Personalised medicine of fluoropyrimidines using DPYD pharmacogenetics

Lunenburg, C.A.T.C.

CHAPTER 10

Comparison of four phenotyping assays for predicting dihydropyrimidine dehydrogenase (DPD) deficiency and severe fluoropyrimidine-induced toxicity: a clinical study

Manuscript in preparation

Carin A.T.C. Lunenburg*, Linda M. Henricks*, Femke M. de Man*, Didier Meulendijks, André B.P. van Kuilenburg, Jan Gerard Maring, Maurice C. van Staveren, Niels de Vries, Hilde Rosing, Jos H. Beijnen, Dick Pluim, Anil Modak, Alexander L.T. Imholz, Jan H.M. Schellens, Jesse J. Swen, Hans Gelderblom, Annemieke Cats, Ron H.J. Mathijssen, Henk-Jan Guchelaar
**Contributed equally*

Abstract

Fluoropyrimidines are widely used anticancer drugs. Prospective *DPYD* (encoding dihydropyrimidine dehydrogenase, DPD, the key metabolic enzyme for degradation of fluoropyrimidines) genotyping followed by dose adjustments in *DPYD* variant allele carriers reduces severe fluoropyrimidine-induced toxicity. However, when using this approach still ~20% of patients experience severe toxicity. We evaluated four DPD phenotyping assays, aiming to determine which is most suitable for identifying patients at risk for severe fluoropyrimidine-induced toxicity, and identifying DPD deficient patients.

Study participants underwent testing of two, three or four DPD phenotyping assays before starting fluoropyrimidine-based therapy; the endogenous dihydrouracil/uracil (DHU/U) ratio, endogenous uracil levels, the oral uracil loading dose, and the 2-^{13}C-uracil breath test. Phenotyping results were associated with the onset of severe toxicity and DPD deficiency according to the DPD enzyme activity measurement in peripheral blood mononuclear cells. Sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV) and F1-score (harmonic mean of sensitivity and PPV) were calculated per phenotyping assay as predictive measures for severe (grade ≥3) fluoropyrimidine-induced toxicity and DPD deficiency.

In total, 1,037 patients participated in this study. Of these, 1,037, 92 and 82 patients underwent two, three or four DPD phenotyping assays, respectively. Two phenotyping assays were analysed on their performance for the prediction of severe fluoropyrimidine-induced toxicity. No differences were identified between wild-type patients who did or did not experience severe toxicity in the mean endogenous DHU/U ratio or mean endogenous uracil levels. The F1-scores of both assays were 10 and 24%, respectively. In the comparison of phenotyping assays in performance for prediction of DPD deficiency, four phenotyping assays were analysed in both wild-type patients and *DPYD* variant allele carriers. The highest F1-score of the phenotyping assays in predicting DPD deficiency was 40% for the oral uracil loading dose.

All four investigated DPD phenotyping assays in this study have been favourably evaluated as predictive test for the occurrence of severe fluoropyrimidine-induced toxicity in previous studies. However, in a first-time prospective head-to-head comparison study we could not show associations with the onset of severe fluoropyrimidine-induced toxicity or DPD deficiency. In order to determine the true clinical value of DPD phenotyping assays, additional research is required.

Introduction

Fluoropyrimidines, including 5-fluorouracil (5-FU) and its oral pro-drug capecitabine, play a key role in the treatment of multiple types of cancer.[1] Although 5-FU has been used for over 60 years, toxicity remains a major clinical problem, as severe fluoropyrimidine-induced side effects occur in up to 30% of patients, resulting in lethal outcome in up to 1% of these patients.[1,2] With over two million patients treated with fluoropyrimidines each year worldwide, many patients are at risk of developing severe toxicity.[3]

Abundant research has been carried out on dihydropyrimidine dehydrogenase (DPD), the key metabolic enzyme of fluoropyrimidines, and the gene *DPYD* encoding DPD. Low DPD activity itself and several *DPYD* variants resulting in low DPD activity have both individually been associated with severe fluoropyrimidine-induced toxicity.[4-6] Prospective phenotyping or genotyping, followed by dose adjustments in DPD deficient patients or *DPYD* variant allele carriers, can reduce the risk for severe toxicity. This was shown for prospective genotyping of *DPYD**2A, c.1679T>G, c.2846A>T and c.1236G>A/HapB3, followed by initially reduced dosages in *DPYD* variant allele carriers.[7,8] However, genotyping to predict severe fluoropyrimidine-induced toxicity has an inherently limited sensitivity, as other genetic and also non-genetic factors are known to play a role in the variability in DPD activity and the onset of severe fluoropyrimidine-induced toxicity. Phenotyping of the DPD enzyme, as a way to determine the DPD activity, can potentially better predict severe fluoropyrimidine-induced toxicity as it takes both pharmacogenetic and other factors influencing DPD activity into account.

A well-established method to determine DPD activity is measurement of DPD enzyme activity in peripheral blood mononuclear cells (PBMCs). The activity in PBMCs is well-correlated to the DPD enzyme activity in the liver, and reference values have been established.[6,9] However, the method is not widely used since feasibility in clinical practice remains challenging due to substantial costs, complex sample logistics and specific equipment required for the radio assay. In addition, the results are influenced by the distribution of blood cells (e.g. monocytes, granulocytes) in the sample,[10] and there is a substantial intra patient variability (up to 25%) in DPD enzyme activity, possibly caused by circadian rhythm.[11,12]

Several DPD phenotyping assays have previously been investigated, focussing on the metabolism of uracil (U) and dihydrouracil (DHU), the endogenous substrate and product of DPD, respectively. Two of these assays are the determination of the endogenous uracil levels and the DHU/U ratio. Several studies have shown an association between the pre-treatment endogenous DHU/U ratio in plasma and 5-FU pharmacokinetics,[13-16] and also with severe fluoropyrimidine-induced toxicity.[15,17-19] In addition, Meulendijks *et al.* have recently shown that high pre-treatment serum uracil concentrations were also strongly related to severe and fatal fluoropyrimidine-induced toxicity.[20] Another DPD phenotyping assay for estimating the *in vivo* DPD activity is the oral uracil loading dose assay.[21,22] In this assay, a high dose of uracil is administered orally, and uracil and DHU levels are measured using a limited sampling strategy.[21] In this way, the DPD enzyme function is utilized to the full capacity. In case of reduced uracil conversion, also partially DPD deficient patients can be identified. Finally, the *in vivo* DPD activity can also be determined using the 2-^{13}C-uracil breath test.[23]

This assay uses a personalized dose of 2-^{13}C uracil, a stable isotope of uracil, and is based on the conversion of 2-^{13}C uracil into $^{13}CO_2$ which can be measured in exhaled breath.[23]

Several of these assays have been evaluated in head-to-head comparisons to DPD enzyme activity measurements in PBMCs in healthy volunteers, or patients selected after experiencing severe toxicity.[11,21,23-27] However, when using an enriched patient cohort for severe toxicity, calculated assay characteristics such as sensitivity or specificity will be biased and not representative for routine clinical care. Strengths and weaknesses of these assays have been reviewed independently of each other;[28,29] however, a head-to-head comparison in clinical practice has never been investigated. Therefore, in this prospective study, we evaluated four DPD phenotyping assays in patients prior to treatment with fluoropyrimidines in order to determine the association with the onset of severe fluoropyrimidine-induced toxicity and detecting DPD deficiency, defined as DPD enzyme activity in PBMCs below the cut-off value.

Materials and methods

Study design and patients

This study was a pre-planned analysis in a large prospective multi-centre clinical trial (clinicaltrials.gov identifier NCT02324452, here referred to as main study cohort or clinical trial) improving the safety of fluoropyrimidines by prospective *DPYD* genotyping.[7] Two out of four phenotyping assays (endogenous DHU/U ratio and endogenous uracil levels) were executed in patients recruited in the seventeen Dutch hospitals participating in the clinical trial. All four DPD phenotyping assays were executed in three hospitals, and three out of four assays were executed in another hospital (excluding the 2-^{13}C-uracil breath test). Patient recruitment for this study was open from 30 April, 2015, until 21 December, 2017. Ethical approval of this study was granted by the medical ethical committee of The Netherlands Cancer Institute, Amsterdam, the Netherlands. All patients provided written informed consent before enrollment in this study.

All assays were executed before start of fluoropyrimidine therapy. 92 patients were asked to participate in all four phenotyping assays, which made intra-patient comparisons possible. Results of the DPD phenotyping assays were determined after start of treatment and were not used for dose individualization. Dose adjustments of the fluoropyrimidine drug were done based on *DPYD* genotype only as per protocol of the clinical trial.

Toxicity was graded according to the National Cancer Institute common terminology criteria for adverse events (CTCAE; version 4.03)[30] and severe toxicity was defined as CTCAE grade ≥3. Only toxicity defined by the treating physician as definitely, probably and possibly related to fluoropyrimidine treatment was taken into account. Patients were followed for toxicity during the entire treatment period and were evaluated for toxicity if they received at least one fluoropyrimidine drug administration. The endpoints of this study were the association of each DPD phenotyping assay with the onset of severe fluoropyrimidine-induced toxicity and DPD deficiency, defined as low DPD activity levels in PBMCs (≤5.9 nmol/[mg*h]).[6]

Inclusion and exclusion criteria for this study were the same as in the clinical trial; eligibility to start with fluoropyrimidine-based therapy, 18 years or older, an adequate performance

status, adequate renal and liver biochemistry and haematological values, and no prior treatment with fluoropyrimidines.

Study procedures

One blood draw for the endogenous DHU/U ratio and endogenous uracil level assays was taken prior to start of treatment. For patients participating in three or four DPD phenotyping assays the study scheme was as follows. During two random days prior to start of fluoropyrimidine-based treatment, all three or four phenotyping assays were performed in each patient (study scheme in Figure 1). On the first day, blood draws for the DPD enzyme activity in PBMCs and two DPD phenotyping assays (endogenous DHU/U ratio and endogenous uracil levels) were taken prior to 9 am. Immediately thereafter, the third phenotyping assay (oral uracil loading dose of 1000 mg uracil) was performed. The U/DHU ratio was assessed at 120 minutes after administration of uracil. At least one day later, but prior to start of fluoropyrimidines, the fourth phenotyping assay (2-^{13}C-uracil breath test with 6 mg/kg 2-^{13}C-uracil) was performed including blood draws for ^{13}C-uracil and ^{13}C-dihydrouracil plasma measurements. The DOB_{50} value from breath samples was correlated to ^{13}C-dihydrouracil plasma levels and the ^{13}C-DHU/U ratio. The oral uracil loading dose and 2-^{13}C-uracil breath test were performed on two separate days to exclude any interference, as uracil was administered orally for both assays. Also, a minimum time interval of 24 hours between the phenotyping assays and start of fluoropyrimidine treatment was taken into account as a safety precaution, although it was expected that the administered uracil would not affect the efficacy and safety of patients when starting their fluoropyrimidine-based treatment, since uracil has a very short half-life of around 40 minutes.[31]

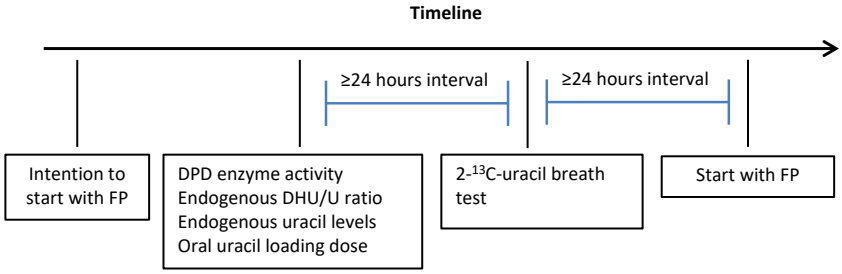


Figure 1. Study scheme

The study scheme per patient. Minimum interval between the tests and between tests and start of fluoropyrimidine-therapy was 24 hours. There was no predefined maximum number of days between assays as patients usually started relatively quickly with therapy when the decision to start was made. *Abbreviations:* FP: fluoropyrimidines; DHU: dihydrouracil; U: uracil.

DPD phenotyping assays

Patients underwent the DPD phenotyping assays in the hospital of their recruitment. Protocols for each DPD phenotyping assay were made available and discussed with executive personnel. In four hospitals, trained personnel was available to execute three or four DPD phenotyping assays. Each DPD phenotyping assay is described in more detail in the

supplementary material. In addition, the time of last food intake prior to the blood draw for the endogenous DHU/U ratio was collected in all patients.

Sample size calculation and statistical analyses

The sample size for comparison of the four phenotyping assays was based on the co-primary aim (the association between the result of a phenotyping assay and severe fluoropyrimidine-induced toxicity). It was calculated to be 240 patients (see detailed description in the supplementary material). The association between the result of a phenotyping assay and DPD deficiency (as determined by the DPD enzyme activity in PBMCs) was investigated as secondary aim.

Patient characteristics or toxicity differences between patient groups were tested using χ^2 test or non-parametric test. The DPD phenotyping assays were executed for the first time in the participating centres, i.e. in a research setting and were not cross-validated. To investigate the effect of centres on the outcome per assay, a mixed model analysis was executed to demonstrate the general reliability of the assay. A univariate analysis of variance was done to compare outcomes per centre. In this analysis, the centre with the highest recruitment rate was chosen as the reference centre. *DPYD* variant allele carriers were excluded in these analyses, as they are underwent a per protocol dose adjustment and hence are at lower risk for toxicity. Age, gender and baseline BSA were additionally taken into account as possible covariates to minimize the risk of biased results of the analyses. When age, gender or baseline BSA were associated with the outcome of the phenotyping assay, the distribution of the covariate was checked between centres using Chi-Square tests or univariate analysis of variance. The possible clinical consequence of the divergent results was discussed per assay and data could be excluded from further analyses.

For assessing clinical validity, measures to determine diagnostic performance (i.e. sensitivity, specificity, positive predictive value [PPV], negative predictive value [NPV] and F1-score [harmonic mean of sensitivity and PPV]) of the assays with regard to the onset of severe toxicity or DPD deficiency were determined. *DPYD* variant allele carriers who received dose reductions based on their identified genotype could only be included in an analysis for association with DPD activity, not for the association with onset of severe toxicity. The level of significance was set at $p<0.05$. Analyses were performed using SPSS, version 23 (IBM SPSS Inc., Chicago, IL, USA).

Results

Patients

In total, 1,037 patients participated in the phenotyping study of which 1,037, 92 and 82 patients underwent two, three and four phenotyping assays, respectively. Patient and treatment characteristics of the 92 patients were similar to those of the main study (*N*=1,103), with the exception that the 92 patients were slightly younger (median age of 60 versus 64 years, p=0.011, Table 1). Details on fluoropyrimidine-induced toxicity are depicted in Supplementary Table 1. In total, 19 out of 92 patients (21%) experienced severe fluoropyrimidine-induced toxicity, which is comparable to the main study in which 264 out of 1,103 patients (24%) experienced severe toxicity (p=0.477).

Phenotyping assays

General performance of assays

For the DPD enzyme activity measurements, the variance was estimated to be 8.6 ± 1.4 nmol/(mg*h). For centre as covariate, the variance was estimated to be 6.5 ± 5.7 nmol/(mg*h). The deviations between centres in general (intra class variation) was 43.0%, therefore the general reliability of the DPD enzyme activity measurements was 57%. For the endogenous DHU/U ratio, endogenous uracil levels, 2-^{13}C-uracil breath test and the oral uracil loading dose assay the general reliabilities were 74.1%, 92.9%, 73.5%, and 94.3%, respectively.

Endogenous DHU/U ratio

The endogenous DHU/U ratio was determined in 1,037 patients. Results of wild-type patients (non-carriers of the four *DPYD* variants, *N*=955) were compared between seventeen study centres. The endogenous DHU/U ratio differed significantly in nine study centres compared to the reference centre (lowest divergent mean DHU/U ratio 5.9, to the highest divergent mean DHU/U ratio 13.9, p<0.001). It appeared that age was significantly associated with the outcome of the DHU/U ratio (p<0.001). Age was differently distributed between the centres (p<0.001). The lowest statistically divergent mean DHU/U ratio was not as low as the suggested DHU/U ratio cut-off value (4.31)[20] for DPD deficient patients, therefore no patients were excluded. The median, interquartile range (IQR) and standard deviation (SD) of each assay are shown in Table 2.

Endogenous uracil levels

Endogenous uracil levels were determined in 1,037 patients. Results of wild-type patients (*N*=955) were compared between seventeen study centres. The endogenous uracil levels differed significantly in four study centres compared to the reference centre (lowest divergent mean uracil level 8.3 ng/ml, to the highest divergent mean uracil level 18.8 ng/ml, p<0.001). It appeared that gender was significantly associated with the outcome of the uracil levels (p=0.030), with lower uracil levels in females. Males and females were significant differently distributed between the centres (p=0.046). The divergent results were substantially higher, even higher than the previously suggested cut-off value (13.9 ng/ml)[20] for DPD deficient patients, therefore the data were considered unreliable and patients recruited in these centres (*N*=172) were excluded from further analyses. The endogenous uracil levels and endogenous DHU/U ratio were correlated to time of last meal that was eaten, to study the influence of food on the uracil levels. No correlation was found (Supplementary Figure 1), therefore time of food intake was not taken into account as covariate in further analyses.

DPD enzyme activity

The DPD enzyme activity in PBMCs was determined in 92 patients. Results of 82 wild-type patients were compared between study centres. The mean DPD enzyme activity was significantly lower in one of the four participating centres (5.23 nmol/(mg*h) versus 10.89 nmol/(mg*h) in the reference centre, p<0.001). These results were substantially lower, even lower than the cut-off value (≤5.9 nmol/[mg*h])[6] for DPD deficiency, therefore the data were considered unreliable and patients recruited in this centre (*N*=19) were excluded from further analyses in which DPD deficiency was taken into account.

Table 1. Baseline characteristics of patients who underwent three or four DPD phenotyping assays and from the main study cohort

Characteristic	Phenotyping assays (*N*=92)	Main study cohort (*N*=1,103)	P-value[a]
Sex			
Male	56 (61%)	593 (54%)	0.189
Female	36 (39%)	510 (46%)	
Age			
Median [IQR]	60 [53-67]	64 [56-71]	0.011
Ethnic origin			
Caucasian	87 (95%)	1048 (95%)	0.786
African descent	1 (1%)	19 (2%)	
Asian	2 (2%)	24 (2%)	
Other	2 (2%)	12 (1%)	
Tumour type			ND
Non-metastatic CRC	38 (41%)	472 (43%)	
Metastatic CRC	23 (25%)	232 (21%)	
BC	7 (8%)	141 (13%)	
GC	7 (8%)	63 (6%)	
Other	17 (18%)	195 (18%)	
Type of treatment regimen			0.854
CAP mono	12 (13%)	205 (19%)	
CAP + RT	23 (25%)	264 (24%)	
CAPOX	37 (40%)	374 (34%)	
CAP other	5 (5%)	72 (7%)	
5-FU mono	-	2 (0%)	
5-FU + RT	6 (7%)	63 (6%)	
FOLFOX	4 (4%)	43 (4%)	
5-FU other	5 (5%)	80 (7%)	
BSA			0.207
Median [IQR]	2.0 [1.79-2.10] (*N*=91)	1.92 [1.77-2.10] (*N*=1075)	
WHO performance status			0.151
0	49 (53%)	554 (50%)	
1	41 (45%)	448 (40%)	
2	1 (1%)	42 (4%)	
NS	1 (1%)	59 (5%)	
Number of treatment cycles			0.987
Median [IQR]	3 [1-6]	3 [1-8]	
DPYD status			0.281
Wild-type	82 (89%)	1018 (92%)	
DPYD variant allele carrier	10 (10.9%)	85 (7.7%)	
c.1236G>A heterozygous	6 (6.5%)	51 (4.6%)	
c.2846A>T heterozygous	3 (3.3%)	17 (1.5%)	
*DPYD*2A heterozygous*	1 (1.1%)	16 (1.5%)	
c.1679T>G heterozygous	-	1 (0.1%)	

[a] All p-values represent a comparison of 92 patients who underwent three or four DPD phenotyping

assays to patients from the main study cohort. We used a non-parametric test for independent samples to compare medians of age, BSA and number of treatment cycles; and a χ^2 test for gender, ethnic origin, treatment regimen and WHO status.
Abbreviations: IQR: interquartile range; CRC: colorectal cancer; BC: breast cancer; GC: gastric cancer; CAP mono: capecitabine monotherapy (with or without bevacizumab); CAP + RT: capecitabine combined with radiotherapy (with or without mitomycin); CAPOX: capecitabine combined with oxaliplatin (with or without bevacizumab); CAP other: capecitabine combined with other anticancer drugs; 5-FU mono: 5-fluorouracil monotherapy; 5-FU + RT: 5-fluorouracil combined with radiotherapy (with or without mitomycin); FOLFOX: 5-fluorouracil combined with oxaliplatin and leucovorin (with or without bevacizumab); 5-FU other: 5-fluorouracil combined with other anticancer drugs; BSA: body surface area; WHO: world health organisation; NS: not specified, either WHO 0, 1 or 2; ND: not done; *DPYD*; gene encoding dihydropyrimidine dehydrogenase.

Table 2. Overview of assay measurements and outcomes
DPD enzyme activity assay and four phenotyping assays are shown, including the number of patients per assay, the number of patients excluded from further analyses due to divergent results of the participating centre in which the assays were executed. After exclusion of these patients, the calculated medians, SD, and the IQR are shown for these patients. The median and SD are also shown for wild-type patients only and *DPYD* variant allele carriers only.

Phenotyping assay	*N* of patients	Patients excluded?	All patients			Wild-type patients		*DPYD* variant allele carriers	
			Median ± SD	IQR	*N* of patients	Median ± SD	*N* of patients	Median ± SD	*N* of patients
DPD enzyme activity	92	Yes (*N*=19)	10.3 ± 3.0 nmol/(mg*h)	8.0-12.9 nmol/(mg*h)	73	10.4 ± 3.0 nmol/(mg*h)	65	9.2 ± 3.1 nmol/(mg*h)	8[a]
Endogenous DHU/U ratio	1037	No	8.7 ± 3.9	6.6-11.6	1037	9.0 ± 3.9	955	7.2 ± 4.0	82
Endogenous uracil levels	1037	Yes (*N*=172)	10.2 ± 8.3 ng/ml	8.0-13.3 ng/ml	865	10.0 ± 8.4 ng/ml	794	13.3 ± 7.3 ng/ml	71
Oral uracil loading dose[b]	92	No	0.72 ± 0.86	0.33-1.17	92	0.58 ± 0.74	82	1.69 ± 1.23	10
2-^{13}C-uracil breath test[c]	82	No	158.9 ± 33.9 ‰	144.2-182.0 ‰	82	161.0 ± 34.0 ‰	74	145.8 ± 28.3 ‰	8

[a] Of these patients, six were carrier of the c.1236G>A variant, one was carrier of the c.2846A>T variant and one was a *DPYD**2A carrier;
[b] Results represent the U/DHU-ratio at 120 minutes;
[c] Results represent the DOB_{50}.
Abbreviations: SD: standard deviation; IQR: interquartile range; U: uracil; DHU: dihydrouracil; DOB_{50}: delta-over-baseline ratio at 50 minutes.

Oral uracil loading dose
The oral uracil loading dose assay was performed in 92 patients. Results of 82 wild-type patients were compared between study centres. The mean U/DHU ratio was significantly lower in one centre (0.622) compared to the reference centre (1.03, p=0.046). It appeared that baseline BSA was significantly associated with the outcome of the U/DHU ratio (p=0.008), a higher baseline BSA was related to a lower U/DHU ratio. Baseline BSA was not differently distributed between the centres (p=0.637). The different mean U/DHU ratio of one centre was far from the cut-off U/DHU ratio (2.4)[21] of DPD deficient patients, therefore no patients were excluded from further analyses.

2-^{13}C-uracil breath test
The 2-^{13}C-uracil breath test was determined in 82 patients.[23] On average, 488 mg 2-^{13}C-uracil was administered, ranging from 312 to 840 mg (6 mg/kg dose). Results of 74 wild-type patients were compared between study centres. The mean delta-over-baseline ratio at t=50 minutes (DOB_{50}) was significantly lower in one centre (137.7 ‰) compared to the other two centres (173.5 ‰ and 168.4 ‰, p<0.009). It appeared that gender was significantly associated with the outcome of the DOB_{50} (p=0.003), with higher DOB_{50} values in females. Males and females were not significant differently distributed between the centres (p=0.263). The significantly different mean DOB_{50} was not as low as the DOB_{50} cut-off value (128.9 ‰)[25] for DPD deficient patients, therefore no patients were excluded. A significant correlation between the DOB_{50} determined in breath samples and the ^{13}C-dihydrouracil plasma levels (r^2=0.178, p<0.001) could be demonstrated, not for the ^{13}C-DHU/U ratio (r^2=0.014, p=0.29). Results are shown in Supplementary Figure 2.

Association with onset of severe toxicity
Clinical validity parameters, i.e. sensitivity, specificity, NPV, PPV and F1-score of the endogenous DHU/U ratio and the endogenous uracil levels for their association with the onset of severe fluoropyrimidine-induced toxicity were calculated and shown in Table 3. The endogenous uracil levels have the highest F1-score of 24%. No significant difference was identified between the median endogenous DHU/U ratio or endogenous uracil level between patients who experienced severe toxicity or not (Figure 2). For the oral uracil loading dose and 2-^{13}C-uracil breath test too few patients were enrolled, therefore the association with the onset of severe toxicity was investigated in an explorative way only for these two phenotyping assays (Supplementary Table 2, Supplementary Figure 3). Yet, the data show similar results in clinical validity parameters and also no difference between medians of patients who experienced severe toxicity or not.

Association with DPD deficiency
DPD deficiency, defined as low DPD activity levels in PBMCs (≤5.9 nmol/[mg*h]),[6] was identified in 7 out of 73 patients (9.6%) or 6 out of 64 patients (9.4%). Clinical validity parameters for association with DPD deficiency are shown in Table 4. High specificity and NPV values were identified, but low sensitivity and PPV values. The oral uracil loading dose has the highest F1-score of 40%. The endogenous uracil levels have the highest sensitivity of 43%.

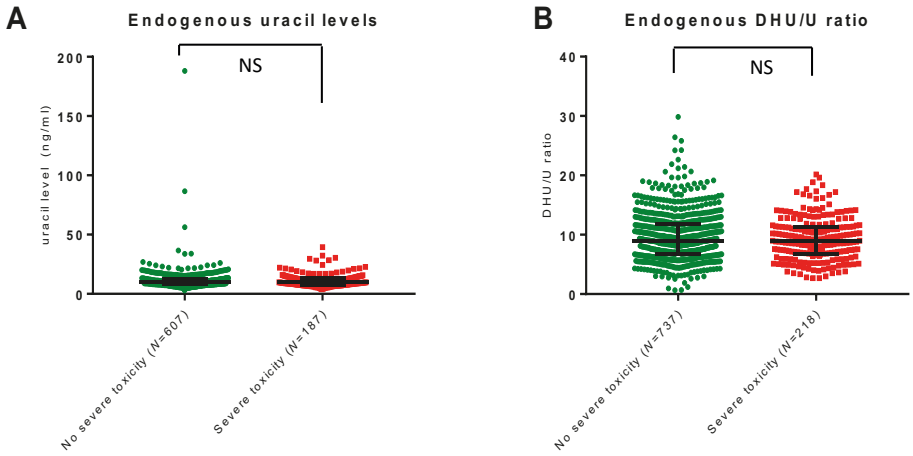


Figure 2. Results of phenotyping assays separated by the occurrence of severe fluoropyrimidine-induced toxicity

Dots represent individual results. Black lines represent the median and 25th and 75th percentile of the data. All *DPYD* variant allele carriers were excluded from the analysis as they received initial dose reductions based on their genotype result. For the endogenous uracil levels, 161 wild-type patients were excluded due to divergent phenotyping assay results of the centre in which these patients were recruited.

Abbreviations: U: uracil; DHU: dihydrouracil; NS: not significant p-value.

Table 3. Comparison of phenotyping assays in performance for prediction of severe fluoropyrimidine-induced toxicity

Clinical validity parameters for the prediction of severe (grade ≥3) fluoropyrimidine-induced toxicity are shown for the endogenous DHU/U ratio and endogenous uracil levels. *DPYD* variant allele carriers were excluded, since *DPYD* variant allele carriers received an initial dose reduction based on their genotype, and therefore bias could exist in the onset of severe fluoropyrimidine-induced toxicity. For the endogenous uracil levels, 161 wild-type patients were excluded due to divergent phenotyping assay results of the centre in which these patients were recruited.

Assay	*N* of patients	Median (IQR)	Cut-off for DPD deficiency	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)	F1-score[a] (%)
Endogenous DHU/U ratio	955	9.0 (6.8-11.6)	≤4.31[b]	6	95	77	26	10
Endogenous uracil levels	794	10.0 (7.8-12.9) ng/ml	≥13.9 ng/ml[20]	21	82	77	27	24

[a] The F1-score represents the harmonic mean of sensitivity and PPV;

[b] This cut-off value is determined by calculating the 6% lower limit of the data, as was described by Meulendijks *et al.* [20]

Abbreviations: IQR: interquartile range; DPD: dihydropyrimidine dehydrogenase; NPV: negative predictive value; PPV: positive predictive value; DHU: dihydrouracil; U: uracil.

Discussion

Despite recent advances by applying prospective *DPYD* genotyping, ~20% of patients treated with fluoropyrimidines still suffer from severe toxicity.[7] These patients are wild-type for the four genotyped *DPYD* variants, yet could still be DPD deficient due to currently untested variants. Therefore, it is of great importance to explore the clinical value of DPD phenotyping assays in order to potentially further reduce the risk of severe fluoropyrimidine-induced toxicity. In this study, we conducted two DPD phenotyping assays in 1,037 patients and 82 patients underwent all four DPD phenotyping assays, in order to rule out inter-individual variation. To the best of our knowledge, this is the first study with this unique design, taking into account that our patient cohort was not selected based on –or enriched for– (severe) toxicity, but represents a patient cohort more representative of routine clinical care. However, in the analyses with severe toxicity we excluded *DPYD* variant allele carriers, thus relatively more wild-type patients were included. Still, some wild-type patients are DPD deficient, indicating that we were able to calculate assay performance measures, such as sensitivity and specificity, for the onset of severe toxicity. In the comparison of DPD deficiency, *DPYD* variant allele carriers were not excluded.

Table 4. Comparison of phenotyping assays in performance for prediction of DPD deficiency

Per phenotyping assay clinical validity parameters are shown for the prediction of DPD deficiency. DPD deficiency is defined as a DPD enzyme activity in PBMCs ≤5.9 nmol/(mg*h), and was identified in 7 out of 73 patients (9.6%) and 6 out of 64 patients (9.4%). The results of the DPD enzyme activity were substantially divergent in one centre. Therefore, these results were considered unreliable and could not be compared to results of the phenotyping assays in predicting DPD deficiency. 19 patients were excluded.

Assay	*N* of patients	Median (IQR)	Cut-off for DPD deficiency	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)	F1-score[a] (%)
Endogenous DHU/U ratio	73	8.3 (6.4-11.1)	≤4.31[b]	14	97	91	33	20
Endogenous uracil levels	73	11.8 (8.9-14.9) ng/ml	≥13.9 ng/ml[20]	43	73	92	14	21
2-^{13}C-uracil breath test	64[c]	DOB_{50}: 159.0 (140.3-181.7) ‰	DOB_{50} ≤128.9 ‰[23,25]	33	86	93	20	25
Oral uracil loading dose	73	U/DHU-ratio at 120 min: 0.61 (0.31-1.15)	U/DHU-ratio at 120 min ≥2.4[21]	29	98	93	67	40

[a] The F1-score represents the harmonic mean of sensitivity and PPV;

[b] This cut-off value is determined by calculating the 6% lower limit of the data, as was described by Meulendijks *et al*.[20];

[c] The 2-^{13}C-uracil breath test was executed in 64 out of 73 patients.

Abbreviations: IQR: interquartile range; DPD: dihydropyrimidine dehydrogenase; NPV: negative predictive value; PPV: positive predictive value; DHU: dihydrouracil; U: uracil; NA: not applicable; DOB_{50}: delta-over-baseline ratio at 50 minutes.

The goal of this study was to explore the clinical value of DPD phenotyping assays to identify DPD deficient patients with an increased risk for severe fluoropyrimidine-induced toxicity.

Previously, high endogenous uracil levels have been associated to the onset of severe toxicity.[20] In our study, there was no difference between the mean endogenous DHU/U ratio or mean endogenous uracil levels between wild-type patients for four *DPYD* variants, who experienced severe toxicity or not. Possibly, when including *DPYD* variant allele carriers a difference would have been visible. Yet, we aimed to identify DPD deficient patients in addition to *DPYD* variant allele carriers who are DPD deficient. In terms of clinical validity parameters, our results for the endogenous uracil levels (sensitivity 21%, specificity 82%, NPV 77%, PPV 27%) were only slightly different from previously published parameters to predict severe fluoropyrimidine-induced toxicity (sensitivity 18%, specificity 95%, NPV 90%, PPV 35%).[20] Taking the limited number of patients for two out of four phenotyping assays into account, none of the phenotyping assays investigated in this study showed a combination of both high PPV and NPV parameters, which could predict the clinical value of a test. Of note, sensitivity and PPV of an assay will remain limited even though there is a high odds ratio, if e.g. adverse events are frequent and deficient patients are rare.[32] This is also the case for DPD deficiency and severe fluoropyrimidine-induced toxicity, and we identified low clinical validity parameters.

Our study is the first study in which several phenotyping assays were compared head-to-head in the same patients. For two out of four assays (endogenous DHU/U ratio and endogenous uracil levels), we recruited over 1,000 patients representative of routine clinical care patients. However, our study has some limitations. The 92 patients who underwent three or four phenotyping assays were a little younger compared to patients from the main study cohort, possibly due to the higher patient burden to participate in multiple DPD phenotyping assays. However, we feel this difference is not clinically relevant and it did not influence the occurrence of severe fluoropyrimidine-induced toxicity in these patients.

Secondly, we identified variation in the results of the phenotyping assays, either possibly caused by differences between centres or baseline characteristics (i.e. age, gender or BSA). Per assay, we have examined divergent results and we have corrected for these variations by excluding patients. While we now attributed the identified variation to differences between study centres, these divergent results might also be caused by already existing fluctuation in phenotyping results due to the character of the assay and measurement method. In addition, variation in the clock time of sampling may have affected uracil levels as the metabolizing enzyme DPD shows significant circadian variation.[11] Variation in phenotyping results might also be caused by a different distribution of DPD deficient patients between centres.

Furthermore, predefined cut-off values per phenotyping assay derived from literature were used to be able to divide patients in DPD deficient and non-DPD deficient patients and calculate clinical validity parameters. Cut-off values are also necessary for clinical use, as it would be difficult to determine which patients would require an initial dose reduction without the use of cut-off values. However, the use of cut-off values limits the interpretation of the data of each phenotyping assay. In addition, DPD deficiency itself does not follow a cut-off at a certain point as its severity varies in gradation between completely DPD deficient, partially DPD deficient or non-DPD deficient. Therefore, it would be better not to use cut-off

values in DPD phenotyping assays, but to analyse the results without categorizing patients.

We discussed differences in DPD phenotyping assays, yet the endpoint toxicity can also influence the results. Variation in the outcome of severe toxicity might be caused by different types of treatment regimens between patients, which we did not correct for. As explained for DPD deficiency, analysing the data by categorizing patients also applies to the categorization of toxicity into severe (grade 3-5) and non-severe (grade 0-2) toxicity, where grade 2 toxicity is a grey area in the assessment of toxicity.

Despite our unique data set, we were unable to show that any of the phenotyping assays was associated with DPD deficiency or the onset of severe fluoropyrimidine-induced toxicity very well. The latter is possibly due to the fact that only ~30-50% of severe fluoropyrimidine-induced toxicity can initially be explained by DPD deficiency.[33] Previously it was described that clinical validity and utility were not yet determined for all phenotyping assays,[29] yet with this study we were unable to fully complement this lack of evidence.

Conclusion

We compared four DPD phenotyping assays (the endogenous dihydrouracil/uracil (DHU/U) ratio, endogenous uracil levels, the oral uracil loading dose, and the 2-^{13}C-uracil breath test) in a first-time head-to-head comparison study. None of the phenotyping assays were associated with DPD deficiency or the onset of severe fluoropyrimidine-induced toxicity very well. In order to determine the clinical value of DPD phenotyping assays additional research is required.

References

1. Rosmarin D, Palles C, Pagnamenta A, et al. A candidate gene study of capecitabine-related toxicity in colorectal cancer identifies new toxicity variants at *DPYD* and a putative role for ENOSF1 rather than TYMS. *Gut.* 2015;64(1):111-120.
2. Saltz LB, Cox JV, Blanke C, et al. Irinotecan plus fluorouracil and leucovorin for metastatic colorectal cancer. Irinotecan Study Group. *N Engl J Med.* 2000;343(13):905-914.
3. Scrip's Cancer Chemotherapy Report. *Scrip world pharmaceutical news London: PJB Publications Ltd.* 2002.
4. Rosmarin D, Palles C, Church D, et al. Genetic markers of toxicity from capecitabine and other fluorouracil-based regimens: investigation in the QUASAR2 study, systematic review, and meta-analysis. *J Clin Oncol.* 2014;32(10):1031-1039.
5. Terrazzino S, Cargnin S, Del Re M, Danesi R, Canonico PL, Genazzani AA. *DPYD* IVS14+1G>A and 2846A>T genotyping for the prediction of severe fluoropyrimidine-related toxicity: a meta-analysis. *Pharmacogenomics.* 2013;14(11):1255-1272.
6. Van Kuilenburg ABP, Meinsma R, Zoetekouw L, Van Gennip AH. Increased risk of grade IV neutropenia after administration of 5-fluorouracil due to a dihydropyrimidine dehydrogenase deficiency: High prevalence of the IVS14+1G>A mutation. *International Journal of Cancer.* 2002;101(3):253-258.
7. Henricks LM, Lunenburg CATC, de Man FM, et al. *DPYD* genotype-guided dose individualisation of fluoropyrimidine therapy in patients with cancer: a prospective safety analysis. *Lancet Oncol.* 2018;19(11):1459-1467.
8. Deenen MJ, Meulendijks D, Cats A, et al. Upfront Genotyping of *DPYD**2A to Individualize Fluoropyrimidine Therapy: A Safety and Cost Analysis. *J Clin Oncol.* 2016;34(3):227-234.
9. Van Kuilenburg ABP, Van Lenthe H, Van Gennip AH. Activity of pyrimidine degradation enzymes in normal tissues. *Nucleosides Nucleotides Nucleic Acids.* 2006;25(9-11):1211-1214.
10. Van Kuilenburg ABP, Van Lenthe H, Tromp A, Veltman PC, Van Gennip AH. Pitfalls in the diagnosis of patients with a partial dihydropyrimidine dehydrogenase deficiency. *Clin Chem.* 2000;46(1):9-17.
11. Jacobs BAW, Deenen MJ, Pluim D, et al. Pronounced between-subject and circadian variability in thymidylate synthase and dihydropyrimidine dehydrogenase enzyme activity in human volunteers. *Br J Clin Pharmacol.* 2016;82(3):706-716.
12. Grem JL, Yee LK, Venzon DJ, Takimoto CH, Allegra CJ. Inter- and intraindividual variation in dihydropyrimidine dehydrogenase activity in peripheral blood mononuclear cells. *Cancer Chemother Pharmacol.* 1997;40(2):117-125.
13. Mueller F, Buchel B, Koberle D, et al. Gender-specific elimination of continuous-infusional 5-fluorouracil in patients with gastrointestinal malignancies: results from a prospective population pharmacokinetic study. *Cancer Chemother Pharmacol.* 2013;71(2):361-370.
14. Gamelin E, Boissdron-Celle M, Guerin-Meyer V, et al. Correlation between uracil and dihydrouracil plasma ratio, fluorouracil (5-FU) pharmacokinetic parameters, and tolerance in patients with advanced colorectal cancer: A potential interest for predicting 5-FU toxicity and determining optimal 5-FU dosage. *J Clin Oncol.* 1999;17(4):1105-1110.

15. Zhou ZW, Wang GQ, Wan de S, et al. The dihydrouracil/uracil ratios in plasma and toxicities of 5-fluorouracil-based adjuvant chemotherapy in colorectal cancer patients. *Chemotherapy.* 2007;53(2):127-131.
16. Jiang H, Lu J, Jiang J, Hu P. Important role of the dihydrouracil/uracil ratio in marked interpatient variations of fluoropyrimidine pharmacokinetics and pharmacodynamics. *J Clin Pharmacol.* 2004;44(11):1260-1272.
17. Boisdron-Celle M, Remaud G, Traore S, et al. 5-Fluorouracil-related severe toxicity: a comparison of different methods for the pretherapeutic detection of dihydropyrimidine dehydrogenase deficiency. *Cancer letters.* 2007;249(2):271-282.
18. Kristensen MH, Pedersen P, Mejer J. The value of dihydrouracil/uracil plasma ratios in predicting 5-fluorouracilrelated toxicity in colorectal cancer patients. *J Int Med Res.* 2010;38(4):1313-1323.
19. Wettergren Y, Carlsson G, Odin E, Gustavsson B. Pretherapeutic uracil and dihydrouracil levels of colorectal cancer patients are associated with sex and toxic side effects during adjuvant 5-fluorouracil-based chemotherapy. *Cancer.* 2012;118(11):2935-2943.
20. Meulendijks D, Henricks LM, Jacobs BAW, et al. Pretreatment serum uracil concentration as a predictor of severe and fatal fluoropyrimidine-associated toxicity. *Br J Cancer.* 2017;116(11):1415-1424.
21. van Staveren MC, van Kuilenburg ABP, Guchelaar HJ, et al. Evaluation of an oral uracil loading test to identify DPD-deficient patients using a limited sampling strategy. *Br J Clin Pharmacol.* 2016;81(3):553-561.
22. van Staveren MC, Theeuwes-Oonk B, Guchelaar HJ, Van Kuilenburg ABP, Maring JG. Pharmacokinetics of orally administered uracil in healthy volunteers and in DPD-deficient patients, a possible tool for screening of DPD deficiency. *Cancer Chemother Pharmacol.* 2011;68(6):1611-1617.
23. Mattison LK, Fourie J, Hirao Y, et al. The uracil breath test in the assessment of dihydropyrimidine dehydrogenase activity: pharmacokinetic relationship between expired 13CO2 and plasma [2-13C]dihydrouracil. *Clin Cancer Res.* 2006;12(2):549-555.
24. Mattison LK, Fourie J, Desmond RA, Modak A, Saif MW, Diasio RB. Increased prevalence of dihydropyrimidine dehydrogenase deficiency in African-Americans compared with Caucasians. *Clinical Cancer Research.* 2006;12(18):5491-5495.
25. Mattison LK, Ezzeldin H, Carpenter M, Modak A, Johnson MR, Diasio RB. Rapid identification of dihydropyrimidine dehydrogenase deficiency by using a novel 2-13C-uracil breath test. *Clin Cancer Res.* 2004;10(8):2652-2658.
26. Thomas HR, Ezzeldin HH, Guarcello V, Mattison LK, Fridley BL, Diasio RB. Genetic regulation of dihydropyrimidinase and its possible implication in altered uracil catabolism. *Pharmacogenet Genomics.* 2007;17(11):973-987.
27. Saif MW, Syrigos K, Mehra R, Mattison LK, Diasio RB. Dihydropyrimidine dehydrogenase deficiency (DPD) in GI malignancies: Experience of 4-years. *Pak J Med Sci.* 2007;23(6):832-839.
28. van Staveren MC, Guchelaar HJ, van Kuilenburg ABP, Gelderblom H, Maring JG. Evaluation of predictive tests for screening for dihydropyrimidine dehydrogenase deficiency. *Pharmacogenomics J.* 2013;13(5):389-395.

29. Meulendijks D, Cats A, Beijnen JH, Schellens JH. Improving safety of fluoropyrimidine chemotherapy by individualizing treatment based on dihydropyrimidine dehydrogenase activity - Ready for clinical practice? *Cancer Treat Rev.* 2016;50:23-34.
30. NCI. National Cancer Institute: Common Terminology Criteria for Adverse Events v4.03. https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_8.5x11.pdf, 5 May 2017.
31. Leyva A, van Groeningen CJ, Kraal I, et al. Phase I and pharmacokinetic studies of high-dose uridine intended for rescue from 5-fluorouracil toxicity. *Cancer Res.* 1984;44(12 Pt 1):5928-5933.
32. Tonk ECM, Gurwitz D, Maitland-van der Zee AH, Janssens A. Assessment of pharmacogenetic tests: presenting measures of clinical validity and potential population impact in association studies. *Pharmacogenomics J.* 2017;17(4):386-392.
33. Hsiao H-H, Lin S-F. Pharmacogenetic syndrome of dihydropyrimidine dehydrogenase deficiency. *Current Pharmacogenomics.* 2007;5(1):31-38.

SUPPLEMENT CHAPTER 10

Comparison of four phenotyping assays for predicting dihydropyrimidine dehydrogenase (DPD) deficiency and severe fluoropyrimidine-induced toxicity: a clinical study

Manuscript in preparation

Carin A.T.C. Lunenburg*, Linda M. Henricks*, Femke M. de Man*, Didier Meulendijks, André B.P. van Kuilenburg, Jan Gerard Maring, Maurice C. van Staveren, Niels de Vries, Hilde Rosing, Jos H. Beijnen, Dick Pluim, Anil Modak, Alexander L.T. Imholz, Jan H.M. Schellens, Jesse J. Swen, Hans Gelderblom, Annemieke Cats, Ron H.J. Mathijssen, Henk-Jan Guchelaar
**Contributed equally*

Supplementary material

Calculation of sample size

Within *DPYD* wild-type patients a variability in DPD enzyme activity exists. We assumed that 95% of the *DPYD* wild-type patients would be classified as having normal enzyme activity and 5% of the *DPYD* wild-type patients would be classified with a low DPD enzyme activity (DPD deficient), with an increased risk of toxicity. This results in an unequal sample size, therefore a total sample size of 240 evaluable patients was required to achieve at least 80% power at significance level α=0.05 to detect an increase in the probability of toxicity from an estimated 20% in non-DPD deficient patients to 60% in DPD deficient patients.

Methods assays

DPD enzyme activity[1,2]

The DPD enzyme activity in PBMCs was determined using a validated radio-assay, which is based on conversion of the radiolabelled probe 4-^{14}C thymine to 4-^{14}C dihydrothymine.[1] As this method is considered the gold standard in DPD measurements in the Netherlands, four phenotyping assays were correlated to this assay. Between 8 and 9 am, after overnight fasting, 20 ml blood (EDTA tube) was drawn, combined with a blood draw for determining the endogenous DHU/U ratio. Depending on the hospital of inclusion (*N*=4), whole blood was either shipped overnight to the Academic Medical Center in Amsterdam for further processing, or was processed at the hospital of blood draw (*N*=3) as described before, to isolate PBMCs.[1] After processing, isolated PMBCs were kept at -80°C before measurement of DPD activity at the Academic Medical Center in Amsterdam.

Endogenous DHU/U ratio and endogenous uracil levels[3,4]

The uracil and DHU levels were determined in plasma using a validated ultra-performance liquid chromatography tandem mass-spectrometry (UPLC-MS/MS) method.[4] All samples were measured at the Netherlands Cancer Institute in Amsterdam. In patients who participated in three or four DPD phenotyping assays, a 4 ml blood (heparin tube) was drawn between 8 and 9 am, after overnight fasting, and centrifuged at 4°C at 1500g for 10 minutes. Plasma was kept at -80°C until measurement. In patients from the clinical trial, blood to determine uracil and DHU levels could be drawn throughout the day and in non-fasting state, but information was collected on how long before the blood draw the patient had eaten a meal, as food status could influence the uracil levels in patients.[5]

Oral uracil loading dose[6,7]

Previously, a loading dose of 500 mg/m^2 uracil was used in this assay. To increase feasibility, a standardized dose of 1,000 mg uracil was administered. Patients had to fast overnight for a minimum of eight hours. Food and drinks had to be abstained for the duration of the assay. Uracil was dissolved in warm water and administered between 8 and 9 am, to minimize effects of circadian rhythm. Four ml blood (EDTA tube) was taken at 60 and 120 minutes after oral intake of uracil. Sample processing consisted of adding 0.15 ml of the DPD inhibitor gimeracil to a 4 ml sample and centrifuging at 4°C at 1,500g for ten minutes. Plasma was kept at -80°C until measurement. Uracil and its metabolite dihydrouracil were determined in

plasma using a high-performance liquid chromatography ultra-violet (HPLC-UV) method in the laboratory of the Department of Pharmacy at the Scheper Hospital in Emmen.

2-^{13}C-uracil breath test[8-10]

A personalized dose of 6 mg/kg 2-^{13}C uracil was administered to patients after overnight fasting (minimum eight hours) and alcohol abstaining (minimum 24 hours). Food and drinks had to be abstained for the duration of the assay as well. The 2-^{13}C uracil was dissolved in hot water and administered between 8 and 9 am, to minimize effects of circadian rhythm. Just prior to the administration of the 2-^{13}C uracil solution the patients had to ingest two tablets of Alka-Seltzer Gold® (containing anhydrous citric acid, potassium bicarbonate and sodium bicarbonate) with water, to stimulate uniform and fast absorption of the 2-^{13}C uracil solution. Breath samples (300 ml in a Otsuka Pharmaceuticals breath bag, Japan®) and blood samples (4 ml in a heparin tube) were taken pre-dose and 50 minutes after administration of uracil. Blood samples were centrifuged immediately at 4°C at 1,500g for ten minutes. Plasma was kept at -80°C until analysis. Quantification of ^{13}C-uracil and ^{13}C-dihydrouracil levels was done using the same UPLC-MS/MS method as for the endogenous DHU/U ratio at the Netherlands Cancer Institute, Amsterdam, but with uracil-$^{13}C_4$,$^{15}N_2$ and dihydrouracil-$^{13}C_4$,$^{15}N_2$ as internal standards. $^{13}CO_2$ and $^{12}CO_2$ concentrations were determined in the exhaled breath samples by infrared spectrometry using the FDA approved POCone IR spectrometer (Photal Electronics, Japan®) at the laboratory of the department of Clinical Pharmacy and Toxicology at the Leiden University Medical Center or at the Division of Pharmacology at the Netherlands Cancer Institute, Amsterdam. A delta-over-baseline (DOB) ratio at 50 minutes was calculated that represents a change in the $^{13}CO_2/^{12}CO_2$ ratio of two breath samples.

Supplementary results

Supplementary Table 1. Toxicity data of patients who underwent three or four DPD phenotyping assays and the main study cohort

Type of event	Phenotyping assays (*N*=92)	Main study cohort (*N*=1,103)	P-value[a]
Overall grade ≥3 toxicity	19 (21%)	264 (24%)	0.477
Grade ≥3 gastrointestinal toxicity	6 (7%)	103 (9%)	0.367
Grade ≥3 hematological toxicity	10 (11%)	78 (7%)	0.180
Grade 3 hand-foot syndrome	4 (4%)	37 (3%)	0.389
Grade ≥3 cardiological toxicity	0	10 (1%)	0.447
Grade ≥3 other treatment-related toxicity	3 (3%)	87 (8%)	0.106
Fluoropyrimidine-related hospitalization	7 (8%)	156 (14%)	0.079
Stop of fluoropyrimidines due to adverse events	20 (22%)	190 (17%)	0.133
Fluoropyrimidine-related death	0	2 (0%)	ND

[a] All p-values represent a comparison of 92 patients who underwent three or four DPD phenotyping assays to patients from the main study cohort. We used χ^2 test or Fisher exact test.

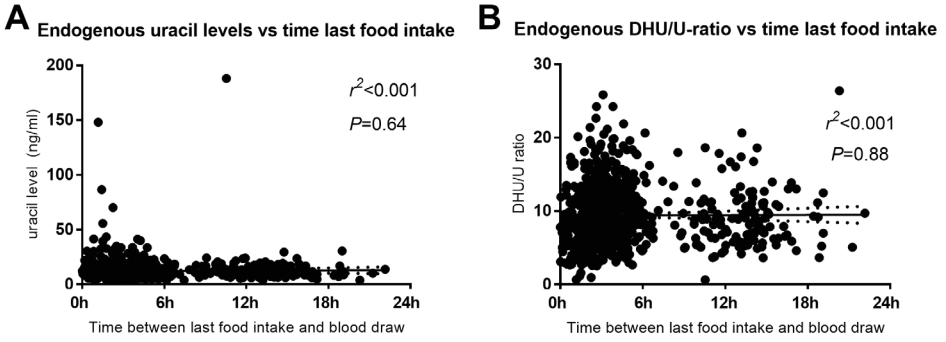
Abbreviations: ND: not done.

Supplementary Table 2. Comparison of phenotyping assays in performance for prediction of severe fluoropyrimidine-induced toxicity

Clinical validity parameters for the prediction of severe fluoropyrimidine-induced toxicity are shown for the 2-^{13}C-uracil breath test and uracil loading assay. *DPYD* variant allele carriers were excluded, since *DPYD* variant allele carriers received an initial dose reduction based on their genotype, and therefore bias could develop in the onset of severe fluoropyrimidine-induced toxicity.

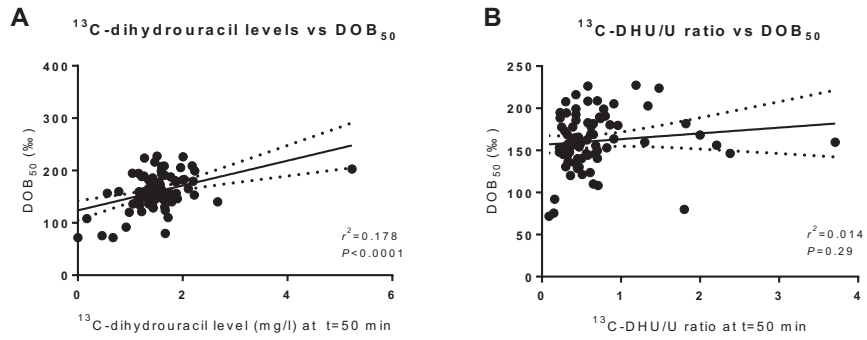
Assay	*N* of patients	Median (IQR)	Cut-off for DPD deficiency	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)	F1-score[a] (%)
2-^{13}C-uracil breath test	74	DOB_{50}: 161 (145.6-186.3) ‰	DOB_{50} ≤128.9 ‰[8,9]	27	89	88	30	29
Oral uracil loading dose	82	U/DHU ratio at 120 min: 0.58 (0.31-1.09)	U/DHU-ratio at 120 min ≥2.4[6]	7	97	82	33	11

[a] The F1-score represents the harmonic mean of sensitivity and PPV.



Supplementary Figure 1. Endogenous uracil levels and endogenous DHU/U ratio plotted against the time between the blood draw and last food intake

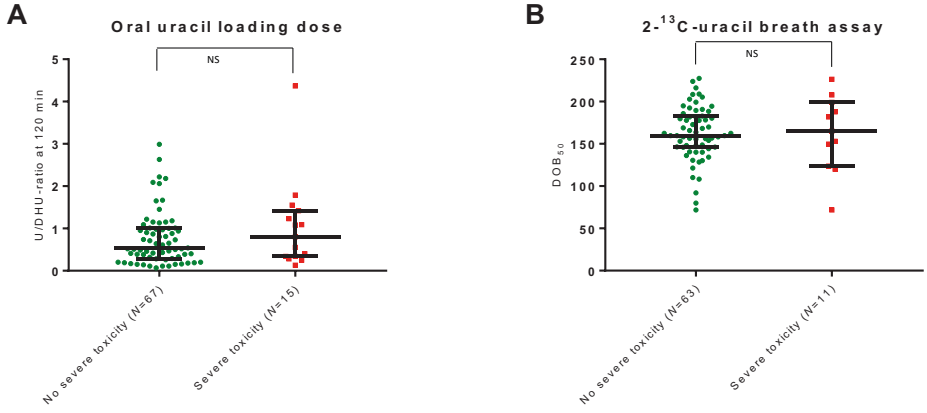
Abbreviations: DHU: dihydrouracil; U: uracil; vs: versus.



Supplementary Figure 2. Correlation between breath samples and plasma samples of the 2-^{13}C-uracil breath test

The association between plasma samples (measured ^{13}C-DHU/U-ratio and ^{13}C-uracil levels at 50 minutes) and breath samples (calculated as DOB_{50}) of the 2-^{13}C-uracil breath test was evaluated by estimating Pearson's correlations coefficients.

Abbreviations: vs: versus; DOB_{50}: delta-over-baseline ratio at 50 minutes; DHU: dihydrouracil; U: uracil.



Supplementary Figure 3. Results of phenotyping assays separated by the occurrence of severe fluoropyrimidine-induced toxicity

Dots represent individual results. Black lines represent the median and 25th and 75th percentile of the data. All *DPYD* variant allele carriers were excluded from the analysis as they received initial dose reductions based on their genotype result.

Abbreviations: U: uracil; DHU: dihydrouracil; DOB_{50}: delta-over-baseline ratio at 50 minutes; NS: not significant p-value.

References

1. Van Kuilenburg ABP, Van Lenthe H, Tromp A, Veltman PC, Van Gennip AH. Pitfalls in the diagnosis of patients with a partial dihydropyrimidine dehydrogenase deficiency. *Clin Chem.* 2000;46(1):9-17.
2. Van Kuilenburg ABP, Meinsma R, Zoetekouw L, Van Gennip AH. Increased risk of grade IV neutropenia after administration of 5-fluorouracil due to a dihydropyrimidine dehydrogenase deficiency: High prevalence of the IVS14+1G>A mutation. *International Journal of Cancer.* 2002;101(3):253-258.
3. Meulendijks D, Henricks LM, Jacobs BAW, et al. Pretreatment serum uracil concentration as a predictor of severe and fatal fluoropyrimidine-associated toxicity. *Br J Cancer.* 2017;116(11):1415-1424.
4. Jacobs BAW, Rosing H, de Vries N, et al. Development and validation of a rapid and sensitive UPLC-MS/MS method for determination of uracil and dihydrouracil in human plasma. *J Pharm Biomed Anal.* 2016;126:75-82.
5. Henricks LM, Jacobs BAW, Meulendijks D, et al. Food-effect study on uracil and dihydrouracil plasma levels as marker for dihydropyrimidine dehydrogenase activity in human volunteers. *Br J Clin Pharmacol.* 2018.
6. van Staveren MC, van Kuilenburg ABP, Guchelaar HJ, et al. Evaluation of an oral uracil loading test to identify DPD-deficient patients using a limited sampling strategy. *Br J Clin Pharmacol.* 2016;81(3):553-561.
7. van Staveren MC, Theeuwes-Oonk B, Guchelaar HJ, Van Kuilenburg ABP, Maring JG. Pharmacokinetics of orally administered uracil in healthy volunteers and in DPD-deficient patients, a possible tool for screening of DPD deficiency. *Cancer Chemother Pharmacol.* 2011;68(6):1611-1617.
8. Mattison LK, Fourie J, Hirao Y, et al. The uracil breath test in the assessment of dihydropyrimidine dehydrogenase activity: pharmacokinetic relationship between expired 13CO2 and plasma [2-13C]dihydrouracil. *Clin Cancer Res.* 2006;12(2):549-555.
9. Mattison LK, Ezzeldin H, Carpenter M, Modak A, Johnson MR, Diasio RB. Rapid identification of dihydropyrimidine dehydrogenase deficiency by using a novel 2-13C-uracil breath test. *Clin Cancer Res.* 2004;10(8):2652-2658.
10. Cunha-Junior GF, De Marco L, Bastos-Rodrigues L, et al. 13C-uracil breath test to predict 5-fluorouracil toxicity in gastrointestinal cancer patients. *Cancer Chemother Pharmacol.* 2013;72(6):1273-1282.