Journal of Pharmacological Sciences 140 (2019) 20–25

Contents lists available at ScienceDirect

Journal of Pharmacological Sciences

journal homepage: www.elsevier.com/locate/jphs

Full Paper

Prediction of irinotecan toxicity in metastatic colorectal cancer patients based on machine learning models with pharmacokinetic parameters

Esther Oyaga-Iriarte [a, *], Asier Insausti [a], Onintza Sayar [a], Azucena Aldaz [b]

[a] Pharmamodelling S.L., Pamplona, Spain
[b] Department of Hospital Pharmacy, Clínica Universidad de Navarra, Pío XII 36, Pamplona, Spain

ARTICLE INFO

Article history:
Received 18 December 2018
Received in revised form
23 February 2019
Accepted 25 March 2019
Available online 4 May 2019

Keywords:
Colorectal cancer
Irinotecan
Machine learning
Pharmacokinetics
Toxicity

ABSTRACT

Irinotecan (CPT-11) is a drug used against a wide variety of tumors, which can cause severe toxicity, possibly leading to the delay or suspension of the cycle, with the consequent impact on the prognosis of survival. The main goal of this work is to predict the toxicities derived from CPT-11 using artificial intelligence methods.

The data for this study is conformed of 53 cycles of FOLFIRINOX, corresponding to patients with metastatic colorectal cancer. Supported by several demographic data, blood markers and pharmacokinetic parameters resulting from a non-compartmental pharmacokinetic study of CPT-11 and its metabolites (SN-38 and SN-38-G), we use machine learning techniques to predict high degrees of different toxicities (leukopenia, neutropenia and diarrhea) in new patients.

We predict high degree of leukopenia with an accuracy of 76%, neutropenia with 75% and diarrhea with 91%. Among other variables, this study shows that the areas under the curve of CPT-11, SN-38 and SN-38-G play a relevant role in the prediction of the studied toxicities.

The presented models allow to predict the degree of toxicity for each cycle of treatment according to the particularities of each patient.

1. Introduction

Irinotecan (CPT-11) with 5 fluorouracil (5-FU), oxaliplatin and folinic acid conform FOLFIRINOX, a frequently used treatment for metastatic colorectal carcinoma and several other tumors (pancreatic cancer, gastric cancer, non-small cell lung cancer, etc.). According to different studies,[1,2] the usual dose in the FOLFIRINOX scheme of CPT-11 is 180 mg/m^2. Posology adjustments are usually made based on the genetic analysis of the UGT1A1 isoform.[3]

The mechanism of action of CPT-11 is the inhibition of the enzyme topoisomerase I, responsible for the replication and transcription of DNA structure. This inhibition causes irreversible defects in the DNA, resulting in cell death.[4,5]

CPT-11 is hydrolyzed to its active potent metabolite, SN-38, mainly on liver tissues and gastrointestinal tract.[6] Next, SN-38 is detoxified into its glucuronide derivative, SN-38-G, by the action of the UDP-glucuronyl transferase system. Finally, after SN-38-G is concentrated in bile and released into the intestinal lumen, most of it is excreted. However, an enterohepatic circulation is produced, reconverting the remaining SN-38-G to SN-38 and allowing it to be reabsorbed. This process is responsible for the characteristic rebounds in the concentration/time profile of the metabolites of this drug and has a great impact in the pharmacodynamics of CPT-11. One of the main adverse effects is gastrointestinal toxicity,[7] caused by a slow transformation of SN-38 to SN-38-G. On occasion, these toxicity episodes are so serious that can lead to treatment suspension. Aside gastrointestinal toxicity, there are other adverse events related to other toxicities that can affect the treatment, such as leukopenia[8] and neutropenia.[9]

For all these reasons and the fact that CPT-11 has a wide inter-individual pharmacokinetic variability, the personalization of this

* Corresponding author.
E-mail address: eoyaga@pharmamodelling.com (E. Oyaga-Iriarte).
Peer review under responsibility of Japanese Pharmacological Society.

https://doi.org/10.1016/j.jphs.2019.03.004

treatment is necessary to guarantee a toxicity-free optimal pharmacotherapy.

This work aimed at predicting the grade of toxicity (specifically, diarrhea, leukopenia and neutropenia) that a patient will suffer once the CPT-11 dose is administered using a machine learning approach.

2. Materials and methods

2.1. Patients

This work was supported with data from the Service of Hospital Pharmacy, Clínica Universidad de Navarra. The inclusion criteria for this study were the following: age ≥18, ECOG ≤2, presence of measurable lesions, life expectancy ≥3 months, time since last cycle of chemotherapy ≥1 month or time since major surgical procedure ≥1 week, leukocytes >3/pL, platelets >75/pL, hemoglobin >9 g/dL and serum creatinine <2 mg/dL or 24-hour urine creatinine clearance >50 mL/min.

The population was formed from 20 patients with advanced colorectal cancer, adding up to 53 cycles of chemotherapy. The treatment followed the FOLFIRINOX scheme in accordance with two regimens, depending on whether the disease was exclusively or predominantly hepatic (first regimen) or not (second regimen). The first regimen consisted in doses of 2500 mg/m^2 of 5-FU (24 h long intravenous infusions on first four days of cycle), 100 mg/m^2 of oxaliplatin (120 min long intravenous infusions on fourth day of cycle) and doses of 250 mg/m^2 of CPT-11 (90 min long intravenous infusions on fourth day of cycle). The second regimen consisted in doses of 2600 mg/m^2 of 5-FU and 500 mg/m^2 of leucovorin (24 h long intravenous infusions on first and fifteenth days of cycle), 100 mg/m^2 of oxaliplatin (120 min long intravenous infusions on first day of cycle) and doses of 250 mg/m^2 of CPT-11 (90 min long intravenous infusions on first day of cycle). This scheme was administered every 28 days until disease progression.

Table 1 summarizes the characteristics of the population prior to each cycle of treatment. In Table 2 we can see the number of cycles in which each toxicity was presented according to its grade (described in Common Toxicity Criteria, CTC).

This observational study was approved by the University Clinic of Navarra ensuring compliance with ethical standards.

Table 1
Baseline characteristics of the patients before each cycle.

Patient characteristics	Descriptive results	
Age (years)	59.00	(52.00–66.00)
Body surface (m^2)	1.75	(1.67–1.92)
Weight (Kg)	70.00	(64.00–80.20)
Irinotecan dose (mg)	455.00	(412.00–531.00)
CA 19.9 (UI/mL)	52.90	(22.60–373.88)
AST (UI/L)	19.00	(11.75–29.00)
ALT (UI/L)	20.50	(12.00–45.25)
ALP (IU/L)	196.00	(164.00–323.50)
GGTP (UI/L)	40.50	(23.25–93.00)
DBil (mg/dL)	0.14	(0.10–0.18)
TBil (mg/dL)	0.55	(0.37–0.74)
Hb (g/dL)	11.60	(10.93–12.78)
Ht (%)	35.40	(33.53–39.33)
MCV (fL)	91.30	(82.75–94.68)
MCH (pg/cell)	29.85	(26.68–31.15)

The variables are represented by median (IQR). IQR: Interquartile range. CEA: Carcinoembryonic antigen. CA 19.9: cancer antigen 19.9. AST: Aspartate aminotransferase. ALT: Alanine aminotransferase. ALP: Alkaline phosphatase. GGTP: Gamma glutamyl transpeptidase. DBil: Direct bilirubin. TBil: Total bilirubin. Hb: Hemoglobin. Ht: Hematocrit. MCV: Mean corpuscular volume. MCH: Mean corpuscular hemoglobin.

Table 2
Number of cycles per toxicity type and grade.

Toxicity	Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
Leukopenia	5	3	10	13	2
Neutropenia	4	3	3	3	11
Diarrhea	12	5	7	5	6

2.2. Pharmacokinetic parameters

Blood samples were collected 30 and 60 min after the beginning of infusion and 5, 15, 30, 60, 120, 240, 360, 720, 1440 and 1800 min after the drug administration was concluded. Plasma samples were drawn in Venoject® heparinized tubes and centrifuged for five minutes at 3000 rpm. Then, they were frozen at −30 °C until they were analyzed.

The method to quantify CPT-11, SN-38 and SN-38-G was High Performance Liquid Chromatography (HPLC), with the methodology described in [10,11].

After the quantification of CPT-11 and its metabolites (SN-38 and SN-38-G), we carried out a non-compartmental pharmacokinetic analysis, using the software Phoenix WinNonlin 8, to obtain the pharmacokinetic parameters for each treatment cycle. These parameters were: maximum concentration (Cmax), maximum time (Tmax) and area under the curve (AUC) (see Table 3 for a descriptive summary).

2.3. Machine learning and statistical analysis

We used demographic data, liver function tests and tumor markers in Table 1, combined with the aforementioned pharmacokinetic parameters, to predict the grade of each toxicity type.

The five grades of toxicity were grouped in two for prediction: low degree of toxicity (grades 0, 1 and 2) and high degree of toxicity (grades 3 and 4).

We implemented four machine learning classification algorithms to predict toxicities after each treatment cycle. The first one was Backward Stepwise Logistic Regression (BSLR) with the inclusion of interactions between variables and non-linearities[12] and we used Akaike information criterion for variable selection.[13] This model is framed within generalized linear models (GLM) and allows to predict the probability of a dichotomous event. Secondly, the C4.5 algorithm[14] was implemented. This algorithm is a decision tree that starts with a single node and then branches into possible outcomes depending on the different variables and the relations among them, based on the information gain criterion. This process is repeated successively, giving it a shape which resembles a tree. The third technique was Random Forest (RF),[15] which is an ensemble of predictive trees with a certain degree of randomness, i.e., the final result is a combination of the predictions of each individual tree. Lastly, we implemented the method called Support Vector Machine (SVM).[16] It consists in constructing a hyperplane in a space of high dimension that permits to classify an event. In this process, a grid search of different kernels (linear, polynomial and radial) and hyperparameters was performed to fine tune the model.

Table 3
Means and deviations of the pharmacokinetic parameters of CPT-11 and its metabolites.

	CPT-11	SN-38	SN-38-G
Cmax (μg/mL)	2.50 (0.92)	0.035 (0.02)	0.11 (0.04)
Tmax (h)	1.58 (0.39)	1.67 (0.7)	2.14 (0.83)
AUC (μg h/mL)	12.59 (7.16)	0.22 (0.16)	1.02 (0.79)

Table 4
Results of the different classification models (BSLR, C4.5, RF and SVM) for each toxicity (leukopenia, neutropenia and diarrhea) in terms of the following indicators: accuracy, specificity, sensitivity, F1 score, positive likelihood rate (LR_+), negative likelihood rate (LR_-) and area under the ROC curve (AUC-ROC). The results of the best classifier for each toxicity are in bold.

Toxicity	Indicator	BSLR	C4.5	RF	SVM
Leukopenia	Accuracy	0.61	0.42	**0.76**	0.55
	Specificity	0.46	0.46	**0.60**	0.46
	Sensitivity	0.72	0.38	**0.89**	0.61
	F1 score	0.67	0.42	**0.8**	0.59
	LR_+	1.33	0.70	**2.22**	1.13
	LR_-	0.61	1.35	**0.18**	0.85
	AUC-ROC	0.67	0.71	**0.74**	0.54
Neutropenia	Accuracy	0.71	0.58	0.58	**0.75**
	Specificity	0.64	0.59	0.71	**0.79**
	Sensitivity	0.80	0.20	0.40	**0.70**
	F1 score	0.70	0.29	0.44	**0.70**
	LR_+	2.22	0.49	1.38	**3.33**
	LR_-	0.31	1.36	0.85	**0.38**
	AUC-ROC	0.75	0.5	0.56	**0.88**
Diarrhea	Accuracy	**0.91**	0.51	0.74	0.66
	Specificity	**1.00**	0.00	0.36	0.18
	Sensitivity	**0.88**	0.75	0.92	0.88
	F1 score	**0.93**	0.68	0.83	0.78
	LR_+	∞	0.75	1.44	1.07
	LR_-	**0.12**	∞	0.22	0.66
	AUC-ROC	**0.95**	0.50	0.64	0.63

Once the optimal predictive models for each toxicity were selected, we studied which of the variables were the most relevant for each model according to their influence in the prediction. Subsequently, we carried out a statistical analysis to see whether there were statistically significant differences between the two classes (high degree or low degree of toxicity). To this account, Kolmogorov–Smirnov, T-student, U-Mann Whitney tests were employed with significance level set to $\alpha = 0.05$ and boxplot diagrams were used for visualization.

2.4. Model diagnostics

We used 5-fold cross-validation[17] to validate the results of BSLR, C4.5 and SVM. The validation of RF was based on bagging,[18] which is a validation procedure embedded in the algorithm.

To select the optimal model, we analyzed different indicators. Specifically, accuracy, specificity, sensitivity, F1 score, positive likelihood rate (LR_+), negative likelihood rate (LR_-) and area under the ROC curve (AUC-ROC).

3. Results

The results of each machine learning algorithm to predict if a patient is going to suffer from high or low degree of toxicity can be seen in Table 4.

Table 5
Most relevant variables for leukopenia, neutropenia and diarrhea according to the models RF, SVM and BSLR, respectively.

Leukopenia	Neutropenia	Diarrhea
$AUC_{SN-38-G}$	Basal DBil	AUC_{SN-38}
Irinotecan dose	$AUC_{SN-38-G}$	Basal MCH
	Basal TBil	Weight
	Basal CA 19.9	Basal AST
	AUC_{CPT-11}	Basal GGTP
	Basal GGTP	

Based on the results of Table 4, the optimal technique for the prediction of leukopenia was RF. For the prediction of neutropenia, SVM with a radial kernel, cost = 1 and $\gamma = 0.02$ stood out. Finally, in the prediction of diarrhea, the method that overcame the rest was BSLR.

In Fig. 1, we can see the ROC curves of the best models for each toxicity. On the one hand, in the case of RF (Fig. 1 (a)), we show the average ROC curve of each decision tree that was generated in the bagging process. On the other hand, in the cases of SVM and BSLR (Fig. 1(b) and (c)), we show a different ROC curve for each validation group in the 5-fold cross-validation process.

The implemented machine learning models allowed to estimate the importance of variables for each toxicity. In the case of RF and SVM, all variables were employed in the prediction, each one with a different weight in the final outcome. On the other hand, BSLR carried out a selection of the most relevant variables during its construction. In Table 5 the most influential variables for the prediction of each toxicity can be seen, sorted by their relevance.

All relevant data in the prediction of each toxicity followed a non-normal distribution with the exception of basal TBil ($p = 0.2$). The results derived from the T-student and U-Mann Whitney tests

Table 6
Results of the statistical analyses to check for significant differences between classes for each of the most relevant variables in each toxicity. P-values reaching statistical significance are in bold. T: T-student test. U: U-Mann Whitney test.

Toxicity	Variable	Test	p-values	Significant differences
Leukopenia	$AUC_{SN-38-G}$	U	0.108	No
	Irinotecan dose	U	**0.033**	Yes
Neutropenia	Basal DBil	U	0.108	No
	$AUC_{SN-38-G}$	U	**0.002**	Yes
	Basal TBil	T	0.06	No
	Basal CA 19.9	U	0.056	No
	AUC_{CPT-11}	U	**0.011**	Yes
	Basal GGTP	U	0.108	No
Diarrhea	AUC_{SN-38}	U	0.370	No
	Basal MCH	U	0.268	No
	Weight	U	0.067	No
	Basal AST	U	0.903	No
	Basal GGTP	U	0.612	No

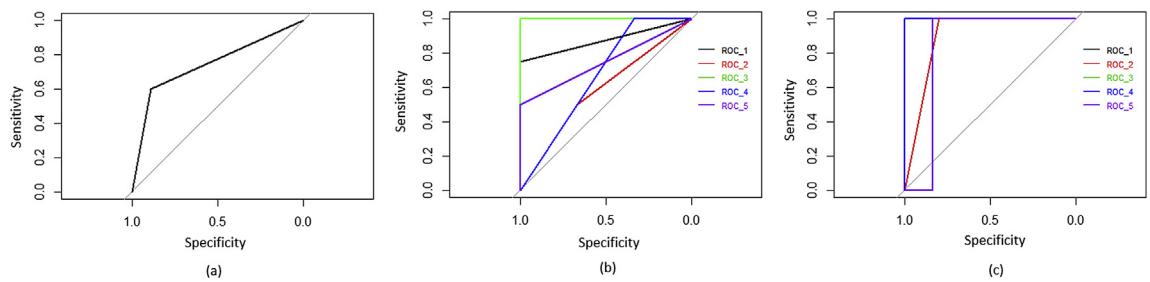

Fig. 1. ROC curves of the best models for each toxicity. (a) Leukopenia: average ROC curve of all decision trees in the bagging process of RF. (b) Neutropenia: a ROC curve for each validation group in the 5-fold cross-validation of SVM. (c) Diarrhea: a ROC curve for each validation group in the 5-fold cross-validation of BSLR.

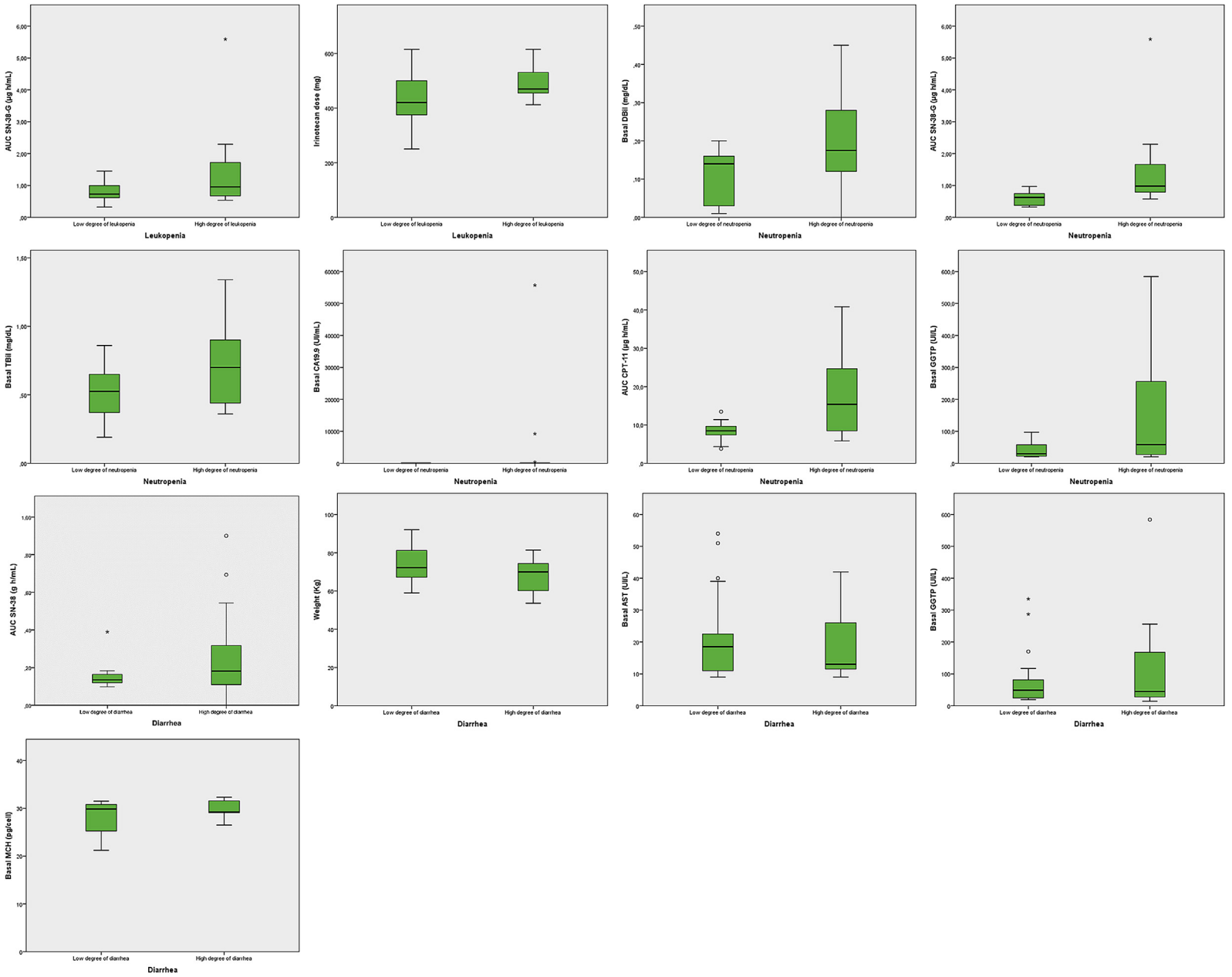

Fig. 2. Boxplots of the most relevant variables in the prediction of each toxicity (see Table 5) for the two classes: low degree of toxicity (left boxplot of each graph) and high degree of toxicity (right boxplot of each graph).

for each of the most relevant variables that described the three toxicities can be seen in Table 6.

Fig. 2 shows the distribution of each variable in Table 5 according to the degree of toxicity that it presented.

4. Discussion

Machine learning techniques have been applied in the literature in the area of health. For instance, in prevention, survival and mortality prediction and cell identification of cancer,[19–22] in early diagnosis of a variety of diseases[23,24] and, in the pharmacological field, for example, in preclinical studies,[25] in drug design[26–28] and in medication adherence.[29] Specifically, in relation to the prediction of toxicities, there are also works in the literature that use machine learning techniques. For example, Modongo et al.[30] implemented a CART prediction tree to predict ototoxicity from pharmacokinetic parameters and other covariates in multidrug resistant tuberculosis patients and Yamazaki et al.[31] developed an algorithm for the prediction of drug induced proarrhythmia.

In this work we predict whether a patient with metastatic colorectal cancer will have high degree of treatment derived toxicity using different personal characteristics. As far as we know, it is the first time that machine learning techniques are used to predict CPT-11 toxicities in colorectal cancer.

In the present day, pharmacokinetic/pharmacodynamic (PK/PD) models and pharmacogenetics are the tools used to predict toxicities in oncological treatments.[32–34] The methodology proposed in this paper, which has been validated in multidisciplinary areas, is a new trend in this field of research. Hence, the results of this work provide a new (and complementary) method to predict toxicities from PK parameters accurately.

In the FOLFIRINOX scheme, being a combination treatment, other components such as 5-FU or oxaliplatin can also produce hematological toxicity. The inclusion of PK parameters of 5-FU and oxaliplatin could lead to improvements in the performance of the models. This fact is patent in the prediction of late diarrhea, which is a toxicity derived solely from CPT-11 and its metabolites, where the accuracy rate (91%) was higher than those of the predictions of leukopenia and neutropenia (76 and 75%, respectively). However, the proposed models permitted to characterize the three studied toxicities solely from the CPT-11 monitoring and the inclusion of anthropometric, analytical and biochemical covariates (Table 1), which are related to the pharmacokinetics of 5-FU and oxaliplatin[35–37] and compensate the lack of PK parameters of these drugs.

In the previous section, Table 4 shows the degree of certainty with which the different applied techniques can predict each of the studied toxicities. In fact, the values of accuracy, using the best technique for each toxicity, went up to 76%, 75% and 91% for leukopenia, neutropenia and diarrhea, respectively. Specificity values were 0.6 (leukopenia), 0.79 (neutropenia) and 1 (diarrhea). These values are clinically relevant because they are related to false negatives, i.e., the number of patients that are classified in the low degree toxicity group, but in fact they will suffer high degrees of toxicity. The case of diarrhea is particularly significant because the specificity value was the highest possible, since the mathematical model did not yield any false negatives. Moreover, F1 score values were 0.8 (leukopenia), 0.7 (neutropenia) and 0.93 (diarrhea), which indicates that the models were correctly balanced.

In the predictions of these three toxicities, the pharmacokinetic parameters of CPT-11, SN-38 and SN-38-G appeared as fundamental variables, as in[38] and,[39] where it was demonstrated that glucuronidation was correlated with diarrhea and neutropenia, respectively. Additionally, several authors have linked high levels of basal bilirubin with high risk of developing grades 3 or 4 of neutropenia.[40,41] In our case, this fact was verified as Basal DBil and Basal TBil were two of the relevant variables in the prediction of neutropenia, along with other hepatic covariates, such as AST and GGTP. Finally, in[42] the dose of CPT-11 was shown to be related to severe diarrhea and, in this work, we have found that the dose of this drug is related to high degrees of leukopenia as well.

This work also shows that the combination of multiple variables enables to obtain more accurate results than those of an individualized analysis of each variable, where only three of the variables presented statistically significant differences.

The utility of the models presented in this work resides in the possibility of knowing if a new patient will suffer from toxicity once the dose is administered. Thus, the clinician is able to anticipate to any of the studied toxicities and make decisions accordingly; both to treat such toxicities and to adjust the dosing for subsequent cycles, in combination with additional information derived from pharmacogenetics[43] and pharmacokinetics.[44]

This work is limited by the size of the population sample and, although the employed validation techniques have been applied in other published works with accredited reliability when the available cohort is small,[21,23–25,30] this fact should be considered when applying these models in clinical practice. Hence, the adequacy of the algorithms to the target population must be checked prior to their application.

5. Conclusions

This study has proposed a machine learning based methodology to predict whether a patient will suffer a high degree of leukopenia, neutropenia and diarrhea after CPT-11 administration. These models give the medical practitioner prior knowledge about the grade of toxicity that a patient might suffer with high accuracy and, thus, they make it possible to take the appropriate measures to achieve optimal pharmacotherapy.

The pharmacokinetic parameters of CPT-11, SN-38 and SN-38-G, specifically AUC, are relevant variables for the prediction of toxicities. It has been determined that $AUC_{CPT\text{-}11}$ has a great impact in the prediction of neutropenia, $AUC_{SN\text{-}38}$ in the prediction of diarrhea and $AUC_{SN\text{-}38\text{-}G}$ in the prediction of both leukopenia and neutropenia. Therefore, controlling the pharmacokinetic parameters of patients is a key factor in the prevention of the CPT-11 derived toxicities.

Funding

This work is partially supported by “Ayuda para Doctorados Industriales del Ministerio de Economía, Industria y Competitividad” (Ref. DI-15-07511).

Conflict of interest

The authors declare that they have no conflicts of interest.

References

1. Conroy T, Desseigne F, Ychou M, et al. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. *N Engl J Med*. 2011;364:1817–1825.
2. Muranaka T, Kuwatani M, Komatsu Y, et al. Comparison of efficacy and toxicity of FOLFIRINOX and gemcitabine with nab-paclitaxel in unresectable pancreatic cancer. *J Gastrointest Oncol*. 2017;8:566–571. https://doi.org/10.21037/jgo.2017.02.02.
3. Oeser SG, Bingham JP, Collier AC. Regulation of hepatic UGT2B15 by methylation in adults of Asian descent. *Pharmaceutics*. 2018;10:1–14. https://doi.org/10.3390/pharmaceutics10010006.

4. Liu LF, Duann P, Lin CT, D'Arpa P, Wu J. Mechanism of action of camptothecin. *Ann N Y Acad Sci*. 1996;803:44–49. https://doi.org/10.1111/j.1749-6632.1996.tb26375.x.
5. Xu Y, Villalona-Calero MA. Irinotecan: mechanisms of tumor resistance and novel strategies for modulating its activity. *Ann Oncol*. 2002;13:1841–1851. https://doi.org/10.1093/annonc/mdf337.
6. Burris H, Fields S. Topoisomerase I inhibitors. An overview of the camptothecin analogs. *Hematol Oncol Clin N Am*. 1994;8:333—355.
7. Dranitsaris G, Shah A, Spirovski B, Vincent M. Severe diarrhea in patients with advanced-stage colorectal cancer receiving FOLFOX or FOLFIRI chemotherapy: the development of a risk prediction tool. *Clin Colorectal Cancer*. 2007;6: 367–373. https://doi.org/10.3816/CCC.2007.n.006.
8. Shiozawa T, Tadokoro J Ichi, Fujiki T, et al. Risk factors for severe adverse effects and treatment-related deaths in Japanese patients treated with irinotecan-based chemotherapy: a postmarketing survey. *Jpn J Clin Oncol*. 2013;43: 483–491. https://doi.org/10.1093/jjco/hyt040.
9. Guo Y, Shi M, Shen X, Yang C, Yang L, Zhang J. Capecitabine plus irinotecan versus 5-FU/leucovorin plus irinotecan in the treatment of colorectal cancer: a meta-analysis. *Clin Colorectal Cancer*. 2014;13:110–118. https://doi.org/10.1016/j.clcc.2013.12.004.
10. Escoriaza J, Aldaz A, Castellanos C, Calvo E, Giráldez J. Simple and rapid determination of irinotecan and its metabolite SN-38 in plasma by high-performance liquid-chromatography: application to clinical pharmacokinetic studies. *J Chromatogr B*. 2000;740:159–168.
11. Oyaga-Iriarte E, Insausti A, Sayar O, Aldaz A. Population pharmacokinetic model of irinotecan and its metabolites in patients with metastatic colorectal cancer. *Eur J Clin Pharmacol*. 2019;75:529–542. https://doi.org/10.1007/s00228-018-02609-6.
12. Le Cessie S, Van Houwelingen JC. Ridge estimators in logical regression. *Appl Stat*. 1992;41:191–201.
13. Bozdogan H. Model selection and akaike's information criterion (AIC): the general theory and its analytical extensions. *Psychometrika*. 1987;52:345–370. https://doi.org/10.1007/BF02294361.
14. Wu X, Kumar V, Ross QJ, et al. Top 10 algorithms in data mining. *Knowl Inf Syst*. 2008;14:1–37. https://doi.org/10.1007/s10115-007-0114-2.
15. Breiman L. Random forests. *Mach Learn*. 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.
16. Vapnik VN. *The nature of statistical learning theory*. New York: Springer-Verlag; 1995.
17. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *International joint conference on artificial intelligence (IJCAI)*. 1995:1137–1143.
18. Breiman L. Bagging predictors. *Mach Learn*. 1996;24:123–140. https://doi.org/10.1007/BF00058655.
19. Izmirlian G. Application of the random forest classification algorithm to a SELDI-TOF proteomics study in the setting os a cancer prevention trial. *N Y Acad Sci*. 2004;1020:154–174.
20. Grumett S, Snow P, Kerr D. Neural networks in the prediction of survival in patients with colorectal cancer. *Clin Colorectal Cancer*. 2003;2:239–244. https://doi.org/10.3816/CCC.2003.n.005.
21. Wang G, Lam KM, Deng Z, Choi KS. Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques. *Comput Biol Med*. 2015;63:124–132. https://doi.org/10.1016/j.compbiomed.2015.05.015.
22. Zhou ZH, Jiang Y, Yang Y Bin, Chen SF. Lung cancer cell identification based on artificial neural network ensembles. *Artif Intell Med*. 2002;24:25–36. https://doi.org/10.1016/S0933-3657(01)00094-X.
23. Exarchos TP, Tzallas AT, Baga D, et al. Using partial decision trees to predict Parkinson's symptoms: a new approach for diagnosis and therapy in patients suffering from Parkinson's disease. *Comput Biol Med*. 2012;42:195–204. https://doi.org/10.1016/j.compbiomed.2011.11.008.
24. Peng L, Chen W, Zhou W, Li F, Yang J, Zhang J. An immune-inspired semi-supervised algorithm for breast cancer diagnosis. *Comput Methods Progr Biomed*. 2016;134:259–265. https://doi.org/10.1016/j.cmpb.2016.07.020.
25. Gao M, Igata H, Takeuchi A, Sato K, Ikegaya Y. Machine learning-based prediction of adverse drug effects: an example of seizure-inducing compounds. *J Pharmacol Sci*. 2017;133:70–78. https://doi.org/10.1016/j.jphs.2017.01.003.
26. Ibrić S, Djuriš J, Parojčić J, Djurić Z. Artificial neural networks in evaluation and optimization of modified release solid dosage forms. *Pharmaceutics*. 2012;4: 531–550. https://doi.org/10.3390/pharmaceutics4040531.
27. Archetti F, Lanzeni S, Messina E, Vanneschi L. Genetic programming for computational pharmacokinetics in drug discovery and development. *Genet Program Evolvable Mach*. 2007;8:413–432. https://doi.org/10.1007/s10710-007-9040-z.
28. Gertrudes JC, Maltarollo VG, Silva RA, Oliveira PR, Honório KM, Da Silva ABF. Machine learning techniques and drug design. *Curr Med Chem*. 2012;19: 4289–4297. https://doi.org/10.2174/092986712802884259.
29. Tucker CS, Behoora I, Nembhard HB, Lewis M, Sterling NW, Huang X. Machine learning classification of medication adherence in patients with movement disorders using non-wearable sensors. *Comput Biol Med*. 2015;66:120–134. https://doi.org/10.1016/j.compbiomed.2015.08.012.
30. Modongo C, Pasipanodya JG, Zetola NM, Williams SM, Sirugo G, Gumboc T. Amikacin concentrations predictive of ototoxicity in multidrug-resistant tuberculosis patients. *Antimicrob Agents Chemother*. 2015;59:6337–6343. https://doi.org/10.1128/AAC.01050-15.
31. Yamazaki D, Kitaguchi T, Ishimura M. Proarrhythmia risk prediction using human induced pluripotent stem cell-derived cardiomyocytes. *J Pharmacol Sci*. 2018;136:249–256. https://doi.org/10.1016/j.jphs.2018.02.005.
32. Jaruratanasirikul S, Wongpoowarak W, Jullangkoon M, Samaeng M. Population pharmacokinetics and dosing simulations of imipenem in serious bacteraemia in immunocompromised patients with febrile neutropenia. *J Pharmacol Sci*. 2015;127:164–169. https://doi.org/10.1016/j.jphs.2014.12.003.
33. Sandström M, Lindtnan H, Nygren P, Lidbrink E, Bergh J, Karlsson MO. Model describing the relationship between pharmacokinetics and hematologic toxicity of the epirubicin-docetaxel regimen in breast cancer patients. *J Clin Oncol*. 2005;23:413–421. https://doi.org/10.1200/JCO.2005.09.161.
34. Matsukura S, Nakamura Y, Hoshiai K, et al. Effects of moxifloxacin on the proarrhythmic surrogate markers in healthy Filipino subjects: exposure-response modeling using ECG data of thorough QT/QTc study. *J Pharmacol Sci*. 2018;136:234–241. https://doi.org/10.1016/j.jphs.2018.01.009.
35. Gusella M, Crepaldi G, Barile C, et al. Pharmacokinetic and demographic markers of 5-fluorouracil toxicity in 181 patients on adjuvant therapy for colorectal cancer. *Ann Oncol*. 2006;17:1656–1660. https://doi.org/10.1093/annonc/mdl284.
36. Sandström M, Lindman H, Nygren P, Johansson M, Bergh J, Karlsson MO. Population analysis of the pharmacokinetics and the haematological toxicity of the fluorouracil-epirubicin-cyclophosphamide regimen in breast cancer patients. *Cancer Chemother Pharmacol*. 2006;58:143–156. https://doi.org/10.1007/s00280-005-0140-2.
37. Nikanjam M, Stewart CF, Takimoto CH, et al. Population pharmacokinetic analysis of oxaliplatin in adults and children identifies important covariates for dosing. *Cancer Chemother Pharmacol*. 2015;75:495–503. https://doi.org/10.1007/s00280-014-2667-6.
38. Gupta E, Lestingi TM, Mick R, Ramirez J, Vokes EE, Ratain MJ. Metabolic fate of irinotecan in humans: correlation of glucuronidation with diarrhea. *Cancer Res*. 1994;54:3723–3725.
39. Innocenti F, Undevia SD, Iyer L, et al. Genetic variants in the UDP-glucuronosyltransferase 1A1 gene predict the risk of severe neutropenia of irinotecan. *J Clin Oncol*. 2004;22:1382–1388. https://doi.org/10.1200/JCO.2004.07.173.
40. Freyer G, Rougier P, Bugat R, et al. Prognostic factors for tumour response, progression-free survival and toxicity in metastatic colorectal cancer patients given irinotecan (CPT-11) as second-line chemotherapy after 5FU failure. *Br J Canc*. 2000;83:431–437. https://doi.org/10.1054/bjoc.2000.1303.
41. Ramchandani RP, Wang Y, Booth BP, et al. The role of SN-38 exposure, UGT1A1*28 polymorphism, and baseline bilirubin level in predicting severe irinotecan toxicity. *J Clin Pharmacol*. 2007;47:78–86. https://doi.org/10.1177/0091270006295060.
42. Hu ZY, Yu Q, Zhao YS. Dose-dependent association between UGT1A1*28 polymorphism and irinotecan-induced diarrhoea: a meta-analysis. *Eur J Cancer*. 2010;46:1856–1865. https://doi.org/10.1016/j.ejca.2010.02.049.
43. Marcuello E, Altés A, Menoyo A, Del Rio E, Gómez-Pardo M, Baiget M. UGT1A1 gene variations and irinotecan treatment in patients with metastatic colorectal cancer. *Br J Canc*. 2004;91:678–682. https://doi.org/10.1038/sj.bjc.6602042.
44. de Man FM, Goey AKL, van Schaik RHN, Mathijssen RHJ, Bins S. Individualization of irinotecan treatment: a review of pharmacokinetics, pharmacodynamics, and pharmacogenetics. *Clin Pharmacokinet*. 2018:1–26. https://doi.org/10.1007/s40262-018-0644-7.