# Pharmacogenetics–Based Preliminary Algorithm to Predict the Incidence of Infection in Patients Receiving Cytotoxic Chemotherapy for Hematological Malignancies: A Discovery Cohort

Matias F. Martinez [1,2,3]*, Enzo Alveal [1], Tomas G. Soto [1,4], Eva I. Bustamante [5], Fernanda Ávila [6], Shrikant I. Bangdiwala [7,8], Ivonne Flores [5], Claudia Monterrosa [5], Ricardo Morales [5], Nelson M. Varela [1,3], Alison E. Fohner [9] and Luis A. Quiñones [1,3]*

[1]Laboratory of Chemical Carcinogenesis and Pharmacogenetics (CQF), Department of Basic and Clinical Oncology (DOBC), Faculty of Medicine, University of Chile, Santiago, Chile, [2]Departamento de Ciencias y Tecnología Farmacéuticas, Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, Santiago de Chile, Chile, [3]Latin American Network for the Implementation and Validation of Pharmacogenomic Clinical Guidelines (RELIVAF-CYTED), Madrid, Spain, [4]Departamento De Ciencias Básicas Santiago, Facultad De Ciencias, Universidad Santo Tomás, Santiago, Chile, [5]Cancer Institute Arturo López Pérez Foundation, Santiago, Chile, [6]Clinical Hospital of the University of Chile, Santiago, Chile, [7]Population Health Research Institute, McMaster University, Hamilton, ON, Canada, [8]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada, [9]Department of Epidemiology and Institute of Public Health Genetics, University of Washington, Seattle, WA, United States

**Introduction:** Infections in hematological cancer patients are common and usually life-threatening; avoiding them could decrease morbidity, mortality, and cost. Genes associated with antineoplastics' pharmacokinetics or with the immune/inflammatory response could explain variability in infection occurrence.

**Objective:** To build a pharmacogenetic-based algorithm to predict the incidence of infections in patients undergoing cytotoxic chemotherapy.

**Methods:** Prospective cohort study in adult patients receiving cytotoxic chemotherapy to treat leukemia, lymphoma, or myeloma in two hospitals in Santiago, Chile. We constructed the predictive model using logistic regression. We assessed thirteen genetic polymorphisms (including nine pharmacokinetic—related genes and four inflammatory response-related genes) and sociodemographic/clinical variables to be incorporated into the model. The model's calibration and discrimination were used to compare models; they were assessed by the Hosmer-Lemeshow goodness-of-fit test and area under the ROC curve, respectively, in association with Pseudo-$R^2$.

**Results:** We analyzed 203 chemotherapy cycles in 50 patients (47.8 ± 16.1 years; 56% women), including 13 (26%) with acute lymphoblastic and 12 (24%) with myeloblastic leukemia.

Pharmacokinetics-related polymorphisms incorporated into the model were *CYP3A4* rs2242480C>T and *OAT4* rs11231809T>A. Immune/inflammatory response-related

polymorphisms were *TLR2* rs4696480T>A and *IL-6* rs1800796C>G. Clinical/ demographic variables incorporated into the model were chemotherapy type and cycle, diagnosis, days in neutropenia, age, and sex. The Pseudo-$R^2$ was 0.56, the *p*-value of the Hosmer-Lemeshow test was 0.98, showing good goodness-of-fit, and the area under the ROC curve was 0.93, showing good diagnostic accuracy.

**Conclusions:** Genetics can help to predict infections in patients undergoing chemotherapy. This algorithm should be validated and could be used to save lives, decrease economic costs, and optimize limited health resources.

# INTRODUCTION

Infections in patients diagnosed with hematological malignancies are common and usually life-threatening (Green 2017); 30% of patients undergoing chemotherapy for the treatment of these illnesses experience infection, and 11% dying as a result (Castagnola et al., 2007; Yilmaz et al., 2008; Dutronc et al., 2009). Several variables have been described as risk or protector factors. In addition to clinical variables associated with a high risk of infection, neutropenia, defined as an absolute neutrophil count lower than 500 cells/mm$^3$ (Klastersky et al., 2016), is common in patients undergoing cytotoxic chemotherapy (Li et al., 2016). Genetics can modify the occurrence of chemotherapy-related neutropenia. A polymorphic variant can alter the metabolism or elimination of the cytotoxic agent; this could increase the plasmatic level of the antineoplastic, and therefore increase the risk of dose-related toxicity (Lyman et al., 2014; Buaboonnam et al., 2019). Among these pharmacokinetics-related genetic factors, CYP3A4 and CYP3A5 are two enzymes that participate in the metabolism of most drugs used in the treatment of hematological malignancies (Lee et al., 2013; Daly 2015), and polymorphic variants in the genes that encode those proteins are known to decrease their expression or functionality, decreasing drug metabolism (Lamba et al., 2012; Park et al., 2014; Zhu et al., 2014). Moreover, polymorphic variants in genes that codify for drug transporters could affect the elimination of antineoplastic medications (Choi et al., 2015). Some of the transporters associated with cancer medications are ATP-Binding Cassette (ABC), specifically ABCB1, ABCC2, and ABCG2 (Jedlitschky et al., 2006; Vasiliou et al., 2009), and Solute Linking Carrier (SLC), specifically *SLC22A11* that codifies for the Organic Anion Transporter 4 (OAT4) (Burckhardt 2012). Genetic variants in these genes could increase the risk of adverse events due to decreased elimination leading to toxic drug concentrations (Vormfelde et al., 2006; Lewis et al., 2013).

Furthermore, some immune/inflammatory response-related proteins can improve the aggregation and survival of neutrophils. Interleukin 6 and 1β (IL6 and IL-1β) participate in the maturation and apoptosis inhibition of white blood cells (WBC) (Rose-John et al., 2017; Chiba et al., 2018). A lower expression of the genes that encode these interleukins could enhance the risk of

neutropenia, make it more severe or prolong the length of neutropenia (Wright et al., 2014; Loft et al., 2018; Badawy et al., 2019). The caspase recruitment domain 8 (CARD8) participates in the activation of IL-1β, and a polymorphism that affects CARD8 functionality could have the same consequences of having less interleukin signal (Paramel et al., 2015). Toll-like receptor 2 (TLR2) is a protein that senses pathogen molecules and develops the intracellular signaling in response to a possible infection (Beutler et al., 2006); a genetic variant that affects the functionality of this receptor could increase the risk of infections (Hawn et al., 2009; Bielinski et al., 2011; Esposito et al., 2014).

No tool can predict the incidence of infections in hematological cancer patients in chemotherapy that use clinical and genetic variables. By accurately predicting which patients are likely to develop an infection while undergoing chemotherapy, dose adjustments and enhanced monitoring could be targeted to prevent infection, thereby reducing morbidity, mortality, and healthcare costs. Our study aims to create an algorithm to predict the incidence of infections among patients undergoing chemotherapy, using pharmacokinetics and immune response-related genetic polymorphisms in addition to clinical variables.

# MATERIALS AND METHODS

## Study Design

We carried out a prospective cohort study from November 2017 to October 2018 at the Oncologic Hospital "Fundación Arturo Lopez Pérez" (FALP) and the Clinical Hospital of the University of Chile (HCUCH) in Santiago, Chile. Patients were enrolled at those clinical centers before the first chemotherapy cycle, and they were followed prospectively through every cycle of chemotherapy. Infection was established by clinical criteria, including fever not explained by chemotherapy, positive bacterial cultures, or compatible imaging. A multidisciplinary team compound by specialist physicians and clinical pharmacists reviewed every febrile episode and decided if it was due to infection or not according to local practice. The occurrence of any infections was recorded.

## Patients and Data Source

The study included patients 18 years or older diagnosed with leukemia or non-Hodgkin's lymphoma and undergoing cytotoxic chemotherapy. We excluded patients regularly taking immunosuppressive medication, pregnant women, and patients with a diagnosis of immunodeficiency. All clinical data were obtained from the medical record. We collect the information that allows clinic, pharmacotherapeutic, morbid, and demographic characterization of the sample.

## Ethics Statement

All patients signed a written informed consent and an agreement to participate in this study. The study was carried out following the strict ethical procedures recommended by the Ethics Committee of the Clinical Hospital of the University of Chile (approval received on July 18, 2017) and the Eastern Metropolitan Health Service (approval received on July 4, 2017), following the procedures suggested in the Declaration of Helsinki, with Chilean Laws 20.120, 20.584, and 19.628 and with the guidelines of the Good Clinical Practices from the World Health Organization.

## Genotyping Analysis

Genomic DNA was isolated from the subjects' peripheral blood samples using the High Pure PCR Template Preparation Kit (Catalog Number, 11796828001; Roche Diagnostics GmbH, Mannheim, Germany). The blood sample was collected after the first remission.

The nine pharmacokinetics-related polymorphisms were *SLC22A11* rs11231809; *ABCB1* rs2032582, rs1045642y rs1128503; *CYP3A4* rs2740574, rs2242480, *CYP3A5* rs15524, *ABCC2* rs12762549, and *ABCG2* rs2231142. The immune response-related polymorphisms were *IL6* rs1800796, *IL1β* rs1143627, *CARD8* rs2043211, and *TLR2* rs4696480. The potential effect in protein was the main criteria to pick the polymorphic variants in addition to the relationship with drug toxicity reported in the literature. Another factor considered was the minor allele frequency of more than 5% in the Latinamerican population when available.

All polymorphisms were analyzed using *TaqMan*® SNP Genotyping Assay (Catalog number, 4362691; Thermo Fisher Scientific, Waltham, MA, United States) in a Stratagene Mx3000p real-time PCR system (Agilent Technologies, Santa Clara, CA, United States). Every sample was analyzed in triplicate to ensure reliability. The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: [10.6084/m9.figshare. 13444211].

## Statistical Analysis

To determine which genetic variants were associated with infection incidence, it was a binary variable, and it was assessed in every chemotherapy cycle. We carried out multivariate logistic regression models to establish the relationship between every polymorphism by itself and the outcome. We assessed the inheritance in a dominant, recessive, and co-dominant model, and they were added

**TABLE 1 |** Sample characteristics.

| Characteristics | n = 50 (%) |
| --- | --- |
| Age, years (median ± IQR) | 40 (31–52) |
| 18–40 years | 17 (34) |
| 41–60 years | 20 (40) |
| >60 years | 13 (26) |
| Female | 27 (54) |
| Male | 23 (46) |
| Diagnosis | |
| Acute lymphoblastic leukemia | 13 (26) |
| Acute myeloblastic leukemia | 12 (24) |
| Other leukemias | 5 (10) |
| Lymphoma | 20 (40) |

depending on their statistical significance. These models included one genetic variable and the following control variables: sex; age (in groups: from 18 to 40 years; from 41 to 60 years and older than 60 years); the number of chemotherapy cycle; kind of chemotherapy scheme (induction or consolidation), diagnosis (acute lymphoblastic leukemia, acute myeloblastic leukemia, other leukemias or lymphoma) and the number of days in profound neutropenia, defined as an absolute neutrophil count of 0 cells/mm$^3$.

Adjusted Odds Ratios (OR) were obtained from exponentiating the coefficient given by the regression model. The ORs were used to determine which variables were risk or protective factors of having an infection.

To develop the final model, we added all the control variables and those genetic factors that had a statically significant association with infection in the previous step. The best model was chosen according to the value of Pseudo-$R^2$, the calibration, and the discrimination of the model. The control variables were chosen based on the univariate relationship or previous reports of the association's association with the event.

We use a mixed-effect model to account for intraindividual variability regarding chemotherapy scheme and cycle and use correlated outcome data (also known as hierarchical models). Here we used a three-level model, so we had a random intercept for three characteristics. The first one was the patient level, the second chemotherapy scheme, and finally cycle level, where infections were assessed.

Calibration is the degree of similarity between the probability given by the model and the observed incidence. The goodness-of-fit test of Hosmer—Lemeshow compares frequencies of cases and controls using a chi2 test. In this case, a higher *p*-value indicates fewer differences between the predicted frequencies and the frequencies observed in the sample.

On the other hand, discrimination is the degree to which the proposed model can distinguish between patients who experience the event from those who do not; that is, the ability to indicate that a patient will experience an infection, and, on the other hand, the ability to predict when the patient will not experience it. Discrimination between cases and non-cases of infection was tested using the area under the Receiver Operating Characteristic curve (AUC-ROC).

**TABLE 2** | Allele and genotypic frequencies of the studied polymorphism.

| Genetic polymorphism | Allele frequency | | Genotypic frequency, n (%) | | |
|---|---|---|---|---|---|
| CARD8 rs2043211 | A | T | A/A | A/T | T/T |
| | 0,64 | 0,36 | 19 (39) | 25 (51) | 5 (10) |
| TLR2 rs4696480 | T | A | T/T | T/A | A/A |
| | 0,63 | 0,38 | 20 (42) | 20 (42) | 8 (17) |
| IL-6 rs1800796 | G | C | G/G | G/C | C/C |
| | 0,50 | 0,50 | 11 (22) | 28 (56) | 11 (22) |
| IL-1β rs1143627 | G | A | G/G | G/A | A/A |
| | 0,54 | 0,46 | 16 (33) | 21 (43) | 12 (24) |
| OAT4 rs11231809 | T | A | T/T | T/A | A/A |
| | 0,38 | 0,62 | 11 (24) | 13 (28) | 22 (48) |
| ABCB1 rs2032582 | C | A | C/C | C/A | A/A |
| | 0,68 | 0,32 | 21 (53) | 12 (30) | 7 (18) |
| ABCB1 rs1045642 | A | G | AA | AG | GG |
| | 0,33 | 0,67 | 3 (6) | 27 (54) | 20 (40) |
| ABCB1 rs1128503 | A | G | AA | AG | GG |
| | 0,38 | 0,62 | 7 (14) | 24 (48) | 19 (38) |
| CYP3A4 rs2740574 | T | C | T/T | T/C | C/C |
| | 0,83 | 0,17 | 31 (74) | 8 (19) | 3 (7) |
| CYP3A4 rs2242480 | C | T | CC | CT | TT |
| | 0,71 | 0,29 | 29 (59) | 12 (24) | 8 (16) |
| CYP3A5 rs15524 | A | G | AA | AG | GG |
| | 0,84 | 0,16 | 37 (74) | 10 (20) | 3 (6) |
| ABCC2 rs12762549 | G | C | G/G | G/C | C/C |
| | 0,62 | 0,38 | 19 (42) | 18 (40) | 8 (18) |
| ABCG2 rs2231142 | G | T | G/G | G/T | T/T |
| | 0,90 | 0,10 | 37 (80) | 9 (20) | (0) |

*Some samples could not be fully genotyped.*

The sample characteristics were mainly presented as proportions. The Shapiro–Wilk test was used to assess the normality of the distribution of continuous variables, and when the distribution was skewed, the variable was presented as median and interquartile range. All the analyses and figures were performed using STATA 15.0 software ®.

# RESULTS

## Sample Characteristics

Participants' median age was 40 years (IQR 31–52), and 27 (54.0%) were women. Half of the patients had acute leukemia [13 (26%) lymphoblastic and 12 (24%) myeloblastic] (**Table 1**). We detected infection in 82 (40.8%) of 203 total chemotherapy cycles, including 31(62.0%) of the 50 patients recruited. No patients died during the follow-up period. The genotype distribution of the thirteen polymorphisms assessed in the study is given in **Table 2**.

# CONSTRUCTION OF THE PREDICTIVE MODEL

## Model With Non-genetic Variables

To assess how genetics can improve the predictive model's performance, we carried out primarily a non-genetic

algorithm. For this, we incorporated chemotherapy type and cycle, diagnosis, days in neutropenia, age, and sex. The *p*-value of the Hosmer – Lemeshow test was 0.29, the AUC-ROC was 0.82, and the Pseudo-$R^2$ was 0.23.

## Model With Genetic Variables

We found that Toll-Like Receptor 2 (*TLR2*), Interleukin 6 (*IL6*), CYP3A4, Solute Linking Carrier family 22 member 11 (*SLC22A11*) or Organic Anion Transporter 4 (*OAT4*), and ATP-Binding Cassette Subfamily C member 2 (*ABCC2*) polymorphisms were associated with the incidence of infection by themselves (**Supplementary Table S1**), and they were added to the final model.

In **Figure 1**, we summarize the OR obtained from the adjusted final model. There we found that TLR2, IL6, OAT4, and CYP3A4 were significantly associated with the occurrence of infection. The Pseudo R2 of the model was 0.5327; i.e., the model explained 53% of the variability of the incidence of infections.

The multilevel model with a random intercept for patient, chemotherapy scheme, and cycle variables gave the same calibration and discrimination performance. As we had a predictive objective and allowing a more straightforward application in clinical practice, we construct the model only with fixed intercepts, i.e., a "normal" logistic regression.
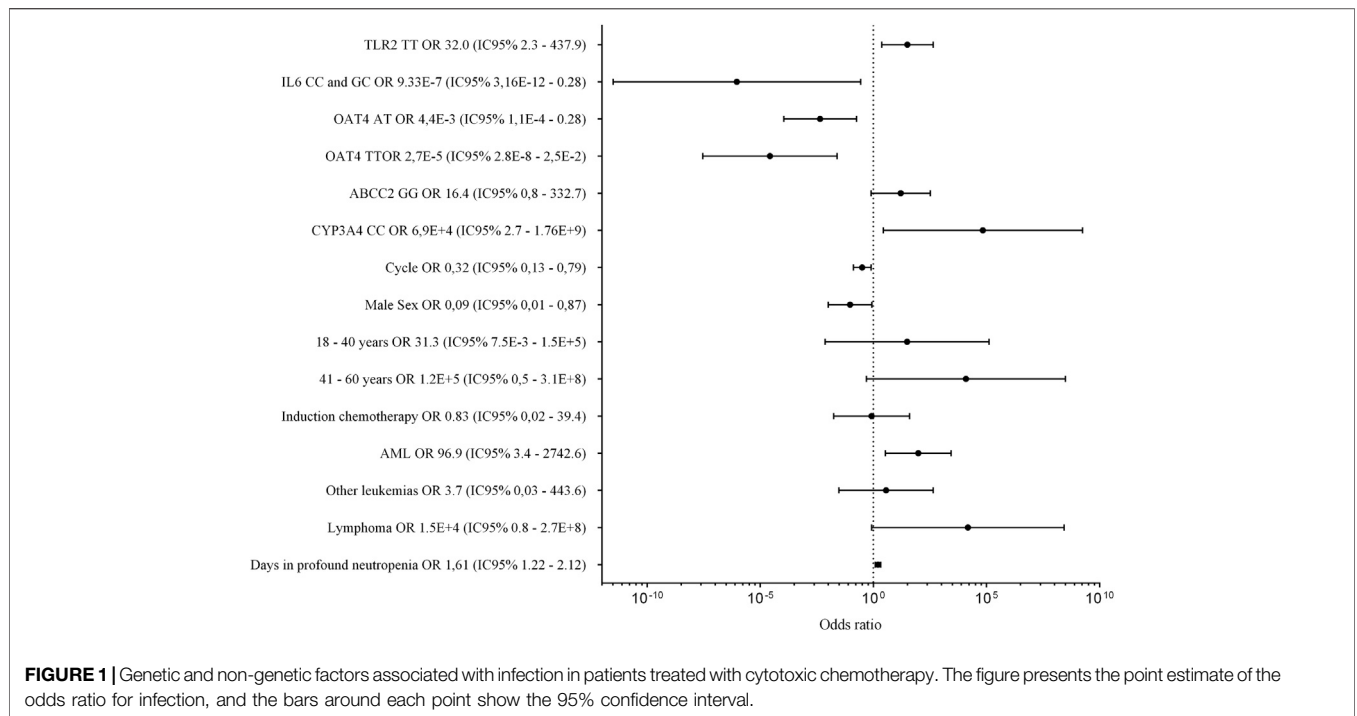
## Calibration of the Model

The *p*-value of the test of Hosmer—Lemeshow was 0.9516; this means that the model had an excellent goodness-of-fit or, in other words, there is no statistical difference between the frequency of cases detected by the model and that observed in the sample.

## Discrimination of the Model

The AUC-ROC indicates the probability of assigning a higher probability of incidence to patients who develop an infection than those who do not. The ROC curve compares sensitivity, that is, the ability to assign the event correctly when it occurred, and one- specificity, that is, the proportion of patients who did not have an infection, but the model indicated that they would present it.

The AUC-ROC curve can range from 0 to 1, with 0.5 indicating that the instrument cannot discriminate outcomes better than chance. A value of one would be for a model that perfectly predicts the occurrence of the event. For our proposed model, the AUC-ROC curve was 0.93. **Figure 2** shows the ROC curve for the model to predict the incidence of infections in the sample. Concerning non-genetic variables, we found that the number of chemotherapy cycles was a protective factor. This result could be interpreted as follows: with each cycle of chemotherapy that passes, the risk of infection decreases, perhaps because if the patient had an infection in an initial cycle, in the following cycles, the prophylaxis would be optimized to avoid futures events. Besides, the male sex was protective compared to the female sex; induction was associated with more risk than consolidation, and the longer neutropenia episodes were associated with more risk of infection.

The coefficients for every variable obtained from the final regression model allow us to build an equation to predict the

**FIGURE 1 |** Genetic and non-genetic factors associated with infection in patients treated with cytotoxic chemotherapy. The figure presents the point estimate of the odds ratio for infection, and the bars around each point show the 95% confidence interval.

probability of infection occurrence. Thus, the model, including genetic and non-genetic variables, is:

$$\ln\left(\frac{p}{1-p}\right) = 3.47 \times (TLR2\ TT) - 13.9 \times (IL6\ GG)\ 5.4 \times (OAT4\ AT) - 10.5 \times (OAT4\ TT) +$$

$$2.8 \times (ABCC2\ GG) + 11.1 \times (CYP3A4\ CC) - 1.1 \times (Cycle) - 2.4 \times (Male\ sex) + 3.4 \times$$

$$(From\ 18\ to\ 40\ years) + 9.4 \times (From\ 41\ to\ 60\ years) + 0.2 \times (Induction) + 4.6 \times (AML) +$$

$$1.3 \times (Other\ leukemias) + 9.6 \times (Lymphoma) + 0.5 \times (Days\ in\ deep\ neutropenia) - 1.9$$

where $p$ is the probability of occurrence of infection, with values ranging from 0 to 1 (or 0%–100% of occurrence probability). Some factors increased the odds of having an infection (positive factors), and others decreased the odds of the event (negative factor), according to the OR.

For genetics, the model works with dichotomous variables (i.e., 0 and 1). Thus when the variant showed in the model is present, we should use a 1, and when it is another genotype, we should use a 0. For example, in the OAT variant, if the genotype TT were determined, the coefficient for OAT AT would be multiplied by 0 (5.4 × 0) because the patient presents another genotype, and the coefficient for TT genotype should be multiplied by 1 (10.5 × 1). Alternatively, if the genotype AA were determined, the coefficients for TT and AT should be multiplied by 0. Days in neutropenia should be used as a continuous variable. Meanwhile, the days in neutropenia increase, also the risk of infection does.

## DISCUSSION

Infections are frequent and potentially lethal events in patients treated with cytotoxic chemotherapy for hematological

malignancies. Neutropenia due to the medication and the illness by itself is common in these patients, and it is one of the more critical factors in the incidence of infections. This study
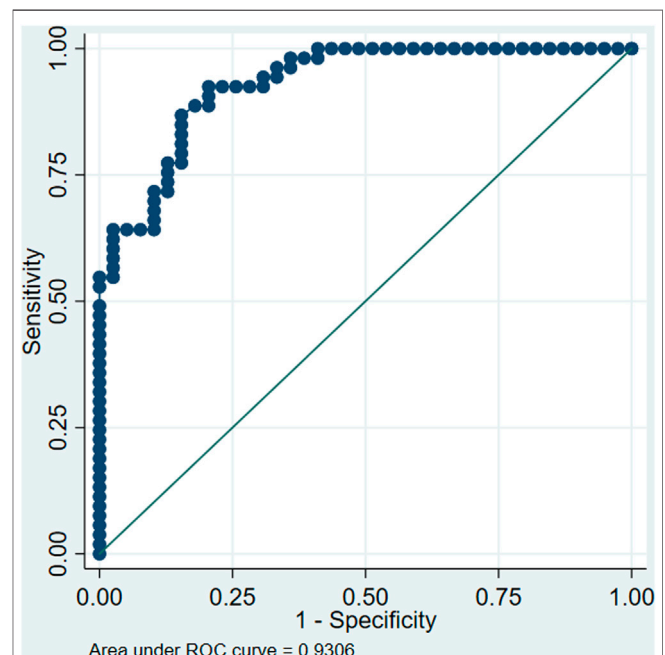


**FIGURE 2 |** Receiver Operating Characteristic (ROC) curve for the model proposed to predict infections in patients undergoing cytotoxic chemotherapy. (References TLR2 TA/AA, IL6 GG, OAT4 AA, ABCC2 CC/CG, CYP3A4 TT and CT, Female sex, 60 years or older, Consolidation chemotherapy, Acute Lymphoblastic Leukemia, respectively for each category).

is the first that aims to predict their occurrence using genes associated with both antineoplastic pharmacokinetics and immune response.

## Relationship Between Genetics and the Incidence of Infections

Concerning immune response-related genes, we found two polymorphisms significantly associated with infections. One of them was *IL6* rs1800796 (−572C>G), which is a variant located in the promoter of the gene and where the G allele is related to a lower expression of the protein (Sharma et al., 2018). We found that the GG genotype was a risk factor for having an infection, agreeing with previous reports (Tang et al., 2014; Badawy et al., 2019), probably due to a lower neutrophil mobilization and activation (Wright et al., 2014; Rose-John et al., 2017).

The other immune response-related polymorphism associated with infections was *TLR2* rs4696480 (c.-373+1614T>A). Although the variant's effect at the protein level is not well established, the T allele has been associated with lowering receptor functionality (Loft et al., 2018), a decreased function of *TLR2* prevents the pathogen recognition and delay the immune response (Hawn et al., 2009). We found the TT genotype was a risk factor of having an infection compared to TA and AA genotypes as we expected because the *TLR2* receptor also promotes neutrophil recruitment and survival (Malcolm et al., 2003; Sabroe et al., 2005).

Concerning pharmacokinetics–related polymorphisms, *SLC22A11* rs11231809 polymorphism was associated with the risk of infection. We found that the genotype TT and AT are protector factors compared to the AA genotype. This finding is consistent with previous reports describing that the A allele is associated with a lower functionality of the protein OAT4. This variant has been associated with a decreased clearance of some drugs (Vormfelde et al., 2006; Lima et al., 2015), so perhaps a lower elimination leads to a higher plasmatic level of antineoplastic that causes dose-related toxicities.

We found the polymorphic variant rs2242480 in the gene *CYP3A4* that encodes the biotransformation enzyme CYP3A4 was related to the incidence of infection. The allele CC genotype has been related to higher levels of some drugs (Zhu et al., 2014), and we found that the CC genotype was associated with an increased risk of infections, so this could be due to higher plasmatic levels of antineoplastic and, therefore, a higher probability of adverse events.

## The Usefulness of the Model

We can calculate the probability of having an infection by combining all factors and coefficients in the model we developed. Nevertheless, the clinical action taken due to the calculated value should be consistent with the clinical objective. An important issue is determining the output probability over which the patient will be considered as a probable case or not. Depending on the probability cutoff selected, the model's sensibility and specificity would change, and it is relevant to ponder which indicator will be better for clinical outcomes, workflows, and resources.
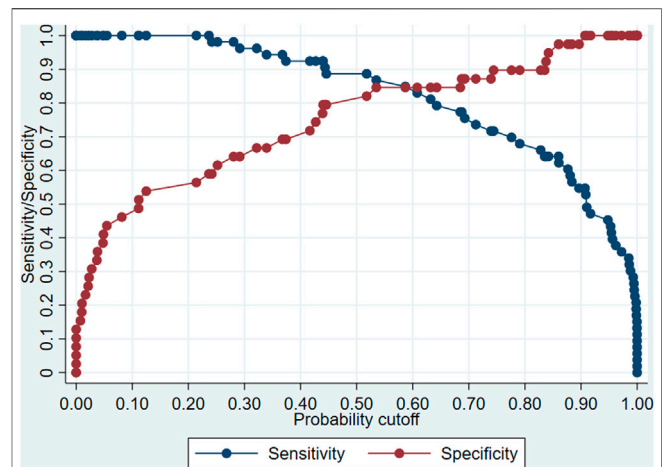


**FIGURE 3 |** Change in sensitivity and specificity of the model proposed to predict infections in patients undergoing cytotoxic chemotherapy according to the cutoff point for the definition of case.

**Figure 3** shows the change of sensitivity (blue line) and specificity (red line) of the proposed model according to different cutoffs. If a balance between sensitivity and specificity is sought, a better cutoff point would be 0.6 (that is, above 0.6, the patient would be classified as having an infection and under the value, as without infection). This cutoff is the intersection point of curves, where both specificity and sensitivity are in values around 85%.

If it is more important that all or a large part of the possible cases are correctly classified, sensitivity should be privileged, and a lower probability cutoff should be chosen. For example, with a cutoff point of 0.3, the sensitivity is 100%, which ensures that most cases (true positives) are detected, with the disadvantage that some patients who will not suffer an infection will be misclassified (false positives). In this example, the model could be used to give medical discharge to those patients who have a low risk or probability of infection.

On the other hand, if it is sought to correctly classify all patients (or the vast majority) who will not suffer an infection (true negatives), a higher cutoff point should be set. For example, by setting a cutoff point of 0.9, we ensure that no patient (or very few) who will not suffer an infection are classified as a case. The disadvantage of this approach is that the rate of those classified as non-cases would increase, even though they will present an infection (false negatives). In this example, the model could help patients at high risk of infection be referred to an isolation room or receive optimized antimicrobial prophylaxis from the start of chemotherapy since the model indicates the patient has a high probability of suffering an infection.

Concerning the performance of the proposed model, the value of the AUC-ROC is high. With an AUC-ROC value over 0.8, the model makes a proper classification of people who will have the event and who will not (Harrell, 2015). The Hosmer–Lemeshow test *p*-value indicates that the frequency of cases predicted by the model was similar to our sample's observed values.

Several studies have tried to predict the incidence of infections in patients with hematological malignancies, but no one has combined genetic and clinical factors in a single model. Webb et al. created an algorithm to predict the incidence of bloodstream infections due to vancomycin-resistant *enterococcus* in patients undergoing leukemia induction and included severe neutropenia as one of the factors adding to the use of some antimicrobials previously; they found an AUC-ROC curve of 0.84 (Webb et al., 2017).

Schalk et al. used a modified Infection Probability Score (mIPS) to predict the incidence of central venous catheter-related bloodstream infections in patients with hematological malignancies. The mIPS includes clinical variables such as the heart and respiratory rates and also the WBC count. They compared the patients' score at the moment of catheter insertion and removal; the AUC-ROC curve was 0.77 (Schalk et al., 2015). Apostolopoulou et al. included in their model the presence and length of neutropenia, and similarly to what we found with longer neutropenia associated with more risk of infection, they also identified chemotherapy as a risk factor (Apostolopoulou et al., 2010).

## Non-Genetic Factors Associated With the Risk of Infections

Other risk factors we identified as associated with the risk of infection were acute lymphoblastic leukemia diagnosis compared to acute myeloblastic leukemia and the number of days in profound neutropenia. The first factor could be due to differences in treatment and even the illness's pathophysiology (Sung et al., 2007; Chandran et al., 2012; Inaba et al., 2017). An absolute neutrophil count lower than 500 cells/mm3 is one of the more significant factors associated with infection, so the longer neutropenia lasts, the higher the risk (Gil et al., 2007). Male sex seems to be a protective factor. It could be because women are more susceptible to hematological toxicities during chemotherapy, increasing the risk of infection (Singh et al., 2005), making them more likely to have a bloodstream infection with hematology-oncology illness (Apostolopoulou et al., 2010). Additionally, women often have smaller body size than men and may be more likely to experience chemotoxicity.

It is necessary to consider non-genetic variables in the analysis because they can explain the intraindividual variability of infection, having genetics unchanged. Differences could be attributed to changes in the chemotherapy scheme or doses, different use of G-CSF or prophylaxis, and other unmeasured variables as food, visits, or bed availability. With this model, we can better understand some of the factors (genetic or not) involved in chemotherapy response.

## Limitations

This study aimed to create a preliminary algorithm and identify genetic and clinical variables associated with infection risk in patients undergoing chemotherapy. Nevertheless, these results could be used as a base for new studies. Because of the small

sample size, we could not split our cohort into a training and test cohort for the model development. As a result, the performance metrics of our prediction model likely reflect a degree of overfitting. The prediction algorithm should not be directly applied to clinical practice. Future studies are needed to validate the model and set cutoff decision parameters to improve clinical care and outcomes.

One of the proposed model's main limitations is that it uses the number of days in profound neutropenia as an explanatory variable and other data obtained directly from the clinical record. However, if we want to use this prediction model at the beginning of the treatment cycle, we should use the expected number of days, causing the model to lose accuracy because it would be an approximate number, not the actual length of neutropenia. A solution to this problem is to generate a second model to predict neutropenia duration and feed its results into the previous model to predict infection risk.

Due to the sample size, we decided not to incorporate some pharmacotherapeutic variables, mainly the specific chemotherapy scheme or the use of granulocyte colony-stimulating factors. In a larger sample, these variables should be included to improve the precision of the model. Besides, other possible not measured explanatory variables, such as the severity of the illness or the gut microbiome, may influence infection risk (Hakim et al., 2018).

The sample size should be analyzed in the context of a developing way country, with a population of 18 million and an incidence of hematological malignancies of 15 cases per 100,000 habitants/year. Also, the prospective character of the study makes the rate of recruitment similar to the incidence rate. This sample size means almost all the new cases in a year in two hospitals in Chile. Although we did not reach a power to discard other polymorphisms, we got the confidence of four polymorphism were related to the event.

## CONCLUSION

*CYP3A4* rs2242480C>T, *SLC22A11* (OAT4) rs11231809T>A, *TLR2* rs4696480T>A, and *IL6* rs1800796C>G genetic polymorphisms are associated with the incidence of infections among patients undergoing cytotoxic chemotherapy. Including these genetic variables with clinical variables leads to a useful prediction tool. This study is the first to use genetic variables in addition to clinical variables to predict the incidence of infections in patients with hematological malignancies undergoing cytotoxic chemotherapy and could lead to improved clinical outcomes for patients.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found on doi.org/10.6084/m9.figshare.13444211. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

# ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Clinical Hospital of the University of Chile (approval received on July 18, 2017) and the Eastern Metropolitan Health Service (approval received on July 4, 2017). The patients/participants provided their written informed consent to participate in this study.

# AUTHOR CONTRIBUTIONS

All authors listed have made a substantial and direct contribution to the work and approved it for publication.

# ACKNOWLEDGMENTS

# SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.602676/full#supplementary-material.

# REFERENCES

Apostolopoulou, E., Raftopoulos, V., Terzis, K., and Elefsiniotis, I. (2010). Infection Probability Score, Apache II and KARNOFSKY scoring systems as predictors of bloodstream infection onset in hematology-oncology patients. *BMC Infect. Dis.* 10 (1), 135. doi:10.1186/1471-2334-10-135

Badawy, A. A., Othman, G., Elabbasy, L., Abd Elsalam, M., Shrief, R., Fahmy, E., et al. (2019). IL-6− 572G/C and− 174G/C polymorphisms association with hepatitis C virus-induced hepatocellular carcinoma. *Br. J. Biomed. Sci.* 76 (4), 201–204. doi:10.1080/09674845.2019.1642562

Beutler, B., Jiang, Z., Georgel, P., Crozat, K., Croker, B., Rutschmann, S., et al. (2006). Genetic analysis of host resistance: Toll-like receptor signaling and immunity at large. *Annu. Rev. Immunol.* 24, 353–389. doi:10.1146/annurev.immunol.24.021605.090552

Bielinski, S. J., Hall, J. L., Pankow, J. S., Boerwinkle, E., Matijevic-Aleksic, N., He, M., et al. (2011). Genetic variants in TLR2 and TLR4 are associated with markers of monocyte activation: the atherosclerosis risk in communities MRI study. *Hum. Genet.* 129 (6), 655–662. doi:10.1007/s00439-011-0962-4

Buaboonnam, J., Sripatanatadasakul, P., Treesucon, A., Glomglao, W., Siraprapapat, P., Narkbunnam, N., et al. (2019). Effect of NUDT15 on incidence of neutropenia in children with acute lymphoblastic leukemia. *Pediatr. Int.* 61 (8), 754–758. doi:10.1111/ped.13905

Burckhardt, G. (2012). Drug transport by organic anion transporters (OATs). *Pharmacol. Ther.* 136 (1), 106–130. doi:10.1016/j.pharmthera.2012.07.010

Castagnola, E., Fontana, V., Caviglia, I., Caruso, S., Faraci, M., Fioredda, F., et al. (2007). A prospective study on the epidemiology of febrile episodes during chemotherapy-induced neutropenia in children with cancer or after hemopoietic stem cell transplantation. *Clin. Infect. Dis.* 45 (10), 1296–1304. doi:10.1086/522533

Chandran, R., Hakki, M., and Spurgeon, S. (2012). "Infections in leukemia," in *Sepsis-an ongoing and significant challenge.* Editors L. Azevedo (London, United Kingdom: IntechOpen).

Chiba, Y., Mizoguchi, I., Hasegawa, H., Ohashi, M., Orii, N., Nagai, T., et al. (2018). Regulation of myelopoiesis by proinflammatory cytokines in infectious diseases. *Cell. Mol. Life Sci.* 75 (8), 1363–1376. doi:10.1007/s00018-017-2724-5

Choi, J. R., Kim, J. O., Kang, D. R., Shin, J. Y., Zhang, X. H., Oh, J. E., et al. (2015). Genetic variations of drug transporters can influence on drug response in patients treated with docetaxel chemotherapy. *Cancer Res. Treat.* 47 (3), 509. doi:10.4143/crt.2014.012

Daly, A. K. (2015). Polymorphic variants of cytochrome P450: relevance to cancer and other diseases. *Adv. Pharmacol.* 74, 85–111. doi:10.1016/bs.apha.2015.03.001

Dutronc, H., Billhot, M., Dupon, M., Eghbali, H., Donamaria, C., Dauchy, F.-A., et al. (2009). Prise en charge de 315 épisodes neutropéniques fébriles dans un centre anticancéreux. *Med. Mal. Infect.* 39 (6), 388–393. doi:10.1016/j.medmal.2008.10.014

Esposito, S., Zampiero, A., Pugni, L., Tabano, S., Pelucchi, C., Ghirardi, B., et al. (2014). Genetic polymorphisms and sepsis in premature neonates. *PLoS One* 9 (7), e101248. doi:10.1371/journal.pone.0101248

Gil, L., Styczynski, J., and Komarnicki, M. (2007). Infectious complication in 314 patients after high-dose therapy and autologous hematopoietic stem cell transplantation: risk factors analysis and outcome. *Infection* 35 (6), 421–427. doi:10.1007/s15010-007-6350-2

Green, M. L. (2017). Viral pneumonia in patients with hematopoietic cell transplantation and hematologic malignancies. *Clin. Chest Med.* 38 (2), 295–305. doi:10.1016/j.ccm.2016.12.009

Hakim, H., Dallas, R., Wolf, J., Tang, L., Schultz-Cherry, S., Darling, V., et al. (2018). Gut microbiome composition predicts infection risk during chemotherapy in children with acute lymphoblastic leukemia. *Clin. Infect. Dis.* 67 (4), 541–548. doi:10.1093/cid/ciy153

Harrell, F. E., Jr (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis.* Berlin, Germany: Springer.

Hawn, T. R., Scholes, D., Li, S. S., Wang, H., Yang, Y., Roberts, P. L., et al. (2009). Toll-like receptor polymorphisms and susceptibility to urinary tract infections in adult women. *PLoS One* 4 (6), e5990. doi:10.1371/journal.pone.0005990

Inaba, H., Pei, D., Wolf, J., Howard, S., Hayden, R., Go, M., et al. (2017). Infection-related complications during treatment for childhood acute lymphoblastic leukemia. *Ann. Oncol.* 28 (2), 386–392. doi:10.1093/annonc/mdw557

Jedlitschky, G., Hoffmann, U., and Kroemer, H. K. (2006). Structure and function of the MRP2 (ABCC2) protein and its role in drug disposition. *Expert Opin. Drug Metab. Toxicol.* 2 (3), 351–366. doi:10.1517/17425255.2.3.351

Klastersky, J., De Naurois, J., Rolston, K., Rapoport, B., Maschmeyer, G., Aapro, M., et al. (2016). Management of febrile neutropaenia: ESMO clinical practice guidelines. *Ann. Oncol.* 27 (Suppl 5), v111–v118. doi:10.1093/annonc/mdw325

Lamba, J., Hebert, J. M., Schuetz, E. G., Klein, T. E., and Altman, R. B. (2012). PharmGKB summary: very important pharmacogene information for CYP3A5. *Pharmacogenet. Genomics* 22 (7), 555. doi:10.1097/FPC.0b013e328351d47f

Lee, J. S., Cheong, H. S., Kim, L. H., Kim, J. O., Seo, D. W., Kim, Y. H., et al. (2013). Screening of genetic polymorphisms of CYP3A4 and CYP3A5 genes. *Korean J. Physiol. Pharmacol.* 17 (6), 479–484. doi:10.4196/kjpp.2013.17.6.479

Lewis, L. D., Miller, A. A., Owzar, K., Bies, R. R., Markova, S., Jiang, C., et al. (2013). The relationship of polymorphisms in ABCC2 and SLCO1B3 with docetaxel pharmacokinetics and neutropenia: CALGB 60805 (Alliance). *Pharmacogenet. Genomics* 23 (1), 29–33. doi:10.1097/FPC.0b013e32835b16d8

Li, Y., Klippel, Z., Shih, X., Reiner, M., Wang, H., and Page, J. H. (2016). Relationship between severity and duration of chemotherapy-induced neutropenia and risk of infection among patients with nonmyeloid malignancies. *Support. Care Cancer* 24 (10), 4377–4383. doi:10.1007/s00520-016-3277-0

Lima, A., Bernardes, M., Azevedo, R., Medeiros, R., and Seabra, V. (2015). Pharmacogenomics of methotrexate membrane transport pathway: can clinical response to methotrexate in rheumatoid arthritis be predicted? *Int. J. Mol. Sci.* 16 (6), 13760–13780. doi:10.3390/ijms160613760

Loft, N., Skov, L., Iversen, L., Gniadecki, R., Dam, T., Brandslund, I., et al. (2018). Associations between functional polymorphisms and response to biological treatment in Danish patients with psoriasis. *Pharmacogenomics J.* 18 (3), 494. doi:10.1038/tpj.2017.31

Lyman, G. H., Abella, E., and Pettengell, R. (2014). Risk factors for febrile neutropenia among patients with cancer receiving chemotherapy: a

systematic review. *Crit. Rev. Oncol. Hematol.* 90 (3), 190–199. doi:10.1016/j.critrevonc.2013.12.006

Malcolm, K. C., Arndt, P. G., Manos, E. J., Jones, D. A., and Worthen, G. S. (2003). Microarray analysis of lipopolysaccharide-treated human neutrophils. *Am. J. Physiol. Lung Cel Mol Physiol* 284 (4), L663–L670. doi:10.1152/ajplung.00094.2002

Paramel, G., Sirsjö, A., and Fransén, K. (2015). Role of genetic alterations in the NLRP3 and CARD8 genes in health and disease. *Mediators Inflamm.* 2015, 846782. doi:10.1155/2015/846782

Park, J.-Y., Cha, Y.-J., and Kim, K.-A. (2014). CYP3A5*3Polymorphism and its clinical implications and pharmacokinetic role. *Transl Clin. Pharmacol.* 22 (1), 3–7. doi:10.12793/tcp.2014.22.1.3

Rose-John, S., Winthrop, K., and Calabrese, L. (2017). The role of IL-6 in host defence against infections: immunobiology and clinical implications. *Nat. Rev. Rheumatol.* 13 (7), 399. doi:10.1038/nrrheum.2017.83

Sabroe, I., Dower, S. K., and Whyte, M. K. (2005). The role of Toll-like receptors in the regulation of neutrophil migration, activation, and apoptosis. *Clin. Infect. Dis.* 41 Suppl 7 (Suppl 7), S421–S426. doi:10.1086/431992

Schalk, E., Hanus, L., Färber, J., Fischer, T., and Heidel, F. H. (2015). Prediction of central venous catheter-related bloodstream infections (CRBSIs) in patients with haematologic malignancies using a modified Infection Probability Score (mIPS). *Ann. Hematol.* 94 (9), 1451–1456. doi:10.1007/s00277-015-2387-y

Sharma, A., Singh, K., Biswas, A., Ranjan, R., Kishor, K., Pandey, H., et al. (2018). Impact of interleukin 6 promoter polymorphisms (-174 G > C, -572 G > C and -597 G > A) on plasma IL-6 levels and their influence on the development of DVT: a study from India. *Hematology* 23 (10), 833–838. doi:10.1080/10245332.2018.1483546

Singh, S., Parulekar, W., Murray, N., Feld, R., Evans, W. K., Tu, D., et al. (2005). Influence of sex on toxicity and treatment outcome in small-cell lung cancer. *J. Clin. Oncol.* 23 (4), 850–856. doi:10.1200/JCO.2005.03.171

Sung, L., Lange, B. J., Gerbing, R. B., Alonzo, T. A., and Feusner, J. (2007). Microbiologically documented infections and infection-related mortality in children with acute myeloid leukemia. *Blood* 110 (10), 3532–3539. doi:10.1182/blood-2007-05-091942

Tang, S., Yuan, Y., He, Y., Pan, D., Zhang, Y., Liu, Y., et al. (2014). Genetic polymorphism of interleukin-6 influences susceptibility to HBV-related hepatocellular carcinoma in a male Chinese Han population. *Hum. Immunol.* 75 (4), 297–301. doi:10.1016/j.humimm.2014.02.006

Vasiliou, V., Vasiliou, K., and Nebert, D. W. (2009). Human ATP-binding cassette (ABC) transporter family. *Hum. Genomics* 3 (3), 281. doi:10.1186/1479-7364-3-3-281

Vormfelde, S. V., Schirmer, M., Hagos, Y., Toliat, M. R., Engelhardt, S., Meineke, I., et al. (2006). Torsemide renal clearance and genetic variation in luminal and basolateral organic anion transporters. *Br. J. Clin. Pharmacol.* 62 (3), 323–335. doi:10.1111/j.1365-2125.2006.02655.x

Webb, B. J., Healy, R., Majers, J., Burr, Z., Gazdik, M., Lopansri, B., et al. (2017). Prediction of bloodstream infection due to vancomycin-resistant enterococcus in patients undergoing leukemia induction or hematopoietic stem-cell transplantation. *Clin. Infect. Dis.* 64 (12), 1753–1759. doi:10.1093/cid/cix232

Wright, H. L., Cross, A. L., Edwards, S. W., and Moots, R. J. (2014). Effects of IL-6 and IL-6 blockade on neutrophil function *in vitro* and *in vivo*. *Rheumatology (Oxford)* 53 (7), 1321–1331. doi:10.1093/rheumatology/keu035

Yilmaz, Ş., Ören, H., Demircioğlu, F., and Irken, G. (2008). Assessment of febrile neutropenia episodes in children with acute leukemia treated with BFM protocols. *Pediatr. Hematol. Oncol.* 25 (3), 195–204. doi:10.1080/08880010801938231

Zhu, X., Yun, W., Sun, X., Qiu, F., Zhao, L., and Guo, Y. (2014). Effects of major transporter and metabolizing enzyme gene polymorphisms on carbamazepine metabolism in Chinese patients with epilepsy. *Pharmacogenomics* 15 (15), 1867–1879. doi:10.2217/pgs.14.142