Dose-Response:
An International Journal
July-September 2020:1-9

DOI: 10.1177/1559325820951367
journals.sagepub.com/home/dos
SAGE

Comparison of 4 Screening Methods for Detecting Fluoropyrimidine Toxicity Risk: Identification of the Most Effective, Cost-Efficient Method to Save Lives

Olivier Capitain[1], Valérie Seegers[1], Jean-Philippe Metges[2], Roger Faroux[3], Claire Stampfli[4], Marc Ferec[5], Tamara Matysiak Budnik[6], Hélène Senellart[1], Valérie Rossi[7], Nadège Blouin[1], Jonathan Dauvé[1], and Mario Campone[1]

Abstract

Background: Fluoropyrimidines (FPs) carry around 20% risk of G3-5 toxicity and 0.2-1% risk of death, due to dihydropyrimidine dehydrogenase (DPD) deficiency. Several screening approaches exist for predicting toxicity, however there is ongoing debate over which method is best. This study compares 4 screening approaches.

Method: 472 patients treated for colorectal, head-and-neck, breast, or pancreatic cancers, who had not been tested pre-treatment for FP toxicity risk, were screened using: *DPYD* genotyping (G); phenotyping via plasma Uracil (U); phenotyping via plasma-dihydrouracil/uracil ratio (UH_2/U); and a Multi-Parametric Method (MPM) using genotype, phenotype, and epigenetic data. Performance was compared, particularly the inability to detect at-risk patients (false negatives).

Results: False negative rates for detecting G5 toxicity risk were 51.2%, 19.5%, 9.8% and 2.4%, for G, U, UH_2/U and MPM, respectively. False negative rates for detecting G4-5 toxicity risk were 59.8%, 36.1%, 21.3% and 4.7%, respectively. MPM demonstrated significantly ($p < 0.001$) better prediction performance.

Conclusion: MPM is the most effective method for limiting G4-5 toxicity. Its systematic implementation is cost-effective and significantly improves the risk-benefit ratio of FP-treatment. The use of MPM, rather than G or U testing, would avoid nearly 8,000 FP-related deaths per year globally (500 in France), and spare hundreds of thousands from G4 toxicity.

Keywords
fluoropyrimidines, DPD deficiency, toxicity, risk assessment, comparison, screening

Introduction

Fluoropyrimidines (FPs), including 5-fluorouracil (5-FU) and its pro-drug capecitabine (CAP) are the most widely used chemotherapeutic agents for many types of solid tumors.[1] Globally, around 1,600,000 patients receive FP treatment per year[2,3] ($\sim$180 K in the US, and $\sim$450 K in the EU). FPs have a narrow therapeutic window with a small difference between the minimum efficacious dose and the maximum tolerable dose, both of which are variable from patient to patient. Thus, while they can be highly effective in treating cancer, severe, sometimes fatal, toxicities occur in a significant minority of patients, which represents an important public health problem.[4] The toxic adverse drug reactions (ADRs) typically manifest 2 to 7 days into the first or second

[1] Institut de Cancérologie de l'Ouest, Angers, France
[2] CHU Hôpital Morvan, Brest, France
[3] CH Départemental Vendée La Roche sur Yon, France
[4] CH Laval, France
[5] CH Pays de Morlaix, Morlaix, France
[6] CHU Hôtel Dieu, Nantes, France
[7] CH Château Gontier, France.

Received 28 April 2020; received revised 3 July 2020; accepted 27 July 2020

Corresponding Author:
Jonathan Dauvé, ICO Paul Papin, 15 rue André Boquel, CS 10059, 49055 ANGERS Cedex 02, France.
Email: Jonathan.Dauve@ico.unicancer.fr

course of chemotherapy, with one or more severe symptoms including mucositis, stomatitis, diarrhea, neutropenia, myelosuppression, skin toxicities (hand and foot syndrome), and at a later stage, leucoencephalitis.[5] The management of these toxicities usually requires an extended stay in intensive care. In the most severe cases, multisystem organ failure occurs, despite intensive care, leading to dehydration, renal failure, and death.[6]

Severe and very severe toxicities (grades 3 and 4 on the US National Cancer Institute's Common Terminology Criteria for Adverse Events (NCI-CTCAE) scale[7]) are reported in 10 to 30% of patients treated with standard doses of FPs, with lethal outcomes (grade 5 toxicity) in 0.2 to 1% of cases.[3,8,9] Even with moderate doses, 3%–5% of patients experience early-onset grade 3 or 4 ADRs, with lethal outcomes in 0.2 to 0.5% of cases.[3] Thus, on a global scale, we estimate that FP treatment is responsible for between 3,200 to 16,000 deaths per year, and that serious, debilitating grade 3 or 4 ADRs occur in up to 320,000 patients per year.

Several pre-treatment screening approaches are used to reduce the risk of FP toxicity, however there is considerable debate regarding the optimal method.[10,11]

Historically, the initial FP dose has been based on the patient's body surface area (BSA), and this approach is still used in many parts of the world. However, this method is known to be inadequate because 43% of patients are not given the correct dose; 33% are underdosed (significantly reduced efficacy) and 10% are overdosed (running a strong risk of very severe early-onset toxicity).[12] This is because it does not consider other information such as the patient's FP metabolic capacity, or the patient's personal characteristics such as age or gender.[13]

Indeed, significantly greater risk of FP toxicity is frequently reported in females and elderly patients, as compared to males and younger people.[13-15] In addition, the patient's FP metabolic capacity is of key importance, and is linked to the initial rate-limiting metabolic step performed by the enzyme dihydropyrimidine dehydrogenase (DPD).[3] DPD converts more than 80% of a 5-FU dose into the inactive metabolite 5-fluoro-5,6- dihydrouracil (5-FDHU).[3] Remarkably, DPD enzyme activity is subject to a wide inter-individual variability, due in part to genetic polymorphisms of the *DPYD* gene coding for DPD.[16] Approximately 3%–10% of the entire population demonstrates partial DPD deficiency and 0.1%–0.5% demonstrate complete deficiency leading to a total inability to detoxify FPs.[17,18] More than 100 single-nucleotide polymorphisms (SNPs) of *DPYD* alleles have been reported.[6,19-21] Only a minority of them have a potentially deleterious impact on DPD enzyme activity, however 5 specific SNPs (*DPYD**2A, *DPYD**7, *DPYD**9B, *DPYD**13 and *HapB3*) are most often implicated in clinically-relevant DPD deficiency, and are present with sufficient population frequency to make their detection useful via genotyping (G).[3,20-24] However not all severe ADRs can be explained by polymorphisms of these 5 *DPYD* gene mutations.[23,24] Thus, genotyping is useful for avoiding toxicity in some patients, but it cannot robustly predict all patients that are at risk of severe FP toxicity.[3,11,23] Other available pre-treatment screening approaches rely on phenotypic testing, to either detect high plasma Uracil (U) levels,[24,25] or to determine the ratio of dihydrouracil (UH_2) to U in plasma (UH_2/U).[3,26] Phenotypic testing is an indirect measure of a patient's metabolic competence, and is a recognized approach for limiting the risk of FP toxicity. However, phenotypic testing suffers from a lack of comprehensive studies addressing its sensitivity and specificity.[11] In addition, the quality of the analytical result is influenced by the need for stringent sample handling conditions and possible biases related to circadian variations and/or food intake.[11]

Another pre-treatment screening approach, known as the Multi-Parametric Method (MPM),[3] generates a prediction of FP toxicity risk by combining a patient's genotypic and phenotypic results with the patient's personal characteristics (sex, age, concomitant treatments, etc). This approach is in routine clinical practice in some parts of France.

Given the current lack of consensus on the best pre-treatment screening method for accurately predicting the risk of FP toxicity, herein we assess and critically compare the performance of the 4 screening approaches outlined above, using real-life data from the same patient population.

Materials & Methods

Patients and Treatments

Between July 2000 and July 2018, blood samples from 472 patients undergoing 1st or 2nd round FP treatment were sent for analysis to the Oncopharmacology Laboratory of the Integrated Center for Oncology, Angers, France (ICO Laboratory), either because i) no screening had been done and the patient subsequently experienced $\geq$ grade 3 ADR (263 patients), or ii) FP treatment was started with a standard dose and samples were sent in to the lab for testing in parallel (209 patients).

The 472 patients had various types of cancer and were treated with a range of drug regimens. Colorectal cancer patients received FPs using the protocols LV5FU2, FOLFIRI, or FOLFOX over 46 hours, with or without EGFR or VEGF monoclonal antibodies. Breast cancer patients were treated by either FEC50, FEC100, or capecitabine. Head & neck cancer patients were treated by continuous infusion of 5-fluorouracil over 96 hours along with cisplatin.

Toxicity Evaluation

The toxicity evaluations were performed by qualified clinicians according to NCI-CTCAE Version 5 scale, where 0 = no toxicity observed and 5 = death.[27]

Blood Samples and Patient Data

Blood samples were collected by qualified personnel and consisted of two 5 mL samples in lithium-heparin tubes. One tube was centrifuged (3000 rpm, 10 min) within 1 hour of collection and the plasma was then decanted and immediately frozen at -20°C. This tube was used for phenotypic analyses. The second tube was kept at ambient temperature and was used for genotypic analysis. All tubes were shipped to the ICO

Laboratory with strict respect to their appropriate storage temperatures.

Upon arrival of samples at the ICO Laboratory, the following patient details were recorded in the laboratory's computer system: patient identity, weight, height, sex, age, type of cancer, chemotherapy treatment (drugs, administration route, dose, treatment length), timing of the sampling (before or after start of treatment, before or after appearance of ADRs), and the recorded NCI-CTCAE toxicity grade (if any).

Study Design

All blood/plasma samples from the 472 patients underwent the following 4 screening tests at the ICO laboratory:

1. *DPYD* genotype mutations (*2A, *7, *9B, *13, HapB3)
2. Plasma Uracil (U) level (ng/mL)
3. Metabolic Index: dihydrouracil/uracil ratio (UH_2/U)
4. Multi-Parametric Method (MPM), which integrates genotyping, U plasma concentration, UH_2 plasma concentration, UH_2/U ratio, and key patient factors (age, sex, treatment, etc.).

The performances of the 4 tests were evaluated and compared with respect to their ability to accurately predict $\geq$ grade 3 toxicity in terms of: False Positives, False Negatives, Sensitivity, Specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), Positive Likelihood Ratio (LR+), Negative Likelihood Ratio (LR-), and Diagnostic Odds Ratio[28] (DOR) = (LR+/LR-).

As per Delacour et al (2009)[29] the clinical diagnostic strength of a test is a function of its LR+, LR- and DOR values.

Very strong strength if: LR+ > 10, LR- < 0.1

Strong strength if: 5 < LR+ < 10, 0.1 < LR- < 0.2

Moderate strength if: 2 < LR+ < 5, 0.2 < LR- < 0.5

Low strength if: 1 < LR+ < 2, 0.5 < LR- < 1

DPYD *Genotype Analysis*

Detection of SNPs was performed using pyrosequencing (PyroMark PSQ96MA, Qiagen, France) according to the method described by Morel et al 2006,[23] which is accredited by the French Comité d'Accréditation (COFRAC 8-3281).

The following SNPs were targeted for detection: *DPYD**2A (rs3918290;IVS14+1G<A), *DPYD**7 (rs72549309; DelTCAT 295-298), *DPYD**9B (rs67376798; D949 V), *DPYD**13 (rs55886062; I560 S) and HapB3 (3 intronic variants c.483+18 G>A-rs56276561, c.680+139 G>A-rs6668296, and c.959-51 T>C-rs115349832).

Phenotype Analyses

Quantification of plasmatic Uracil (U) and dihydrouracil (UH_2) was performed via UPLC with UV detection (210 nm for UH_2 and 260 nm for U) derived from the method described by Remaud et al 2005[30] and Boisdron-Celle et al 2007.[24] The method is accredited by COFRAC 8-3281.

Regarding Plasma Uracil (U) Levels

According to recommendations published in December 2018 by the French Institut National du Cancer (INCa) and the French *Haute Autorité de Santé* (HAS),[31] a plasma Uracil level < 16 ng/ml, is normal and indicates the absence of a DPD enzyme deficit. Plasma Uracil between 16 ng/ml and 150 ng/ml, indicates partial DPD enzyme deficit: the dose of FP should be lower than the standard dose. Plasma Uracil $\geq$ 150 ng/ml indicates total DPD enzyme deficit: FPs should be contra-indicated for the patient.

Plasma U results for the 427 patients in this study were evaluated using these recommendations, to evaluate the pertinence of the HAS-INCa-recommended thresholds for indicating toxicity risk.

Regarding the Metabolic Index (UH_2/U):

Boisdron-Celle et al 2007,[24] demonstrated that the UH_2/U ratio is inversely correlated to the observed toxicity, and that the correlation is highly significant ($p < 0.001$). The following thresholds were established with respect to risk of FP toxicity: If $UH_2/U \geq 6$ the patient is considered non-DPD enzyme deficient. If UH_2/U is between 2 and 6: the patient is considered partially DPD enzyme deficient and a lowered FP dose is proposed based on the UH_2/U value. If $UH_2/U < 2$: patient considered totally DPD enzyme deficient and use of FPs is contra-indicated.

The metabolic index results for the 427 patients in this study were evaluated against these thresholds with respect to their pertinence for indicating FP toxicity risk.

Regarding the Multi-Parametric Method (MPM)

The multi-parametric method (MPM) utilizes the results of each patient's genotype and phenotype analyses, as well as the patient's personal characteristics: age (years), sex, weight (kg), height (cm), intended FP treatment protocol (including concomitant drugs), the patient's Eastern Cooperative Oncology Group (ECOG) score, and the intended duration of the perfusion of fluoropyrimidines (hours). Each patient's overall data sets were analyzed with the mathematical scores provided by 5-FU$^{ODPMTox™}$, version v3.3.1.5 ~ m1,[32] CE marking: CE-DMDIV 03/10/2011 DT11061.2.

The algorithms are composed of 178 mathematical scores which utilize the patient's full data set to:

1. Predict the risk of grade 3 or higher FP toxicity
2. Allow standard FP dosing if $\geq$ grade 3 toxicity is predicted to be unlikely
3. Propose a lower dose of FP, if grade 3 or 4 toxicity with a standard dose is likely
4. Contraindicate treatment with FPs if the patient is at risk of grade 5 toxicity.

The 5-FU$^{ODPMTox™}$ program was developed and validated at the ICO, (Angers, France), between 2000-2013. None of the patient data in this study was used to develop or train the

Table 1. Patient Characteristics.

Total population: N = 472 Characteristic	Number of patients, and (%) of total	Patients with toxicity, grade 4 or 5: N = 169 and % of subset
Sex: Number and (%)		
Male	222 (47%)	66 (39%)
Female	250 (53%)	103 (61%)
Age (years)		
Average ±SD	66.7 ± 12.0	63 ± 12.0
Range (min-Max)	19-88	23-87
Tumor Localization: number and (%)		
Colorectal	337 (71.4%)	98 (57.9%)
Pancreas	9 (1.9%)	3 (1.77%)
Breast	36 (7.6%)	22 (13%)
Head and Neck	33 (6.99%)	19 (11.2%)
Gastric	15 (3.2%)	10 (5.9%)
Other	9 (1.9%)	2 (1.18%)
Not specified	33 (6.99%)	15 (8.9%)
Treatment type: number and (%)		
Adjuvant	64 (13.6%)	22 (13%)
Metastatic	49 (10.4%)	17 (10%)
Neo-adjuvant	4 (0.85%)	3 (1.8%)
Not specified	355 (2.2%)	127 (75%)

program. The latest version (version v3.3.1.5 ~ m1) was validated in 2014 using data from 4859 patients and has been in routine clinical use in France since 2015. The 5-FU$^{ODPMTox™}$ program is in direct liaison with the laboratory's computer system that holds all of the patient data (genotype and phenotype results and the patient's personal characteristics) thus avoiding risk of data transfer error.

Statistical Analysis

Quantitative data is presented as the average, standard deviation, median, 25th and 75th percentiles.

Categorical (binary) data are summarized by their number and percentage for each modality. To compare quantitative data between paired groups (the same samples were analyzed using the 4 different methods), non-parametric tests for paired data were used: Mann-Whitney or Friedman, depending on the groups compared. To compare categorical data, the McNemar test or Cochran's Chi-squared test was used, depending on the number of groups compared.

Reported p-values are the raw p-values adjusted with the Benjamini and Hockberg method[33] to account for the risk of a higher alpha value due to multiple testing on the same samples.

The diagnostic strength of the methods was determined using the epiR package (version 0.9-99).[34] All statistical tests were performed with R (version 3.5.3), from the R Foundation.[35]

Results

Patient Characteristics

The characteristics of the 472 patients are presented in Table 1.

Toxicity Evaluation

The following toxicity grades were recorded for the 472 patients involved in this study:

Grade 0, 1 or 2: 206 patients: 202 After Chemotherapy (AC), 4 After Toxicity (AT)

Grade 3: 97 patients: 1 AC and 96 AT,

Grade 4: 128 patients: 5 AC and 123 AT,

Grade 5: 41 patients: 1 AC and 40 AT.

The most frequent type of grade 4 toxicity was digestive (42.2%), followed by haematological (32.0%) and skin-mucosal (21.9%). Of the patients that died from the ADR (grade 5 toxicity), 43.9% presented multi-organ toxicity.

Genotype Analysis

Table 2 shows the prevalence of *DPYD* mutations, and the genotypic profile of patients who suffered grade 4 or 5 toxicity.

Table 2. Prevalence of *DPYD* Mutations, and Genotypic Profiles of Patients With ≥Grade 3 Toxicity.

Gene mutation	Prevalence of the mutation in the patient population Number & (%) N = 472 patients	Patients with Grade 3 toxicity Number & (%) † N = 97 patients	Patients with Grade 4 toxicity Number & (%) † N = 128 patients	Patients with Grade 5 toxicity Number & (%)† N = 41 patients
DPYD*2A	37 (7.84%)	10 (27.0%)	19 (51.4%)	6 (16.2%)
Heterozygous (h)	3 (0.63%)	0	1 (33.0%)	2 (66.0%)
Homozygous (H)				
hDPYD*9B	42 (8.89%)	8 (19%)	22 (52.4%)	8 (19.0%)
hDPYD*7	3 (0.63%)	0	2 (66.0%)	1 (33.0%)
hDPYD*13	3 (0.63%)	1 (33.33%)	2 (66.0%)	0
hDPYD*2A+ hDPYD*9B	1 (0.21%)	0	0	1 (100%)
hDPYD*2A+ hDPYD*13	1 (0.21%)	0	0	1 (100%)
hHapB3	8 (1.69%)	1 (12.5%)	2 (25.0%)	1 (12.5%)
None of the above gene mutations	374 (79.2%)	77 (20.6%)	80 (62.5%)	21 (51.2%)

† The percentage value is the proportion of patients with the mutation that demonstrated toxicity of the stated grade.
h = heterozygous

Table 3. Percentage of Patients With Grade ≥3 Toxicity as a Function of Their Plasma U Level. Median, 25th et 75th Percentile (Pctl).

	Number of patients (of 472)	Patients with toxicity grade 3 number (%)	Patients with toxicity grade 4 number (%)	Patients with toxicity grade 5 number (%)
U < 16ng/mL	280	47 (16.8%)	53 (18.9%)	8 (2.9%)
Median [U] ng/mL	-	9.97	8.90	12.48
25th Pctl_75th Pctl	-	6.55_13.96	6.3_12.7	11.1_14.8
16ng/ml ≤ U < 150ng/mL	178	50 (28.1%)	73 (41.0%)	21 (11.8%)
Median [U] ng/mL	-	23.60	25.28	33.51
25th Pctl_75th Pctl	-	19.63_31.81	21.0_31.8	25.0_58.8
U ≥ 150 ng/mL	14	0	2 (14.3%)	12 (85.7%)
Median [U] ng/mL	-	-	-	2083
25th Pctl_75th Pctl	-	-	-	417.5_3263.2

"-" = not applicable

Table 4. Percentage of Patients That Suffered Grade ≥3 Toxicity as a Function of Their Metabolic Index (UH_2/U). Median, 25th et 75th Percentile (Pctl).

	Number of patients (of 472)	Patients with toxicity grade 3, number (%)	Patients with toxicity grade 4, number (%)	Patients with toxicity grade 5, number (%)
UH_2/U ≥ 6	267	31 (11.6%)	32 (12.0%)	4 (1.5%)
Median UH_2/U	-	8.76	8.73	6.87
25th Pctl_75th Pctl	-	7.01_9.74	7.3_10	6.54_8.25
2 ≤ UH_2/U < 6	180	66 (36.7%)	88 (48.9%)	20 (11.1%)
Median UH_2/U ± SD	-	4.93	4.40	3.72
25th Pctl_75th Pctl	-	4.26_5.55	3.55_5.09	3.06_4.64
UH_2/U < 2	25	0	8 (32.0%)	17 (68.0%)
Median UH_2/U ± SD	-	-	1.55	0.90
25th Pctl_75th Pctl	-	-	0.95_1.70	0.01_1.37

"-" = not applicable

Phenotype Analysis

Plasma Uracil levels (U). Table 3 displays plasma Uracil (U) levels for patients that suffered ≥ grade 3 toxicity. Twenty-one point eight percent (21.8%) of patients with plasma Uracil < 16ng/ml presented with grade 4 or 5 toxicity, and 2.9% died. Fifty-two-point eight percent (52.8%) of patients with intermediate plasma U level (16ng/ml < U < 150ng/ml) suffered grade 4 to 5 toxicities, of whom 11.8% died. In addition, 28% presented with grade 3 toxicity.

All of the patients with a plasma Uracil level ≥ 150ng/ml presented with ≥ grade 4 toxicity, and 85.7% of them died from their ADR.

Metabolic index (UH_2/U). Table 4 shows the metabolic index results for patients that suffered ≥ grade 3 toxicity.

Thirteen point five percent (13.5%) of patients with UH_2/U ≥ 6 suffered grade 4 or 5 toxicity, and 1.5% died. Sixty percent (60%) of patients with a UH_2/U ratio between 2 and 6 suffered grade 4 or 5 toxicity, and 36.7% suffered grade 3 toxicity.

One hundred percent (100%) of the patients that had a UH_2/U < 2 suffered grade 4 or 5 toxicity, and 68% of them died because of their ADR.

Multi-Parametric Method (MPM):

Table 5 shows the percentage of patients who presented ≥ grade 3 toxicity as a function of their DPD-enzyme deficit status, as determined by the multi-parametric method.

Three point eight percent (3.8%) of patients were evaluated as being non-DPD deficient, yet presented ≥ grade 4 toxicity, and 1 (0.5%) died.

Including grade 3 toxicities, only 5.8% of patients were incorrectly identified as being non-DPD deficient, yet experienced ≥ grade 3 toxicity, using MPM.

For patients evaluated as partially DPD-deficient, 56.9% presented with grade 4-5 toxicity and 38.9% with grade 3.

One hundred percent (100%) of the patients evaluated as totally DPD-deficient presented ≥ grade 4 toxicities, and all 17 deaths could have been avoided via use of MPM.

Comparison of the 4 Different Screening Approaches

The screening approaches were compared for their ability to correctly predict toxicity.

Capacity to correctly detect risk of grade ≥ 3 toxicity. Table 6 compares the performances of the 4 DPD screening approaches to correctly predict ≥ grade 3 toxicity.

Table 5. Percentage of Patients That Presented ≥Grade 3 Toxicity as a Function of Their DPD-Enzyme Deficit Status, as Determined by the Multi-Parametric Method.

	Number of patients (472)	Patients with toxicity grade 3, number (%)	Patients with toxicity grade 4, number (%)	Patients with toxicity grade 5, number (%)
Non-deficient	208	4 (1.9%)	7 (3.4%)	1 (0.5%)
Partially deficient	239	93 (38.9%)	113 (47.3%)	23 (9.6%)
Totally deficient	25	0	8 (32.0%)	17 (68.0%)

Table 6. Comparison of DPD Screening Approaches to Correctly Predict Grade 3 or Higher FP Toxicity.

Statistical property	Genotyping	Phenotyping Uracil ≥ 16 ng/ml	Phenotyping UH_2/U ratio	Multi-parametric method
N	472	472	472	472
False Positives, n (%)	10 (2.1%)	34 (7.2%)	6 (1.3%)	10 (2.1%)
False Negatives, n (%)	178 (36.4%)	108 (22.9%)	67 (14.2%)	12 (2.5%)
Sensitivity (%)	33.08	59.40	74.81	95.49
Specificity (%)	95.15	83.50	97.09	95.15
Positive Predictive Value (%)	89.80	82.29	97.07	96.21
Negative Predictive Value (%)	52.41	61.43	74.91	94.23
Positive Likelihood ratio (LR+)	6.82	3.59	25.68	19.67
Negative Likelihood ratio (LR-)	0.70	0.48	0.26	0.047
Diagnostic Odds Ratio (DOR)	9.69	7.48	98.76	418.51
Clinical Diagnostic Strength	Moderate	Moderate	Moderately Strong	Very Strong
Deaths (Grade 5 Toxicity)	**Genotyping**	**Phenotyping Uracil ≥ 16 ng/ml**	**Phenotyping UH_2/U ratio**	**Multi-Parametric Method**
N	41	41	41	41
Sensitivity, n (%)	20 (48.78%)	33 (80.5%)	37 (90.2%)	40 (97.6%)
False negatives, n (%)	21 (51.2%)	8 (19.5%)	4 (9.8%)	1 (2.4%)

"-" = not applicable

Genotyping (alone) is highly specific (95%) but has insufficient sensitivity because it only detects 33% of patients at risk of ≥ grade 3 toxicity. In other words, 67% of patients at risk of severe toxicity and/or death are not detected.

Measuring plasma Uracil levels (alone), as recommended by HAS-INCa (France), has better sensitivity (59%) than genotypic analysis, but generates a high number of false negatives (22.9%) which leaves patients at risk of severe toxicity and/or death.

With respect to the Likelihood ratios: Genotyping (alone) has an LR+ of 6.8 and an LR- of 0.70 which gives it a moderate clinical diagnostic strength. The same is true of plasma Uracil measurement which has an LR+ of 3.59 and a LR- of 0.48. Both UH_2/U and MPM have strong clinical diagnostic strength with an LR+ > 10, and low LR- values, particularly MPM which has an LR- close to zero.

Only UH_2/U and MPM have a Diagnostic Odds Ratio of > 20, and thus they are the only 2 approaches that can be considered to be sufficiently discriminating for the correct detection of a true risk of severe FP toxicity (≥ grade 3).

Capacity to correctly detect risk of ≥ grade 4 toxicity. Table 7 in the Supplemental section shows the comparison of the 4 different screening approaches to correctly predict toxicity ≥ grade 4.

Genotyping (alone) has insufficient sensitivity because it is only able to detect ∼40% of patients at risk of grade 4 or 5 toxicity.

Measuring plasma Uracil levels (alone), as recommended by HAS-INCa, has better sensitivity (64%) than genotypic analysis, but it generates a high percentage of false negatives (36.1% of patients) who are at risk of very severe toxicity and/or death.

In contrast, the Multiparametric Method shows very high sensitivity (95.3%) and the lowest percentage of false negatives (4.7%) for predicting toxicity ≥ grade 4.

As shown in Table 6, of the 41 patients who died due to ADR to FP treatment (grade 5 toxicity), screening via:

Genotyping, would have permitted 20 lives to be saved.

Plasma Uracil determination would have permitted 33 lives to be saved.

UH_2/U ratio would have permitted 37 lives to be saved.

The Multi-Parametric Method would have permitted 40 lives to be saved.

Discussion

In accordance with the results of Jansman et al,[13] the incidence of grade 4 or 5 toxicity was substantially higher in female patients compared to male patients (61% of the 169 patients that suffered grade 4 or 5 toxicity were female).

Given that genotyping alone has insufficient sensitivity (detection of only 40% of patients at risk of grade 4 or 5 toxicity), 60% of patients at risk of very severe toxicity and/or death are not detected (cf. Tables 2 and 6). The Clinical Diagnostic Strength of genotyping (alone) was found to be low.

Tables 3 and 6 reveal that the plasma Uracil method used alone is a relatively poor method for predicting FP toxicity, because it is prone to a high level of false negative results: 21.8% of patients with a plasma Uracil concentration of < 16ng/ml (who according the INCa-HAS threshold carried little or no risk of severe FP toxicity) suffered ≥ grade 4 toxicity, and 2.9% of them died (Table 3). Furthermore, in the intermediate group (16 ng/ml < U < 150 ng/ml), 19.2% of patients did not present with severe toxicity, and therefore would potentially have been under-dosed if the current dose reduction recommendation of HAS-INCa had been followed. Moreover, 29 of the 458 patients (6.3%) with a plasma Uracil level <150 ng/ml had grade 5 toxicity and died. Thus, the INCa-HAS thresholds of 16 ng/ml for predicting lack of toxicity risk, and 150 ng/ml for predicting high risk of death from a standard FP dose, are not reliable.

DPD deficiency screening via metabolic index (UH_2/U ratio) was found to be a better method for predicting ≥ grade 4 toxicity than plasmatic Uracil (U): Only 13.5% of patients with a UH_2/U ratio ≥ 6 presented grade 4 or 5 toxicity, compared to 21.8% of patients that had plasma Uracil < 16 ng/ml. In the intermediate group (6 < UH_2/U ≤ 2), 96.7% of patients were correctly predicted to have a risk of grade 3 or higher toxicity, leaving just 3.3% of patients who would potentially be underdosed, as compared to 19% of underdosed patients if the plasma Uracil result is used alone. The Clinical Diagnostic Strength of the metabolic index (UH_2/U ratio) approach was found to be moderately strong (its LR- and DOR values narrowly missed the "strong" level).

Of the 4 methods evaluated, the multi-parametric method (MPM) showed the highest sensitivity (95.3%) and the lowest number of false negatives (4.7%) for prediction of ≥ grade 4 toxicity. Tables 6 and 7 illustrate its superior performance compared to other methods. With respect to the patient population in this study, we predict that the use of MPM would have spared an additional 53 patients (out of 169) from very severe grade 4 or 5 toxicity, as compared to the use of the plasma Uracil approach (where 61 "at risk" patients were not detected). Furthermore, we predict that the use of MPM would have saved all but 1 of the 41 deaths, whereas as the plasma Uracil approach (alone) would have failed to protect 8 patients' lives. Overall, the Clinical Diagnostic Strength of MPM was found to be very strong (Table 6). Combining the patient's genotypic and phenotypic results with their personal characteristics is clearly of value in the prediction model.

The MPM used in the current study is already in routine clinical practice in some parts of France and has a proven track-record of saving lives and avoiding debilitating ADRs. The latest version of the $5FU^{ODPMTox™}$ program (version v3.3.1.5∼m1) was validated in 2014/2015 using data from 4859 patients. Within this data set there were just 3 occurrences of grade 5 toxicity, 2 of which occurred because FP dosing was started without having the screening results (i.e. the patients were not screened before treatment began) and one because the doctor chose not to follow the dose reduction suggested by the MPM. These 3 deaths were potentially avoidable if the pre-treatment screening and dose adjustment provided by the program had been followed appropriately.

In comparison, based on the incidence of false negatives given by the plasma Uracil screening approach in the current study, it can be estimated that if the plasma Uracil method had been used to screen these 4859 patients, up to 544 patients would potentially have experienced grade 4 toxicity, and there could have been up to 82 deaths, because these at-risk patients would not have been recommended for dose-adjustment based on their plasma U result (alone).

In our opinion, limiting DPD deficiency screening to one parameter (e.g. plasma U), is not only illusory, but dangerous. As shown herein, there is a high risk of false negatives when employing a genotyping or plasma Uracil method, alone. This leads to a false sense of security for both the clinician and the patient.

Global potential of the MPM to reduce FP-related mortality and grade 4 toxicity. Around 1.6 million patients are treated with FPs per year globally. Assuming a mortality rate of 0.5%, treatment with FPs cause at least 8,000 deaths per year. The false negative rates (i.e. failure to correctly detect patients at risk of grade 5 toxicity) for the plasma Uracil and genotyping approaches (used alone) were 19.5% and 51.2%, respectively (cf. Table 7 in the supplemental section). Thus, if 100% of patients were screened, the false negative results for routine screening using the plasma Uracil or genotyping approach (alone) would lead to at least 1,560 to 4,096 deaths per year globally, due to sub-optimal test performance. The false negative rate for detecting grade 5 toxicity risk with the MPM is 2.4%. Thus, routine use of the MPM test in 100% of patients, would reduce the number of deaths from FP-treatment to just 192 patients per year globally, saving up to 1,368 lives per year compared to the other methods.

Taking a conservative incidence for grade 4 toxicity of 10%, treatment with FPs leads to at least 160,000 grade 4 toxicities per year, globally. The false negative rates found herein for the plasma Uracil and genotyping approaches to predicting risk of grade 4 toxicity were 36.1% and 59.8%, respectively. Thus, routine pre-treatment screening of 100% of patients via the genotyping or plasma Uracil approaches (alone) would lead to at least 57,760 to 95,680 grade 4 toxicities per year, globally. The false negative rate for detecting grade 4 toxicity risk with the MPM is 4.7%. Thus, routine use of the MPM would reduce the number of grade 4 toxicities to around 7,520 per year globally, avoiding over 152,500 grade 4 toxicities per year.

Cost efficiency. In France, plasma Uracil dosing currently costs around 40€ per patient, genotyping around 52-177€ per patient, and the MPM costs around 150€ per patient. For a screening test to be cost efficient, the total cost of screening 100% of

patients must outweigh the incurred costs to the healthcare system of treating the severe toxic events, the hospitalizations in ICUs, and the deaths, that would be avoided by the screening.

The cost-effectiveness of screening using the MPM has been evaluated by our group.[36] The analysis was performed prospectively in 2010/2011 on 2 groups of patients. The first group (Group A, n = 886) received a standard dose of 5-FU (2400 mg/m^2) without prior screening of their DPD status. The second group (Group B, n = 856) were screened with an early version of the 5FU$^{ODPMTox™}$ program, followed by a dose adjustment if required. Patients were followed throughout 2 cycles of chemotherapy. In group A (no MPM screening), only the cost of treating the toxicities that arose were considered. In group B (MPM screening), the cost of the screening tests (153€ /patient) plus the cost of treating any toxicities that arose despite dose adjustment were taken into account. In group A, the incidence of grade 4 or 5 toxicity was 5.80% and 6.90% in the 1st and 2nd cycle of 5-FU treatment, respectively.

One patient in group A died from their ADR. In group B (MPM screening, with dose-adjustment) the incidence of grade 4 toxicity was 0.5% and 0.9% in the 1st and 2nd cycle of 5-FU treatment, respectively. No patients died of ADRs in group B. The cost benefit of implementing the MPM screening and dose adjustment was calculated to be 426 € per patient (508€ per patient with standard treatment versus 195€ per patient using the MPM screening). Pre-treatment screening with the MPM was thus shown to reduce the incidence of toxicity (and patient suffering) associated with 5-FU, avoid unnecessary deaths, and provide high cost efficiency.

Conclusion

Herein we demonstrate that a multi-parametric method (MPM) for predicting risk of FP toxicity significantly out-performs ($p < 0.0001$) the use of genotypic or phenotypic (plasma U or UH_2/U) information alone. Its specificity and sensitivity are both $\geq$ 95% for predicting $\geq$ grade 3 toxicity and reaches 98% for predicting (and thus avoiding) grade 5 toxicity (as opposed to 48.8%, 80.5% and 90.2% for genotyping, plasma [U] and plasma UH_2/U, respectively).

Routine, systematic use of the MPM (instead of genotyping or phenotyping, alone) is estimated to be capable of saving over 7,800 unnecessary deaths per year and would spare over 150,000 patients from debilitating grade 4 toxicity.

The improved performance and the cost-effectiveness of the MPM approach gives it real potential to go beyond currently recommended screening methods and to substantially increase the clinical benefit-risk ratio of FP treatment. MPM should therefore be considered for adoption as a Standard of Care for all cancer patients treated with fluoropyrimidines.

Acknowledgments

The authors thank Dr Paul Baldwin, a medical writer and translator, for help in preparing this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Ligue Contre le Cancer comité départemental du Maine et Loire, France.

Supplemental Material

Supplemental material for this article is available online.

References

1. Meyerhardt JA, Mayer RJ. Systemic therapy for colorectal cancer. *N Engl J Med*. 2005;352(5):476-487.
2. Rich TA, Shepard RC, Mosley ST. Four decades of continuing innovation with fluorouracil: current and future approaches to fluorouracil chemoradiation therapy. *J Clin Oncol*. 2004;22(11): 2214-2232.
3. Boisdron-Celle M, Capitain O, Faroux R, et al. Prevention of 5-fluorouracil-induced early severe toxicity by pre-therapeutic dihydropyrimidine dehydrogenase deficiency screening: assessment of a multiparametric approach. *Semin Oncol*. 2017;44(1):13-23.
4. Takimoto CH. The clinical pharmacology of the oral fluoropyrimidines. *Curr Probl Cancer*. 2001;25(3):134-213.
5. Gamelin E, Boisdron-Celle M, Guérin-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.
6. Mounier-Boutoille H, Boisdron-Celle M, Cauchin E, et al. Lethal outcome of 5-fluorouracil infusion in a patient with a total DPD deficiency and a double *DPYD* and *UTG1A1* gene mutation. *Br J Clin Pharmacol*. 2010;70(2):280-283.
7. National Cancer Institute, Division of Cancer Treatment & Diagnosis. Common Terminology Criteria for Adverse Events v.5.0. Published 2017. Accessed April 8, 2020. https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf
8. Froehlich TK, Amstutz U, Aebi S, Joerger M, Largiade CR. Clinical importance of risk variants in the dihydropyrimidine dehydrogenase gene for the prediction of early-onset fluoropyrimidine toxicity. *Int J Cancer*. 2015;136(3):730-739.
9. Meulendijks D, Henricks LM, Sonke GS, 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*. 2015;16(16):1639-1650.
10. van Staveren MC, Guchelaar HJ, van Kuilenburg AB, Gelderblom H, Maring JG. Evaluation of predictive tests for screening for dihydropyrimidine dehydrogenase deficiency. *Pharmacogenomics J*. 2013;13(5):389-395.
11. Ciccolini J, Del Re M, Danesi R, Milano G, Schellens JHM, Raymond E. Predicting fluoropyrimidine-related toxicity: turning

wish to will, the PAMM-EORTC position. *Ann Oncol*. 2018; 29(9):1893-1894.
12. Gamelin E, Delva R, Jacob J, et al. Individual fluorouracil dose adjustment based on pharmacokinetic follow-up compared with conventional dosage: results of a multicenter randomized trial of patients with metastatic colorectal cancer. *J Clin Oncol*. 2008; 26(13):2099-2105.
13. Jansman FG, Sleijfer DT, Coenen JL, De Graaf JC, Brouwers JR. Risk factors determining chemotherapeutic toxicity in patients with advanced colorectal cancer. *Drug Saf*. 2000;23(4):255-278.
14. Grem JL. 5-Fluorouracil: forty-plus and still ticking. A review of its preclinical and clinical development. *Invest New Drugs*. 2000; 18(4):299-313.
15. Tsalic M, Bar-Sela G, Beny A, Visel B, Haim N.Severe toxicity related to the 5-fluorouracil/leucovorin combination (The Mayo Clinic Regimen): a prospective study in colorectal cancer patients. *Am J Clin Oncol*. 2003;26(1):103-106.
16. Amstutz U, Froehlich TK, Largiader CR. Dihydropyrimidine dehydrogenase gene as a major predictor of severe 5-fluorouracil toxicity. *Pharmacogenomics*. 2011;12(9):1321-1336.
17. Johnson MR, Yan J, Shao L, Albin N, Diasio RB. Semi-automated radioassay for determination of dihydropyrimidine dehydrogenase (DPD) activity. Screening cancer patients for DPD deficiency, a condition associated with 5-fluorouracil toxicity. *J Chromatogr B Biomed Sci Appl*. 1997;696(2):183-191.
18. Seck K, Riemer S, Kates R, et al. Analysis of the *DPYD* gene implicated in 5-fluorouracil catabolism in a cohort of Caucasian individuals. *Clin Cancer Res*. 2005;11(16):5886-5892.
19. Diasio RB, Beavers TL, Carpenter JT. Familial deficiency of dihydropyrimidine dehydrogenase. Biochemical basis for familial pyrimidinemia and severe 5-fluorouracil-induced toxicity. *J Clin Invest*. 1988;81(1):47-51.
20. Morel A, Boisdron-Celle M, Fey L, Laine-Cessac P, Gamelin E. Identification of a novel mutation in the dihydropyrimidine dehydrogenase gene in a patient with a lethal outcome following 5-fluorouracil administration and the determination of its frequency in a population of 500 patients with colorectal carcinoma. *Clin Biochem*. 2007;40(1-2):11-17.
21. Amstutz U, Henricks LM, Offer SM, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for dihydropyrimidine dehydrogenase genotype and fluoropyrimidine dosing: 2017 Update. *Clin Pharmacol Ther*. 2018;103(2): 210-216. https://www.ncbi.nlm.nih.gov/pubmed/29152729
22. Caudle KE, Thorn CF, Klein TE, et al. Clinical Pharmacogenetics Implementation Consortium Guidelines for dihydropyrimidine genotype and fluoropyrimidine dosing. *Clin Pharmacol Ther*. 2013;94(6):640-645.
23. Morel A, Boisdron-Celle M, Fey L, et al. Clinical relevance of different dihydropyrimidine dehydrogenase gene single nucleotide polymorphisms on 5-fluorouracil tolerance. *Mol Cancer Ther*. 2006;5(11):2895-2904.
24. 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 Lett*. 2007;249(2):271-282.
25. Capitain O, Boisdron-Celle M, Poirier AL, Abadie-Lacourtoisie S, Morel A, Gamelin E. The influence of fluorouracil outcome parameters on tolerance and efficacy in patients with advanced colorectal cancer. *Pharmacogenomics J*. 2008;8(4):256-267.
26. 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.
27. *Common Terminology Criteria for Adverse Events (CTCAE). SERVICES USDOHAH, ed. Version 4.0*. NIH Publication; 2010:80.
28. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PMM. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol*. 2003;56(11):1129-1135.
29. Delacour H, François N, Servonnet A, Gentile A, Roche B.Likelihood ratios: a powerful tool for biological test's interpretation. *Immunoanal Biol Spe*. 2009;24(2):92-99. doi:10.1016/j.immbio. 2009.01.002
30. Remaud G, Boisdron-Celle M, Hameline C, Morel A, Gamelin E. An accurate dihydrouracil/uracil determination using improved high-performance liquid chromatography method for preventing fluoropyrimidines-related toxicity in clinical practice. *J Chromatog B Analyt Technol Biomed Life Sci*. 2005; 823(2):98-107.
31. Giraud C, et al. Haute Autorité de Santé, Screening for dihydropyrimidine dehydrogenase deficiency to decrease the risk of severe toxicities related to fluoropyrimidines. INAHTA Brief, Saint-Denis La Plaine: HAS. Published 2018. Accessed April 2020. https://www.has-sante.fr/upload/docs/application/pdf/2019-01/inahta_brief_dpd_5fu.pdf
32. Gamelin E, Morel A, Boisdron-Celle M. https://bases-brevets.inpi.fr/fr/document/FR2884257.html?s=1561902718198&p=5&cHash=0e216e1cdc4520016d8d1be604f59712. Accessed April 8, 2020.
33. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995;57(1):289-300.
34. Stevenson M, Nunes T, Heuer C, et al. epiR: tools for the analysis of epidemiological data. R package version 1.0-15. Published 2020. https://rdrr.io/cran/epiR/
35. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. version 4.0.2 Published 2020 . https://www.R-project.org/
36. Traoré S, Boisdron-Celle M, Hunault G, et al. DPD deficiency: medico-economic evaluation of pre-treatment screening of 5-FU toxicity. *J Clin Oncol*. 2012;30(4):410.