxicity. *Br. J. Cancer* **116**, 1415–1424 (2017).
26. Henricks, L.M. *et al*. Food-effect study on uracil and dihydrouracil plasma levels as marker for dihydropyrimidine dehydrogenase activity in human volunteers. *Br. J. Clin. Pharmacol.* **84**, 2761–2769 (2018).
27. Jiang, H., Lu, J. & Ji, J. Circadian rhythm of dihydrouracil/uracil ratios in biological fluids: a potential biomarker for dihydropyrimidine dehydrogenase levels. *Br. J. Pharmacol.* **141**, 616–623 (2004).
28. van den Wildenberg, S.A.H. *et al*. Quantification of uracil, dihydrouracil, thymine and dihydrothymine for reliable dihydropyrimidine dehydrogenase (DPD) phenotyping critically depend on blood and plasma storage conditions. *J. Pharm. Biomed. Anal.* **221**, 115027 (2022).