A case-control study of a combination of single nucleotide polymorphisms and clinical parameters to predict clinically relevant toxicity associated with fluoropyrimidine and platinum-based chemotherapy in gastric cancer

Miguel Cordova-Delgado, María Loreto Bravo, Elisa Cumsille, Charlotte N. Hill, Matías Muñoz-Medel, Mauricio P. Pinto, Ignacio N. Retama, María A. Lavanderos, Juan Francisco Miquel, María Rodriguez-Fernandez, Yuwei Liao, Zhiguang Li, Alejandro H. Corvalán, Ricardo Armisén, Marcelo Garrido, Luis A. Quiñones and Gareth I. Owen

Abstract

Background: Fluoropyrimidine plus platinum chemotherapy remains the standard first line treatment for gastric cancer (GC). Guidelines exist for the clinical interpretation of four DPYD genotypes related to severe fluoropyrimidine toxicity within European populations. However, the frequency of these single nucleotide polymorphisms (SNPs) in the Latin American population is low (< 0.7%). No guidelines have been developed for platinum. Herein, we present association between clinical factors and common SNPs in the development of grade 3–4 toxicity.

Methods: Retrospectively, 224 clinical records of GC patients were screened, of which 93 patients were incorporated into the study. Eleven SNPs with minor allelic frequency above 5% in GSTP1, ERCC2, ERCC1, TP53, UMPs, SHMT1, MTHFR, ABC2 and DPYD were assessed. Association between patient clinical characteristics and toxicity was estimated using logistic regression models and classification algorithms.

* Correspondence: lquinone@uchile.cl, gowen@bio.puc.cl

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
**Results:** Reported grade ≤ 2 and 3–4 toxicities were 64.6% (61/93) and 34.4% (32/93) respectively. Selected DPYD SNPs were associated with higher toxicity (rs1801265; OR = 4.20; 95% CI = 1.70–10.95, p = 0.002), while others displayed a trend towards lower toxicity (rs1801159; OR = 0.45; 95% CI = 0.19–1.08; p = 0.071). Combination of paired SNPs demonstrated significant associations in DPYD (rs1801265), UMPs (rs1801019), ABCC2 (rs717620) and SHMT1 (rs1979277). Using multivariate logistic regression that combined age, sex, peri-operative chemotherapy, 5-FU regimen, the binary combination of the SNPs DPYD (rs1801265) + ABCC2 (rs717620), and DPYD (rs1801159) displayed the best predictive performance. A nomogram was constructed to assess the risk of developing overall toxicity.

**Conclusion:** Pending further validation, this model could predict chemotherapy associated toxicity and improve GC patient quality of life.

**Keywords:** Predictive models, Single nucleotide polymorphism, Chemotherapy toxicity, Fluoropyrimidines, Platinums

**Introduction**
Globally, gastric cancer (GC) is the sixth most common malignancy and the third leading cause of cancer death [1–4]. Current standard first-line treatment for GC patients consists of chemotherapy regimens that combine fluoropyrimidines and platinum compounds. Therapeutic responses and associated toxicity with these regimens can vary significantly among patients ranging from moderate to severe [5–7]. Indeed, gastrointestinal, hematological and neurological toxicities are commonly observed under these regimens, often leading to treatment discontinuation, a reduction in quality of life, and in some extreme cases to death [8, 9]. Hence, severe toxicity becomes an essential obstacle to treatment completion and predictive models of toxicity may improve patient quality of life by avoiding severe toxicity.

Previous studies have demonstrated that single nucleotide polymorphisms (SNPs) are associated with chemotherapy-associated toxicity [10–12]. This can be explained by gene variations that alter the enzymatic activity of key proteins affecting pharmacokinetic and pharmacodynamic processes [13, 14]. In this regard, platinum-based compounds can trigger cell arrest or apoptosis by forming Pt-DNA adducts [15]. Within our bodies, kidneys can excrete these compounds without undergoing biotransformation via B1/C2/G2 type ABC (ATP Binding Cassette) transporters [16, 17]. Within cells, metabolizing enzymes including GSTP1, GSTM1, NQO1 and SOD1 decrease intracellular levels of platinum compounds [18–21]. Intracellularly platinum compounds target the DNA forming DNA-Pt complexes. Damaged DNA is recognized by HMGB, an enzyme that coordinates DNA repair by nucleotide excision repair enzymes [19, 22]. On the other hand, 5-fluorouracil (5-FU) and its pro-drug capecitabine undergo a series of enzymatic transformations prior to exert their effects [23]. Although the precise mechanism is still unclear, 5-FU is known to inhibit thymidylate synthase (TYMS) suppressing the conversion of uracil into thymidine, leading to the inhibition of DNA/RNA synthesis and eventually to cell death [24]. The metabolism of 5-FU occurs mainly in the liver, where DPYD metabolizes ~ 80% of the drug, producing 5,6 dihydroxy-5-FU (an inactive metabolite) [25]. It is widely documented that decreased DPYD activity is associated with severe toxicity [26–28]. Previous reports have also associated TYMS and MTHFR gene variations with 5-FU toxicity; however their clinical relevance is undetermined [29].

To date, the most reliable markers of fluoropyrimidine toxicity are DPYD*2A (rs3918290), DPYD-c.2846A > T (rs67376798), DPYD-Hap-B3 (rs56038477) and DPYD*13 (rs55886062). In fact, these are variants that have a well-documented association with severe toxicity associated with fluoropyrimidines, and there is a Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline that recommends avoiding or reducing the dose of fluoropyrimidines if a patient carries any of these variants [30]. Unfortunately, given their low frequencies in the general population the use of these variants identifies only a small fraction of potentially at-risk patients. For example, their frequencies in the 1000-Genome (1000-G) project and gnomAD databases for American or Latino/Admixed American population are much lower than that of those published for European cohorts. According to these databases, the Latin-American frequencies of the risk allele for: DPYD * 2 is 0.1 and 0.2% (1000-G Project and gnomAD, respectively) yet an order of magnitude greater in frequency in a Finnish cohort (2.5%) [31, 32]. In Latin-America the DPYD-c.2846A > T frequency is 0.3 and 0.2% (1000-G Project and gnomAD, respectively); DPYD-Hap-B3 is 0.6 and 0.7% (1000-G Project and gnomAD, respectively); DPYD*13 is 0 and 0.007% (1000-G Project and gnomAD, respectively). Therefore, we hypothesize that common SNPs in the Latin American population can potentially explain the clinically relevant toxicity in patients with fluoropyrimidines and platinum-based treatment.

As previously mentioned, SNP variants have been previously correlated with chemotherapy toxicity [33–35]. Other factors such as chemotherapy scheme, dosage, sex, and age have also been implicated in the
development and severity of toxicity [7, 36–38]. However, only a few studies have developed comprehensive models that incorporate genetic and non-genetic factors to predict toxicity [39–41]. Herein, we developed and tested several models based on clinical factors, treatment regimens and candidate-SNPs. Our best performance model was used to construct a nomogram.

Materials and methods

Patients and study design

A retrospective, observational and case/control study was carried out. A total of 224 gastric cancer patients diagnosed between April 2005 and March 2018 were registered at the UC-CHRISTUS Cancer Center in the Pontificia Universidad Católica de Chile (PUC). Previously, this group of patients has been clinically and molecularly characterized [42]. After applying the inclusion and exclusion criteria, a total of 93 CG patients were analyzed in this study. Eligibility criteria were: a) histologically confirmed GC, b) chemotherapy regimen based on fluoropyridines and/or platinum compounds, c) adequate patient renal, hepatic and bone marrow function, determined by the treating physician at the time of starting chemotherapy, d) at least 2 cycles of chemotherapy, e) availability of biological sample for extraction of genetic material and f) adults (>18 yr-old). Patients with neuropathies or hematological damage caused by other diseases were excluded. Clinical-pathological characteristics of patients included: age, sex, stage, ECOG, histological classification, treatment schemes used and comorbidities. The Ethics Committee at “PUC” approved this study (#16–046, April 21st, 2016) [4]. All participants signed an informed consent to participate in this study. A waiver of consent was granted to include deceased patients. All data were anonymized to protect patients’ privacy. This study strictly adhered to the Code of Ethics of the World Medical Association (Declaration of Helsinki, 1964).

Toxicity graduation

Toxicities grades were determined following the National Cancer Institute Common Toxicity Criteria 4.0 (NCI-CTC 4.0). Data on anemia, neutropenia, febrile neutropenia, thrombocytopenia, nausea, vomiting, diarrhea, stomatitis, hand-foot syndrome, and peripheral neuropathy were collected. Then they were categorized into hematological, gastrointestinal, and neurological toxicity, and if they presented any of the above as “overall toxicity”. All association analysis evaluated grade 0–2 vs grade 3–4 for hematological, gastrointestinal, neurological and overall toxicity. Treatment schemes and supportive care are shown in Supplementary Data.

SNP selection, DNA extraction and genotyping

A total of 11 SNPs were assessed, and a detailed description of this process is provided in Supplementary Data. Genotypic/allelic frequencies of analyzed SNPs are shown in Supplementary Table S8. Nucleic acids were extracted from paraffin-embedded tumor tissues using the “AllPrep DNA/RNA Mini Kit” kit (Cat#AM1975, Thermo Fisher). DNA was quantified by “Qubit® dsDNA HS Assay” (Thermo Fisher). Candidate SNPs were genotyped by TaqMan® SNP Genotyping Assay technology on an Applied Biosystems® 7500 Fast Real-Time PCR System (Thermo Fisher). Samples were randomly reanalyzed for confirmation. TaqMan® probes are shown in Supplementary Table S9.

Statistical analysis

Association analysis of SNPs

The association between SNPs and grade 3–4 toxicity was analyzed using univariate logistic regression models, reporting Odds Ratio (OR) values with 95% confidence interval (95% CI). These analyses were tested using 3 inheritance models: dominant, codominant and recessive, based on the parameters of AIC and BIC, the best inheritance model was chosen for each SNP [43]. To choose the SNP combinations in the first step, a multivariate logistic regression analysis was performed using the 11 SNPs. To reduce the number of combinations and avoid overfitting, we applied the AIC-based “Stepwise algorithm”. On selected SNPs, binary combinations were performed and their association with grade 3–4 toxicity was established using their respective inheritance models.

Model building

We developed a total of 4 models for severe toxicity. Based on literature evidence or a p-value < 0.2, potentially relevant variables were included in each model. Multivariate logistic regression models were built incorporating variables that maximized Pseudo R² (McFadden); fit indicator of the variables to the model [41, 44]. For example, for overall toxicity Model 1 included: age, sex, peri-operative chemotherapy and scheme. Model 2 included the following SNPs: ERCC2 (rs13181), DPYD (rs2297595), DPYD (rs1801159), DPYD (rs1801265) and GSTP1 (rs1695). Model 3 included: age, sex, peri-operative chemotherapy, 5-FU containing scheme, ERCC2 (rs13181), DPYD (rs2297595), DPYD (rs1801159), DPYD (rs1801265) and GSTP1 (rs1695). Finally, model 4 included: age, sex, peri-operative chemotherapy, 5-FU containing scheme, paired SNPs DPYD (rs1801265) + ABCC2 (rs717620), plus DPYD (rs1801159). For every type of toxicity, the selected variables are depicted in the corresponding table.
Model evaluation and nomogram construction

Obtained models were evaluated using classification algorithms \cite{45, 46} including: Logistic Regression (LR); Support Vector Machine (SVM); Naïve Bayesian (NB); K-Nearest Neighbor (KNN); Artificial Neural Network (ANN); Random Forest (RF); Decision Tree (DT) (Details in Supplementary Data). Using as a basis the coefficients of the multivariate analysis of model 4 we constructed a nomogram using \textit{rms} package \cite{47}. In addition, for discriminatory capacity, 1000 bootstrap replications served as internal validation subsets to estimate the bias - corrected c - index calibration.

General statistical analysis

Continuous variables were compared using ANOVA. Kaplan-Meier method was used for survival analysis and log-rank tests for comparison. Significance was set at \( p < 0.05 \). According to the number of cases and controls (31 and 62 patients respectively), assuming a power of 80\%, an \( \alpha \) error of 5\% and a frequency of common polymorphisms (i.e. \textit{DPYD} rs1801265, \textit{GSTP1} rs1695) of 30\%, Odd Ratios could be detected with values of 3.9 and 0.1 (high and low). Association analysis of SNPs and overall toxicity were performed by SNPstat program \cite{43}. Uni/multivariate logistic regression models were built using the \textit{stats} and \textit{DescTools} packages. Classification algorithms were constructed using the \textit{caret} and \textit{ROCR} packages. For survival analysis, the \textit{survival} and \textit{survminer} packages were used. All analysis were performed in R software v3.5.1 (The R Foundation, Vienna, Austria). Full datasets used in this study can be found in Supplementary Data File 1.

Results

SNPs selection

SNPs were selected based on: (1) scientific evidence regarding the SNPs/toxicity relationship, using the PharmGKB database \cite{48}; (2) allelic and genotypic frequency of the SNPs in the American population \cite{32}; (3) relationship of the SNPs with the toxicity collected in our patients literature-based criteria; (4) functional impact of SNPs at the protein level according to PolyPhen \cite{49} and SIFT \cite{50}. Briefly, in a first approximation 27 SNPs in 11 genes and 7 SNPs in 6 genes for fluoropyrimidines and platinums were reviewed, respectively. Then, based on a score system for fluoropyrimidines, 14 SNPs in 7 different genes were candidates, while for platinums, 4 SNPs in 4 different genes were candidates. Finally, eleven SNPs with an allelic frequency greater than 5\% were genotyped (Table 1). A detailed description of this process is provided in Supplementary Data.

General characterization of patients

Main clinicopathological characteristics of the patients are summarized in Supplementary Table S1. Briefly, patients were predominantly male (62.4\%) and advanced

Table 1 Brief description of the SNPs analyzed in this study

| Gen  | SNP ID | AA Change | SNPs effect | Genotype associated to toxicity | Allelic frequency American population (n)\( ^a \) | Toxicity                        |
|------|--------|-----------|-------------|-------------------------------|---------------------------------|---------------------------------|
| GSTP1| rs1695 | A > G     | I105V       | Protein alteration AA and AG   | Allele A: 0.52                  | ↑ risk of hematological toxicity|
| ERCC2| rs13181| T > G     | K751Q       | Protein alteration GG and GT   | Allele G: 0.21                  | ↑ risk of hematological toxicity|
| TPS3 | rs1042522| C > G   | P72R        | Protein alteration CG and GG   | Allele G: 0.32                  | ↑ risk of hematological toxicity|
| UMP5 | rs1801019| G > C  | G213A       | Protein alteration CC          | Allele C: 0.26                  | ↑ risk of gastrointestinal toxicity|
| SHMT1| rs1979277| G > A | L474F       | Protein alteration AA          | Allele A: 0.27                  | ↑ risk of toxicity               |
| MTHFR| rs1801131| T > G  | E429A       | Protein alteration TT          | Allele T: 0.15                  | ↑ risk of toxicity               |
| ABCC2| rs717620| C > T   | –           | Change in 5'-UTR               | Allele T: 0.17                  | ↑ risk of neurotoxicity          |
| ERCCI| rs11615| A > G    | N118=       | Synonymous change AA           | Allele A: 0.39                  | ↑ risk of hematological toxicity|
| DPYD | rs2297595| T > C  | M166V       | Protein alteration CC and CT   | Allele C: 0.06                  | ↑ risk of severe toxicity        |
| DPYD | rs1801159| T > C  | I543V       | Protein alteration CC and CT   | Allele C: 0.27                  | ↑ risk of severe toxicity        |
| DPYD | rs1801265| A > G  | C29R        | Protein alteration GG and AG   | Allele G: 0.22                  | ↑ risk of severe toxicity        |

\( ^a \) Frequencies of genotypes associated to toxicity in American population according 1000-G project
stage III / IV (71.0%). Most tumors were located in the stomach corpus (31.2%). Histologically, 32.3% were diffuse-type by Lauren and 77.3% were gastric adenocarcinomas. Survival curves are shown in Supplementary Fig. S1. Median overall survival (OS) for the entire group was 29 months (Supplementary Fig. S1A). Males displayed better OS versus females, however these differences did not reach statistical significance (Log Rank \( p = 0.45 \), Supplementary Fig. S1B). As expected, stage had a significant impact on OS. Advanced stage patients showed lower median OS (30 or 13 months for stages III or IV, respectively) versus early stage (62 months in stage II, not reached for stage I) (Supplementary Fig. S1C, Log Rank \( p < 0.0001 \)). Supplementary Table S2 summarizes treatment characteristics. FOLFOX was the most frequently used schema (49.5%) followed by CAPEOX and CF (18.3 and 14.0%, respectively). Adverse reactions are summarized in Supplementary Table S3. Data were grouped according to type of toxicity. Peripheral neuropathy was the most common grade 1 toxicity (34.1%). Nausea was the most predominant grade 2 toxicity (31.5%). Among grade 3 toxicities, neutropenia was dominant (22.58%) followed by diarrhea (20%). Finally, we registered a total of 5 patients with grade 4 events among these 3 out of 5 corresponded to neuropathy (60%). No toxicity-related deaths were registered. Clinically relevant toxicities (grade \( \geq 3 \)) were more frequently associated to digestive problems such as diarrhea and stomatitis, with a total of 19 registered events. Among hematological toxicities, a total of 17 grade \( \geq 3 \) events, principally neutropenia or febrile neutropenia, were registered (Supplementary Table S3).

Table 2: Univariate logistic regression analysis for the association between clinicopathological variables and overall toxicity grade \( \geq 3 \)

| Characteristic          | Control (n = 61) | Case (n = 32) | OR    | 95% CI          | p-value |
|-------------------------|-----------------|--------------|-------|-----------------|---------|
| Age median (years)      |                 |              |       |                 |         |
| < 59                    | 32 (52.5)       | 12 (37.5)    | Ref.  |                 | 0.16    |
| \( \geq 59 \)           | 28 (45.9)       | 20 (62.5)    | 1.83  | (0.77–4.49)     |         |
| Sex                     |                 |              |       |                 |         |
| Male                    | 41 (67.2)       | 19 (59.4)    | Ref.  |                 | 0.45    |
| Female                  | 20 (32.8)       | 13 (40.6)    | 1.40  | (0.57–3.40)     |         |
| Stage                   |                 |              |       |                 |         |
| I-II                    | 20 (32.8)       | 7 (21.9)     | Ref.  |                 | 0.26    |
| III-IV                  | 41 (67.2)       | 25 (78.1)    | 1.74  | (0.67–4.97)     |         |
| ECOG                    |                 |              |       |                 |         |
| 0                       | 28 (45.9)       | 14 (43.8)    | Ref.  |                 | 0.44    |
| 1                       | 26 (42.6)       | 11 (34.4)    | 0.84  | (0.32–2.19)     |         |
| 2                       | 1 (1.6)         | 2 (6.3)      | 4.0   | (0.35–90.53)    |         |
| NA                      | 6 (9.8)         | 5 (15.6)     |       |                 |         |
| Lauren histotype        |                 |              |       |                 |         |
| Diffuse                 | 20 (32.8)       | 10 (31.3)    | Ref.  |                 | 0.98    |
| Intestinal              | 17 (27.9)       | 9 (28.1)     | 1.05  | (0.34–3.22)     |         |
| Mixed                   | 7 (11.5)        | 4 (12.5)     | 1.14  | (0.25–4.77)     |         |
| NA                      | 17 (27.9)       | 9 (28.1)     |       |                 |         |
| Signet-ring cells       |                 |              |       |                 |         |
| No                      | 35 (57.4)       | 19 (59.4)    | Ref.  |                 | 0.85    |
| Yes                     | 26 (42.6)       | 13 (40.6)    | 0.92  | (0.38–2.18)     |         |
| Comorbidities           |                 |              |       |                 |         |
| Chronic heart disease   | 5 (8.2)         | 2 (6.3)      | 0.74  | (0.10–3.69)     | 0.73    |
| Chronic liver disease   | 1 (1.1)         | 1 (3.3)      | 1.93  | (0.07–50.06)    | 0.64    |
| Thromboembolic event    | 1 (1.1)         | 1 (3.3)      | 1.93  | (0.07–50.06)    | 0.64    |
| Diabetes                | 9 (14.8)        | 2 (6.3)      | 1.10  | (0.26–3.97)     | 0.88    |

NA not available, ECOG Eastern Cooperative Oncology Group
Significance: \( P < 0.05 \)
Association between non-genetic factors and overall toxicity
A 34.4% of patients (32/93) displayed grade ≥ 3 overall toxicity. Non-genetic factors and ≥ 3 grade overall toxicity associations are shown in Tables 2 and 3. No significant associations were found between clinicopathological variables and serious adverse reactions. In agreement with previous reports, elderly patients have an OR = 1.83 (95% CI = 0.77–4.49), female patients had an OR = 1.40 (95% CI = 0.57–3.40) and ECOG = 2 patients displayed an OR = 4.00 (95% CI = 0.35–90.53). However, none of these associations reached statistical significance (Table 2). Regarding treatment associated factors (Table 3), patients that received peri-operative chemotherapy regimens had higher grade 3–4 toxicity rates compared to adjuvant treatment (OR = 2.71, 95% CI = 0.91–8.62, p = 0.07, peri vs. adj). Given the heterogeneity of the used schemes in our study, we classified them according to drug contents; 5-FU-containing schemes were associated to higher toxicity OR = 2.26 (95% CI: 0.79–7.48, p = 0.12). In contrast, capecitabine-containing schemes displayed lower toxicity OR = 0.52 (95% CI = 0.15–1.50, p = 0.23). However, none of these associations reached statistical significance.

Genetic variants associated with overall toxicity
Binary associations between overall toxicity and SNPs were assessed using three inheritance models (codominant, dominant and recessive) and are summarized in

Table 3 Univariate logistic regression analysis for association between treatment variables and overall grade ≥ 3 toxicity

| Characteristic                  | Control (n = 61) | Case (n = 32) | OR [95% CI] | p-value |
|--------------------------------|-----------------|---------------|-------------|---------|
| **Chemotherapy regimen**       |                 |               |             |         |
| Adjuvant                       | 19 (31.1)       | 7 (21.9)      | Ref.        | 0.21    |
| Peri-operative                 | 16 (26.2)       | 16 (50)       | 2.71 (0.91–8.62) | 0.07   |
| Palliative                     | 22 (36.1)       | 8 (25)        | 0.98 (0.29–3.30) |         |
| CMT + RDT Adjuvant             | 3 (4.9)         | 1 (3.1)       | 0.90 (0.04–8.52) |         |
| CMT + RDT Peri-operative       | 1 (1.6)         | 0             | NA          |         |
| **Chemotherapy scheme**        |                 |               |             |         |
| FOLFOX                         | 29 (47.5)       | 17 (53.1)     | Ref.        | 0.07    |
| CAPEOX                         | 12 (19.7)       | 5 (15.6)      | 0.71 (0.19–1.05) |         |
| CF                             | 7 (11.5)        | 6 (18.8)      | 1.46 (0.40–5.12) |         |
| DCFm                           | 5 (8.2)         | 0             | NA          |         |
| ECF                            | 0               | 2 (6.3)       | NA          |         |
| EOX                            | 2 (3.3)         | 0             | NA          |         |
| Capecitabine                   | 2 (3.3)         | 0             | NA          |         |
| FLOT                           | 0               | 1 (3.1)       | NA          |         |
| RDT + S-FU/Leu                 | 2 (3.3)         | 1 (3.1)       | 0.85 (0.03–9.55) |         |
| RDT + CAPEOX                   | 1 (1.6)         | 0             | NA          |         |
| RDT + Cis/Cape                 | 1 (1.6)         | 0             | NA          |         |
| **Previously treated**         |                 |               |             |         |
| No                             | 53 (86.9)       | 30 (93.7)     | Ref.        | 0.29    |
| Yes                            | 8 (13.1)        | 2 (6.3)       | 0.44 (0.06–1.90) |         |
| **Scheme contains**            |                 |               |             |         |
| RDT                            | 4 (6.6)         | 1 (3.1)       | 0.45 (0.02–3.27) | 0.46    |
| S-FU                           | 43 (70.5)       | 27 (84.4)     | 2.26 (0.79–7.48) | 0.12    |
| Capecitabine                   | 16 (26.2)       | 5 (15.6)      | 0.52 (0.15–1.50) | 0.23    |
| Oxaliplatin                    | 44 (72.1)       | 23 (71.9)     | 0.81 (0.30–2.21) | 0.67    |
| Cisplatin                      | 13 (21.3)       | 8 (25.0)      | 1.23 (0.43–3.33) | 0.68    |
| Docetaxel                      | 5 (8.2)         | 3 (9.4)       | 0.36 (0.01–2.37) | 0.31    |

CMT, chemotherapy; RDT, Radiotherapy; FOLFOX, 5-fluorouracil + oxaliplatin + leucovorin; CAPEOX, capecitabine + oxaliplatin; CF, cisplatin + 5-fluorouracil; DCFm, docetaxel + cisplatin + 5-fluorouracil; ECF, etoposide + cisplatin + 5-fluorouracil; FLOT, 5-fluorouracil + oxaliplatin + docetaxel + leucovorin; SFU, 5-fluorouracil; Leu, leucovorin; Cis, cisplatin; Cape, capecitabine; NA, not applicable; Ref. Reference

Significance: P < 0.05
Using a dominant model, the AG/GG genotypes of SNPs in the DPYD (rs1801159) were associated with higher toxicity; OR = 4.20 (95% CI = 1.70–10.95, \( p = 0.002 \)). Also, we found a borderline association between lower toxicity and DPYD (rs1801159) with an OR = 0.45 (95% CI = 0.19–1.08; \( p = 0.071 \)). Potential associations in DPYD (rs2297595), ERCC2 (rs13181) and GSTP1 (rs1695) SNPs were also analyzed. However, no significant association was found by univariate analysis.

Combination of genetic variants associated with overall toxicity

Next, we performed a multivariate logistic regression analysis incorporating the 11 SNPs to establish potential associations between combined SNPs and overall toxicity. We applied a “Stepwise algorithm” based on Akaike information criterion (AIC) [51] to reduce the number of combinations and avoid overfitting. Based on this we selected, 5 SNPs and their respective inheritance models to test binary combinations between SNPs (Table 5). In respect to SNP combinations, those patients that carry the AG/GG + GG/GC genotype combination in DPYD (rs1801265) and UMPS (rs1801019) are at a higher risk to develop toxicity; OR = 4.22 (95% CI = 1.66–11.40, \( p = 0.0031 \)) versus AA + GG / GC genotype patients. Furthermore, the DPYD (rs1801265)/SHMT1 (rs1979277) combination displayed a borderline association with grade 3–4 overall toxicity. In this case, the combination AG/GG + CT/TT genotype showed a higher probability of developing toxicity; OR = 11.25 (95% CI = 1.25–245.45, \( p = 0.047 \)) against the AA + CC genotype.

### Table 4

| SNPs and overall grade \( \geq 3 \) toxicity associations in GC patients treated with platinum/fluoropyrimidines-based chemotherapy |
|---|
| **Gen SNP ID** | **Model** | **Genotypes** | **Control (n = 61)** n (%) | **Case (n = 32)** n (%) | **OR [95% CI]** | **p-value** |
| **ERCC2** | Dom | T/T | 39 (63.9%) | 16 (50%) | Ref. | 0.20 |
| rs13181 T > G | | T/G-G/G | 22 (36.1%) | 16 (50%) | 1.77 (0.74–4.22) | |
| **DPYD** | – | T/T | 57 (95%) | 28 (87.5%) | Ref | |
| rs2297595 T > C | | C/T | 3 (5%) | 4 (12.5%) | 2.71 (0.57–12.97) | 0.21 |
| **DPYD** | Dom | T/T | 22 (36.7%) | 18 (56.2%) | Ref | |
| rs1801159 T > C | | C/T-C/C | 38 (63.3%) | 14 (43.8%) | 0.45 (0.19–1.08) | 0.071 |
| **DPYD** | Dom | A/A | 40 (66.7%) | 10 (32.3%) | Ref | |
| rs1801265 A > G | | A/G-G/G | 20 (33.3%) | 21 (67.7%) | 4.20 (1.70–10.95) | 0.002 |
| **GSTP1** | Dom | A/A | 22 (36.1%) | 14 (43.8%) | 0.72 (0.30–1.74) | 0.47 |

Dom: Dominant inheritance model
Significance: \( P < 0.05 \)

### Multivariate analysis for the development of prediction models

Next, we sought to determine if the addition of multiple genetic factors and clinical/treatment information delivered a better prediction. We developed four models with different variables using the multivariate logistic regression. Model 1 was restricted to clinical/treatment variables. Model 2 included only SNPs. Model 3 incorporated a combination of clinical and treatment variables plus SNPs. Finally, model 4 was a mixture of clinical/treatment variables and paired SNPs (see methods). Selection of variables for each model was based on maximum Pseudo R². Variables, ORs and Pseudo R² for each model are shown in Table 6. Interestingly, the addition of clinical variables and genotypes improved model performance. For model 1 Pseudo R² values were 0.073 and 0.15, respectively. In contrast, for combined models (model 3 and 4) values were 0.21 and 0.21, respectively. Furthermore, when we analyzed the SNPs in DPYD (rs1801159) in model 2 OR increased from 3.74 to 4.55 (\( p = 0.004 \)) after adding clinical variables (model 3). Interestingly, sex has been reported to be a determining factor in the effects of polymorphisms on DPYD [7]. In line with the literature, the association with grade 3–4 toxicity of AG/GG genotypes in relation to AA had an OR = 7.78 (95% CI = 2.31–31.79, \( p = 0.001 \)) or 1.85 (95% CI = 0.43–8.26, \( p = 0.40 \)) for male or female patients, respectively (see Supplementary Table S4). On the other hand, in model 4 the DPYD (rs1801265) + ABC2C (rs717620) pair displayed a strong association, in this case the combination AG/GG + CT/TT genotype showed an OR = 18.00 (95% CI = 1.66–439, \( p = 0.027 \)) versus AA + CT/TT.
Evaluation of toxicity models

To assess the predictive power of our models we employed a variety of classification algorithms. Figure 1 shows the sensitivity, specificity, accuracy and AUC of each model. For example, model 1 had a high specificity (range = 0.82–1.0), but low sensitivity (range = 0–0.17), in most tested algorithms. Accuracy reached a maximum value of 0.71 using the RL method and an AUC of 0.74 in the RL and ANN classification algorithms (Fig. 1A). Comparing different algorithms for model 2, were found a promising specificity (range = 0.61–0.87) but low sensitivity (range = 0.17–0.42). In this model maximum accuracy (0.69) was achieved with the KNN method, and the most favorable AUC was 0.68 with the DT method (Fig. 1B). Model 3 showed a relatively high specificity (range = 0.70–0.91) and a moderately low sensitivity (range = 0.17–0.41), reaching its maximum value with the RL method. On the other hand, the maximum accuracy was 0.69 with the SVM and KNN methods, and a maximum AUC of 0.68 achieved with the DT method (Fig. 1C). Model 4 showed a relatively high specificity (range = 0.74–1.0), with the sensitivity ranging between 0 and 0.66 among the classification algorithms. Maximum accuracy was achieved with the RL method (0.80); AUC was 0.82 achieved with the same classification algorithm (Fig. 1D). In summary, our data suggest model 4 was the best predictor of grade 3–4 toxicity (using the RL method). This supports the notion that combined models provide better predictive power versus individual variable models.

Multivariate analysis for type of toxicity

In line with the methodology employed for general toxicity, multivariate analysis was performed on clinical and genetic factors in the development of severe toxicity for independent hematological, gastrointestinal or neurological toxicities. For hematological toxicity, when the clinical/treatment variables were integrated with SNPs, a better fit of the logistic regression model was achieved. For example, for model 1 and model 2, the Pseudo R² returned values of 0.07 and 0.14, respectively, while for the variable’s integration models vary between 0.23 and 0.29 (Supplementary Table S5). Analysis of model 4 (Supplementary Table S5) shows that the variables of sex and the binary combination between the SNPs in ERCC1 (rs11615) and GSTP1 (rs1695) plus the SNP in DPYD (rs1801265) were either associated significantly or at least showed a strong tendency with hematological toxicity. For gastrointestinal toxicity, we observed that the clinical variables have a low influence on the presentation of this toxicity (Pseudo R² = 0.04). An improved fit is observed in upon the examination of the 3 SNPs (Pseudo R² value of 0.15; model 2). Upon combination of the clinical/treatment variables and genotypes, there is a modest increase in the Pseudo R² value to 0.20 and 0.18, for models 3 and 4, respectively (Supplementary Table S6). Applying the same analysis for neurological toxicity, only 3 models could be developed, where model 1, which includes only clinical factors / treatment, had a better fit (Pseudo R² = 0.15) than model 2 (Pseudo R² = 0.15) which included only the SNPs in TP53.
Table 6 Models for overall grade ≥ 3 toxicity in gastric cancer patients treated with platinum/fluoropyridines-based chemotherapy using multivariate analysis

| Characteristics          | Model 1 OR [95% CI] p-value | Model 2 OR [95% CI] p-value | Model 3 OR [95% CI] p-value | Model 4 OR [95% CI] p-value |
|--------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Age (above median)       | 1.71 (0.67–4.48) 0.25       | –                           | 2.33 (0.81–7.15) 0.12       | 2.24 (0.77–6.96) 0.14       |
| Sex (Female)             | 1.83 (0.68–5.05) 0.23       | –                           | 2.12 (0.68–7.01) 0.19       | 2.18 (0.70–7.29) 0.18       |
| CMT peri-op (Yes)        | 2.66 (1.00–7.35) 0.05       | –                           | 2.26 (0.73–7.28) 0.15       | 3.01 (0.94–10.31) 0.06      |
| S-FU based (Yes)         | 1.73 (0.55–6.10) 0.35       | –                           | 2.03 (0.58–8.06) 0.28       | 1.65 (0.46–6.50) 0.44       |
| **SNPs**                 |                             |                             |                             |                             |
| ERCC2 (dom. TG/GG)       | –                           | 1.56 (0.57–4.31) 0.37       | 1.24 (0.41–3.63) 0.68       | –                           |
| DPYD (CT) (rs2297595)    | –                           | 1.56 (0.26–9.95) 0.61       | 1.88 (0.25–14.77) 0.52      | –                           |
| DPYD (dom. TT) (rs1801159) | –                          | 2.32 (0.89–6.24) 0.08       | 2.53 (0.91–7.44) 0.07       | –                           |
| DPYD (dom. AG/GG) (rs1801265) | –                       | 3.74 (1.46–10.08) 0.006     | 4.55 (1.64–13.79) 0.004     | –                           |
| GSTP1 (dom. AG/GG) (rs1695) | –                        | 0.71 (0.26–1.90) 0.50       | 0.59 (0.19–1.74) 0.34       | –                           |
| **SNPs combination**     |                             |                             |                             |                             |
| DPYD dom. + ABCC2 dom. (rs1801265 + rs717620) | –                     | –                           | –                           | –                           |
| AA + CT/TT               | –                           | –                           | 1.45 (0.18–8.68) 0.68       | –                           |
| AG/GG + CC               | –                           | –                           | 4.33 (1.43–14.68) 0.012     | –                           |
| AG/GG + CT/TT            | –                           | –                           | 18.00 (1.66–439) 0.027      | –                           |
| DPYD (dom. TT) (rs1801159) | –                          | –                           | 2.37 (0.87–6.75) 0.094      | –                           |
| **Pseudo R²**            | 0.07                        | 0.15                        | 0.21                        | 0.21                        |

CMT Chemotherapy, Dom Dominant inheritance model
Significance: P < 0.05

Interestingly, in neurological toxicity, model 3, which includes both types of variables, had a Pseudo R² value of 0.19, the highest value among these models (Supplementary Table S7). Taken together, these results suggest that it is the combined models that provide higher Pseudo R² values. However, given the reduced number of variables (low number of SNP incidence) when individual toxicities are analyzed the associations may be relatively imprecise (reflected in the confidence intervals) and thus must be interpreted with caution until a larger cohort is studied.

Nomogram for predicting general toxicity

As an approximation for future validation of our results we developed a multivariate logistic regression-based nomogram that estimates the probability of a given patient to experience grade 3–4 overall general toxicity (Fig. 2A). This model is well calibrated (Supplementary Fig. S2) and has an acceptable discriminatory capacity, with an optimism-corrected c-index of 0.72 (95% CI, 0.72–0.92). Figure 2B shows the distribution of nomogram values for each patient. As expected, the median values for low or high-toxicity groups were significantly different (p < 0.0001). In addition, we established a different cut-off according to the points on delivered by the nomogram. Thereby ≤45-point patients have a 10% probability of developing toxicity. Encouragingly, 93% of patients in the lower range group correspond to the low-toxicity group. In contrast, patients with > 136-accumulated
Fig. 1 (See legend on next page.)
points are at a higher risk to develop toxicity and 73% of them are in the high-toxicity group.

**Discussion**

Chemotherapy treatment-related toxicity remains a critical problem in GC patients. Unfortunately, the relative benefit in terms of patient survival associated with targeted therapies is rather modest [52–55]. Therefore, optimizing chemotherapy regimens becomes crucial to improve GC patient survival. Within this context, the elaboration of reliable models that predict treatment-related toxicities might ensure patient safety in intervention studies. While guidelines exist for genotypes relating to fluoropyrimidine toxicity, our study also demonstrates associations related to platinum presence. The Clinical Pharmacogenetics Implementation Consortium has delivered guidelines for the clinical interpretation of four $DPYD$ genotypes related to severe fluoropyrimidine toxicity within European populations [56]. However, while the frequency of these single nucleotide polymorphisms (SNPs) could reach ~10% in some European populations the reported frequency in the Latin-American populations is below 0.7% [57] (Suarez-Kurtz 2020; Nugent et al. 2019). This may be primarily due to the underrepresentation of the Latin-American population in genetic studies [58]. In accordance, the SNPs in $DPYD$ (rs55886062), included in the CPIC guide, were not mutated homozygous nor were the heterozygous genotypes detected in the Chilean patients used in this study (data not shown). Given the world population that is not derived from European ancestry, the identification of new associations between the genome and fluoropyrimidine and platinum toxicity is of the utmost importance and may complement the current CPIC guidelines once further validation has been completed. Our findings are one of first to frame toxicity pharmacogenetics in an underexplored Latin American population and given the inherent global differences in SNP distribution, it is not beyond the realms of imagination to envisage that future pharmacogenetic tests are applied in a regional or populational manner.

In line with previous GC reports, patients in our cohort were predominantly males [59, 60] and advanced stage [42, 60, 61]. Similarly, median overall survival, histological type and overall toxicity were in agreement with the current literature (Supplementary Table S1 and Supplementary Fig. S1). Regarding age at diagnosis, the association with combined chemotherapy toxicity is probably explained by the age of recruited participants in most GC-trials that range between 50 and 60 years [62]. Interestingly, the inclusion of age in our final model increased the predictive power. A study reported no significant differences in the incidence of grade 3–4 toxicities in gastro-esophageal cancer patients comparing ≥70 vs < 70 year-old participants [63]. However, in many cases a higher prevalence of toxicity in older patients leads to chemotherapy discontinuation [64]. In contrast, a pooled analysis concluded that chemotherapy-related serious adverse events were significantly higher in > 65 year-old patients [65]. Accordingly, a recent study demonstrated that older GC patients (≥ 70) experience more severe toxicities versus younger patients [66]. A number of studies, including meta-analyses, have shown an increased risk of severe toxicity associated to fluoropyrimidine/platinum-based chemotherapy in female gastric and colorectal cancer patients [7, 36, 37, 67]. In accordance, we found a trend towards higher toxicity among females in our study (Table 2).

Compared to intravenous 5-FU, oral capecitabine (5-FU pro-drug) increases OS and response rates in combination with platinum compounds. Also, 5-FU/cisplatin is associated with greater toxicity [68–70]. In line with these findings, we observed that incorporation of certain regimens improved the predictive power of our models (Table 3). A recent study in colorectal cancer patients demonstrated that FOLFOX was associated with a significant increase in stomatitis and neutropenia, but decreased diarrhea and hand-foot syndrome versus CAPEOX [71].

Our model 2 includes the most relevant associations between selected SNPs and overall toxicity (Table 4). Several reports confirm $DPYD$ is a reaction-limiting enzyme for 5-FU catabolism. In fact, $DPYD$-deficiency is commonly associated with a lower drug-clearance and increased toxicity [72]. In our analysis, AG/GG $DPYD$ (rs1801265) genotypes were associated with higher grade 3–4 toxicity (OR = 4.20, $p = 0.002$). This variant causes a Cys 29 to Arg substitution that reduces $DPYD$ enzymatic activity and increases 5-FU-related toxicity [73]. Functional studies demonstrate that AG and GG genotypes of $DPYD$ (rs1801265) have a significantly lower 5-FU degradation rates (5-FUDR) compared to AA, with a profound effect for GG [73]. Likewise, the CT $DPYD$ (rs2297595) genotype was associated with grade 3–4 toxicity versus TT, although individually was not significant (OR = 2.71, $p = 0.21$). The same study reported that the
(A) 

Points

Age (above median)
No

Sex
Female

CMT peri-ope.
No

5-FU based
Yes

DPYD 6 + ABCC2
AA + CT/TT
AG/GG + CT/TT
AA + CC
AG/GG + CC

DPYD 3
CT/CC

Total points

Probability of Toxicity

(B) 

Nomogram score overall toxicity

Severe tox. probability

N | % Low Tox. | % High Tox.
--- | --- | ---
15 | 26.7% | 73.3%
31 | 51.6% | 48.4%
31 | 83.9% | 16.1%
16 | 93.8% | 6.2%

p < 0.0001

median = 79
IC95% (64.3 - 84.9)
n = 61

median = 115
IC95% (104.1 - 131)
n = 32

Fig. 2 (See legend on next page.)
CT DPYD (rs2297595) genotype had a significantly lower 5-FUDR versus TT [73]. On the other hand, the TC/CC DPYD (rs1801159) genotypes were associated with a lower probability of grade 3–4 overall toxicity versus the TT genotype (OR = 0.45, p = 0.071). However, studies on this polymorphism are somewhat inconsistent and some have reported an association with increased severe toxicity [28, 74, 75] or no association [76–78]. A potential explanation for our finding is the high frequency of the C allele in this subset, reaching 34% (Supplementary Table S8). In sharp contrast, European cohorts report 18.3% for the C allele (n = 157) [73]. Similarly, an Asian study reports a 27% frequency (n = 362) [75]. Notably, C allele frequency in the American population is 27%, whereas in East Asian, European, African and South Asian populations is 27, 19, 15 and 8% respectively [32]. Therefore, differences can be attributed to specific geographical/ethnic factors.

Our study also found an association between GSTP1/ERCC2 SNPs and overall toxicity. These are linked to the formation of DNA-adducts. The AG/GG GSTP1 (rs1695) genotypes were associated with a lower probability of grade 3–4 toxicity compared to AA. This “protective” role of the G allele has been previously reported in gastro-esophageal [79], colorectal [78], ovarian [80], testicular [41] and lung cancer [81]. A potential mechanism to explain this protective role could be the activation of the JNK pathway [79, 82] that increases cell defense mechanisms. Conversely, GT/GG genotypes in ERCC2 (rs13181) were associated with a higher probability of grade 3–4 overall toxicity (Table 4). These polymorphic variants decrease repair efficacy and may thus increase DNA adducts [83, 84]. This suggests that increased toxicity may be mediated by platinum damage to normal cells [85].

In line with previous publications, our study found a significant association between DPYD (rs1801265) SNPs and grade 3–4 overall toxicity only in male patients (Supplementary Table S4), [7]. Our paired-SNP analysis found a strong association between DPYD (rs1801265) and ABCC2 (rs717620) SNPs and overall toxicity (see Table 5). In particular, AG/GG (DPYD) and CT/TT (ABCC2) patients had a high probability of developing grade 3–4 overall toxicity (OR = 11.25, 95% CI = 1.25–245.45). The ABCC2 (rs717620) polymorphism is located in the promoter region of the gene, and has been previously associated with decreased protein expression in vitro [86]; ABCC2 also mediates the export/elimination of glutathione-oxalipatin conjugates [87] therefore an impaired function could decrease export of the drug leading to toxicity. Previous studies in colorectal and lung cancer [88, 89] have associated this polymorphism to severe fluoropyrimidine/platinum-related hematological toxicity. Thus, the DPYD (rs1801265)/ABCC2 (rs717620) SNP combination could potentiate fluoropyrimidines and/or oxalipatin derived toxicities. Again, given the SNP frequency in our analysis, this finding requires further validation by a larger cohort.

Utilizing a classification algorithm that involved 28 SNPs and 1 clinical variable (histology) a study by Yin et al. reported that the best prediction of toxicity was achieved in lung cancer patients that received platinum-based therapies [39]. Moreover, these authors demonstrated that the ABCG2 rs2231142-CESSA rs3859104 SNPs combination was strongly associated with grade 3–4 platinum toxicity (adjusted OR = 8.044, p = 4.350 × 10–5) [90]. In this regard, our models displayed better adjustments (based on Pseudo R²) after adding clinical/treatment factors and SNPs (models 3 and 4, Table 6). Our Model 4 was the best-fitted model in terms of sensitivity, specificity, accuracy and AUC (Fig. 1D). Previous studies have used also this strategy with consistent results [7, 40, 41, 91].

The frequency of grade 3–4 toxicity is observed in only 10–15% of gastric cancer patients as medical oncologists often make alterations to treatment protocols when lower toxicities start to manifest. The number of cases and controls incorporated into this study allowed statistically significant differences to be observed, however despite over 223 medical records being screened, we recognize as a limitation that the number of patients was a limiting factor in further interpreting our data and thus validation in a larger cohort is required before these models can be considered in a clinical setting. A future cohort will permit better precision analysis of the combinatorial SNPs. Certain SNPs, despite not achieving the individual significance standard of p < 0.05, were included in the models due to their influence or effect on the development of severe toxicity has been previously reported, or their inclusion improved the Pseudo R² values. Interestingly, according to the CPIC guide, the polymorphisms for DPYD rs1801265 (also known as DPYD*9A) and DPYD rs1801159 (also known as DPYD*5), which were both incorporated into our model,
are classified to not affect the DPD function in a clinically relevant manner in the context of 5-fluorouracil-related toxicity [30]. Interestingly, in accordance with our results, recent reports have shown an association with severe toxicity and high levels of 5-FU post-treatment for DPYD-rs1801265 [92–94].

A further option is a model / nomogram comparison with other similar studies. Schwab et al. [7] and Botticelli et al. [95] previously proposed nomograms to predict toxicity to 5-FU, and thus could be incorporated as a reference in future validation studies. A potential comment on this study could be the heterogeneity of treatments received by patients and that a prospective study considering a limited number of regimens would allow analyses the actual doses and duration of chemotherapy as potential variables associated with toxicity. While this is true, our gastric cancer patients and their treatments are a reflection of standard clinical practice. This heterogeneity in treatment was present despite the patients being part of the same recruiting clinical center and being treated by the same group of medical oncologists. Thus, this heterogeneity observed among our fluoropyrimidine and platinum-based treatments will always exist in the oncology clinic and thus any predictive model or algorithm will be required to be effective in face of this variable. The real-world treatment and clinical outcomes in this study have allowed a “proof of concept” of a model which integrates clinical and pharmacogenetic (three SNPs in two different genes) variables to improve the prediction of toxicity associated with fluoropyrimidine and platinum-based chemotherapy. Furthermore, recent studies have demonstrated that genetic elements outside the coding region of genes are potential regulators of pharmacokinetic and pharmacodynamic processes. In the TYMS gene, variants in UTR regions of 5 VNTR 28 bp-repeat (rs45445694) and 3’UTR 6 bp-indel (rs11280056) have been associated with severe toxicity in patients receiving fluoropyrimidine based treatments [77]. Furthermore, regulatory molecules of the non-coding RNA type, such as circular RNA (circRNAs) and Long Non-Coding RNA (lncRNAs), have been correlated to clinical variables (TNM stage, presence of metastasis and diagnosis) in patients with GC [96]. Interestingly, polymorphisms in lncRNAs of ANRIL (rs1333049) and MEG3 (rs116907618) genes were associated with severe overall and gastrointestinal toxicity in patients with lung cancer treated with platinum-based chemotherapy [97]. In addition, another regulator of gene expression are the Micro RNA (miRNAs), where genetic variations in miRNA binding sites are associated with an altered drug response [98]. In a similar vein, a recent publication by Powell et al. reported the mapping of miRNA-mRNA interactions in several pharmacogenes, in fact, the authors identified an hsa-mir-27b-DPYD interaction at a previously validated binding site, which may suggest the existence of additional elements that contribute to the individualism of drug response [99].

Given that the purpose of study is to predict toxicity to chemotherapy, it would be interesting in further validation studies to test our proposed clinical variables and individual and combinations of SNPs, together with emerging variables such as changes in expression, sequence, and binding sites of non-coding RNAs. This may give us a more complete picture on how to predict severe toxicity associated with chemotherapy and thus improve patient quality of life and survival.

Conclusions
In summary, in the absence of reliable markers and clinically relevant models to predict patient toxicity derived from fluoropyrimidine/platinum-based chemotherapy, herein we present for future validation a logistic regression-based model that integrates clinical, treatment and common SNPs.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12885-021-08745-0.

Additional file 1: Supplementary Material S1. Supplementary Fig. S1. Overall survival rates in the study cohort; Supplementary Fig. S2. Calibration plot for the prognostic model, Supplementary Table S1. Demographic and clinic-pathological characteristics of study population (N = 93), Supplementary Table S2. Platinum plus fluoropyrimidine-based chemotherapy combined treatments used in gastric cancer patients (N = 93), Supplementary Table S3. Grades of toxicity in gastric cancer patients by the Common Toxicity Criteria for Adverse Events (CTCAE) 4.0, Supplementary Table S4. Sex subgroup association analysis of the SNPs DPYD (rs1801265), Supplementary Table S5. Models for hematological grade ≥ 3 toxicity in gastric cancer patients treated with platinum/flouropyridines-based chemotherapy using multivariate analysis. Supplementary Table S6. Models for gastrointestinal grade ≥ 3 toxicity in gastric cancer patients treated with platinum/flouropyridines-based chemotherapy using multivariate analysis. Supplementary Table S7. Models for neurological grade ≥ 3 toxicity in gastric cancer patients treated with platinum/flouropyridines-based chemotherapy using multivariate analysis. Supplementary Table S8. Genotypic and allelic frequencies for the analyzed polymorphisms. Supplementary Table S9. ID assay for each of the analyzed polymorphisms. Supplementary Table S10. SNPs selection based in score for fluoropyrimidines. Supplementary Table S11. Final score for fluoropyrimidines. Supplementary Table S12. SNPs selection based in score for platinum. Supplementary Table S13. Final score for platinum. Supplementary Table S14. Sensitivity, specificity and accuracy calculations from a 2 × 2 confusion matrix. Supplementary Methods: Details of chemotherapy schemes, SNPs selection and classifications algorithm used. Supplementary Data File 1. All raw data used in this study.

Acknowledgments
We would like to thank the patients and their families for participating in this study, to the RED UC Christus and the department of pathology at the “Hospital Clinico UC” for the access to patient data and paraffin embedded samples and to the Biomedical Research Consortium of Chile for the initial support in this project.
Authors’ contributions
All authors made substantial contributions to the conception and design of this research. A detail of specific contributions is shown. Investigation and experimentation: MCD, MLB, EC, CNH, MMM, INR, MAL, MRF, YL, AHC. Conceptualization: MCD, MPP; JFM, ZL, RA, MG, LAQ, GIO. Funding acquisition: AHC, MG, GIO. All authors have read and agreed to the published version of the manuscript.

Funding
This study was funded by the grants issued by the Government of Chile: CONICYT FONDAP-15130011 (GIO), BMRC CTU66 (GIO), Millennium Institute on Immunology & Immunotherapy IMI09/016-F (GIO), FONDECYT 1180241 (GIO), FONDECYT 1180173 (MG) and FONDECYT 1191928 (AHC). The first author (MCD) is a CONICYT doctoral scholarship recipient (#21150695).

Availability of data and materials
All data generated or analyzed during this study are included in this published article and it supplementary information files.

Declarations

Ethics approval and consent to participate
The Internal Review Board and the Ethics and Scientific Committee at the School of Medicine of the Pontificia Universidad Católica de Chile approved this research (CEC MED UC approval number 16–046, resolution dated April 21, 2016) [4]. All participants signed an informed consent to participate in this study. Also, Internal Review Board and the Ethics and Scientific Committee at the School of Medicine of the Pontificia Universidad Católica de Chile granted the waiver for the informed consent and the permission to access patients’ electronic medical records to include deceased patients. All data were anonymized to protect patients’ privacy. This study strictly adhered to the Code of Ethics of the World Medical Association (Declaration of Helsinki, 1964).

Consent for publication
Not applicable.

Competing interests
MG is an advisor for Merck Sharp & Dohme and Novartis, participates at the speakers’ bureau for Bristol-Myers & Squibb and Bayer. Also receives research funding from Bristol-Myers & Squibb (Inst) and Novartis (Inst). Also, has received travel accommodations and expenses from Roche. The funders had no role in the design of the study; in the analysis, opinions, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. The other authors have no conflicts of interest to declare.

Author details
1Faculty of Chemical and Pharmaceutical Sciences, Universidad de Chile, 8380494 Santiago, Chile. 2Department of Physiology, Faculty of Biological Sciences, Pontificia Universidad Católica de Chile, 8331150 Santiago, Chile. 3Department of Hematology and Oncology, Faculty of Medicine, Pontificia Universidad Católica de Chile, 8330032 Santiago, Chile. 4Millennium Institute on Immunology and Immunotherapy, 8331150 Santiago, Chile. 5Faculty of Dentistry, Universidad de Los Andes, 7620001 Santiago, Chile. 6Laboratory of Chemical Carcinogenesis and Pharmacogenetics, Department of Basic and Clinical Oncology, Faculty of Medicine, Universidad de Chile, 8380494 Santiago, Chile. 7Latin American Network for Implementation and Validation of Clinical Pharmacogenomics Guidelines (REILIVAF-CYTED), Madrid, Spain. 8Escuela de Química y Farmacia, Facultad de Ciencias Médicas, Universidad Bernardo O’Higgins, Santiago, Chile. 9Department of Gastroenterology, Faculty of Medicine, Pontificia Universidad Católica de Chile, 8330032, Santiago, Chile. 10Institute for Biological and Medical Engineering, Schools of Engineering, Medicine and Biological Sciences, Pontificia Universidad Católica de Chile, Santiago, Chile. 11Central Laboratory, Yangjiang People’s Hospital, Guangdong Province, Yangjiang, China. 12Center of Genome and Personalized Medicine, Institute of Cancer Stem Cell, Dalian Medical University, Dalian, China. 13National Institute on Aging, National Institute of Health, Baltimore, USA. 14Advanced Center for Chronic Diseases (ACCDIS), 8330034 Santiago, Chile. 15Instituto de Ciencias e Innovación en Medicina, Facultad de Medicina, Clínica Alemana, Universidad del Desarrollo, 7590943 Santiago, Chile.

References
1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424. https://doi.org/10.3322/caac.21492.
2. Ferro A, Peleteiro B, Malvezzi M, Bosetti C, Bertuccio P, Levi F, et al. Worldwide trends in gastric cancer mortality (1980–2011), with predictions to 2015, and incidence by subtype. Eur J Cancer. 2014;50(7):1330–44. https://doi.org/10.1016/j.ejca.2014.01.029.
3. de la lla JI, Bastari G, Ferreccio C, Moscoso C, Sagués S, Cid C, et al. A snapshot of cancer in Chile: analytical frameworks for developing a cancer policy. Biol Res. 2015;48:10.
4. Owen GI, Pinto MP, Retamal IN, Fernández MF, Cisternas B, Mondaca S, et al. Chilean Gastric Cancer Task Force: A study protocol to obtain a clinical and molecular classification of a cohort of gastric cancer patients. Medicine (Baltimore). 2018;97:e0419.
5. Cammona-Bayonas A, Jiménez-Fonseca P, Lorenzo MLS, Ramchandani A, Martínez EA, Custodio A, et al. On the effect of triplet or doublet chemotherapy in advanced gastric Cancer: results from a National Cancer Registry. J Natl Compr Canc Netw. 2016;14(11):1379–88. https://doi.org/10.6004/jnccn.2016.0148.
6. Park H, Jin RU, Wang-Gillam A, Suresh R, Rigden C, Amin M, et al. FOLFRINOX for the treatment of advanced Gastrointestinal cancers: a phase 2 nonrandomized clinical trial. JAMA Oncol. 2020;6(8):1231–40. https://doi.org/10.1001/jamaoncol.2020.2020.
7. Schwab M, Zanger UM, Marx C, Schaeffeler E, Klein K, Dippon J, et al. Role of genetic and nongenetic factors for fluorouracil treatment-related severe toxicity: a prospective clinical trial by the German FU toxicity study group. J Clin Oncol. 2008;26(13):2151–8. https://doi.org/10.1200/JCO.2006.14.182.
8. López Sobella M, Criado Ilana MT, Esteban Herrera B, Lopez Arranz MC. Severe 5-fluorouracil induced toxicity associated with dihydropterdine dehydrogenase deficiency. Farm Hosp. 2008;32(1):54–6. http://www.ncbi.nlm.nih.gov/pubmed/18426703. https://doi.org/10.11067/s0034-98872008-01-0-1.
9. Pachman DR, Qin R, Seisler DK, Smith EM, Beutler AS, Ta LE, et al. Clinical course of Oxaliplatin-induced neuropathy: results from the randomized phase III trial NO8CB (alliance). J Clin Oncol. 2015;33(30):3416–22. https://doi.org/10.1200/JCO.2014.58.8533.
10. Erichsen HC, Chanock SJ. SNPs in cancer research and treatment. Br J Cancer. 2004;90(4):477–51. https://doi.org/10.1038/sj.bjc.6601574.
11. Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. Science. 1999;286:483–91.
12. Lopez-Cortes A, Guerrero S, Redal MA, Alvarado AT, Quiliones LA. State of art of cancer pharmacogenomics in Latin American populations. Int J Mol Sci. 2017;18(6). https://doi.org/10.3390/ijms18066369.
13. Quiliones L, Roco A, Cayín JP, Escalante P, Miranda C, Varela N, et al. Clinical applications of pharmacogenomics. Rev Med Chil. 2017;145(4):483–500. https://doi.org/10.4067/S0034-98872017000400009.
14. Relling MV, Evans WE. Pharmacogenomics in the clinic. Nature. 2015; 526(7573):343–50. https://doi.org/10.1038/nature15817.
15. Wang D, Lippard SJ. Cellular processing of platinum anticancer drugs. Nat Rev Drug Discov. 2005;4(4):307–20. https://doi.org/10.1038/nrd1691.
16. Sakaeda T, Nakamura T, Okumura K. MRDR genotype-related pharmacokinetics and pharmacodynamics. Biol Pharm Bull. 2002;25(11):1391–600. http://www.ncbi.nlm.nih.gov/pubmed/12419946. https://doi.org/10.1248/bpb.25.1391.
17. Scripture CD, Figg WD. Drug interactions in cancer therapy. Nat Rev Cancer. 2006;6(7):546–58. https://doi.org/10.1038/nrc1887.
18. Sun Y, Pan J, Tong X, Chen E, Yan W, Wu M, et al. Glutathione S-transferases genes variants and chemotherapy efficacy in gastrointestinal cancer patients: a meta-analysis based on 50 pharmacogenetic studies. J Cancer. 2019;10(13):2915–26. https://doi.org/10.7150/jca.31130.
19. Marsh S, McLeod H, Dolan E, Shukla SJ, Rabik CA, Gong L, et al. Platinum pathway. Pharmacogenet Genomics. 2009;19(7):563–4. https://doi.org/10.1097/FCG.0b013e32832e0ed7.
20. Palmiotto R, Carella C, Silvestris E, Cives M, Stucci SL, Tucci M, et al. SNPs in predicting clinical efficacy and toxicity of chemotherapy: walking through
92. Varma KA, Jayanthi M, Dubashi B, Shewade DG. Influence of DPYD*9A, DPYD*6 and GSTP1 ile105val genetic polymorphisms on Capecitabine and Oxaliplatin (CAPOX) associated toxicities in colorectal cancer (CRC) patients. Asian Pac J Cancer Prev. 2019;20(10):3093–100. https://doi.org/10.31557/APJCP.2019.20.10.3093.

93. Khushman M, Patel GK, Hosein PJ, Laurini JA, Cameron D, Clarkson DR, et al. Germline pharmacogenomics of DPYD*9A (c.85T>C) variant in patients with gastrointestinal malignancies treated with fluoropyrimidines. J Gastrointest Oncol. 2018;9:416–24.

94. Varma A, Jayanthi M, Dubashi B, Shewade DG, Sundaram R. Genetic influence of DPYD*9A polymorphism on plasma levels of 5-fluorouracil and subsequent toxicity after oral administration of capecitabine in colorectal cancer patients of south Indian origin. Drug Metab Pers Ther. 2020;0(0). https://doi.org/10.1515/dmpt-2020-0133.

95. Botticelli A, Onesti CE, Strigari L, Occhipinti M, Di Pietro FR, Cerbelli B, et al. A nomogram to predict 5-fluorouracil toxicity: when pharmacogenomics meets the patient. Anti-Cancer Drugs. 2017;28. https://doi.org/10.1097/CAD.0000000000000492.

96. Russano M, Napolitano A, Ribelli G, Iuliani M, Simonetti S, Citarella F, et al. Liquid biopsy and tumor heterogeneity in metastatic solid tumors: the potentiality of blood samples. J Exp Clin Cancer Res. 2020;39(1):95. https://doi.org/10.1186/s13046-020-01601-2.

97. Gong W, Peng J, Yin J, Li X, Zheng W, Xiao L, et al. Association between well-characterized lung cancer IncRNA polymorphisms and platinum-based chemotherapy toxicity in Chinese patients with lung cancer. Acta Pharmacol Sin. 2017;38(4):581–90. https://doi.org/10.1038/aps.2016.164.

98. Leone P, Buonavoglia A, Fasano R, Solimando AG, De Re V, Cicco S, et al. Insights into the regulation of tumor angiogenesis by Micro-RNAs. J Clin Med. 2019;8(2). https://doi.org/10.3390/jcm81202030.

99. Powell NR, Zhao H, Ipe J, Liu Y, Skaar TC. Mapping the miRNA-mRNA Interactome in Human Hepatocytes and Identification of Functional miRSNPs in Pharmacogenes. Clin Pharmacol Ther. 2021;n/a n/a; 10.1002/cpt.2379.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.