



HAL
open science

Mycophenolic Acid Exposure Prediction Using Machine Learning

Jean-baptiste Woillard, Marc Labriffe, Jean Debord, Pierre Marquet

► **To cite this version:**

Jean-baptiste Woillard, Marc Labriffe, Jean Debord, Pierre Marquet. Mycophenolic Acid Exposure Prediction Using Machine Learning. *Clinical Pharmacology and Therapeutics*, 2021, 110 (2), pp.370-379. 10.1002/cpt.2216 . hal-04414364

HAL Id: hal-04414364

<https://unilim.hal.science/hal-04414364>

Submitted on 26 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mycophenolic acid exposure prediction using machine learning

Jean-Baptiste Woillard^{1,2}, Marc Labriffe^{1,2}, Jean Debord^{1,2}, Pierre Marquet^{1,2}

¹Pharmacology & Transplantation, UMR1248, INSERM, Université de Limoges, Limoges, France

²Department of Pharmacology, Toxicology and Pharmacovigilance, University Hospital of Limoges, Limoges, France

ORCID: 0000-0003-1695-0695

Corresponding author:

Jean-Baptiste Woillard

Title: PharmD, PhD

Department of Pharmacology, Toxicology and Pharmacovigilance, CBRS, University Hospital of Limoges, 2 Avenue Martin Luther King, 87042 Limoges, France

Phone: +33 5 55 05 61 40

Fax: +33 5 55 05 61 62

Email: jean-baptiste.woillard@unilim.fr

Conflict of interest: None of the authors have any conflicting interests to declare for this work.

Keywords: machine learning; mycophenolic acid; Xgboost; therapeutic drug monitoring

Abstract

Therapeutic drug monitoring of mycophenolic acid (MPA) based on area under the curve (AUC) is well established and machine learning (ML) approaches could help to estimate AUC. The aim of this work is to estimate the AUC of MPA in organ transplant patients using Xgboost ML models.

A total of 12,877 MPA AUC_{0-12h} requests from 6884 patients sent to our ISBA expert system (www.pharmaco.chu-limoges.fr/) for AUC estimation and dose recommendation based on MPA concentrations measured at least at 3 sampling times (approx. 20min, 1 and 3h after dosing) were used to develop 2 ML models based on 2 or 3 concentrations. Data were split into a training set (75%) and a test set (25%) and the Xgboost models in the training set with the lowest RMSE in a ten-fold cross-validation experiment were evaluated in the test set and in 4 independent full-pk datasets from renal or heart transplant patients.

ML models based on 2 or 3 concentrations, differences between these concentrations, relative deviations from theoretical times of sampling, presence of a delayed absorption peak and 5 covariates (dose, type of transplantation, associated immunosuppressant, age and time between transplantation and sampling) yielded accurate AUC estimation performances in the test datasets (relative bias<5% and relative RMSE<20%) and better performance than MAP Bayesian estimation in the 4 independent full-pk datasets.

The Xgboost ML models described allow accurate estimation of MPA AUC_{0-12h} and can be used for routine exposure estimation and dose adjustment and will soon be implemented in a dedicated web interface.

Introduction

Mycophenolate mofetil (MMF) is the prodrug of mycophenolic acid (MPA), an immunosuppressive drug inhibiting inosine monophosphate dehydrogenase which is an enzyme involved in the de novo pathway for purine biosynthesis¹. Although, MMF was initially marketed as a fixed dose drug, it exhibits a large inter-individual pharmacokinetic variability, poor relationships between dose vs. mycophenolic acid (MPA) area-under-the-curve (AUC) or through blood level (C₀) vs. AUC; and a better correlation between MPA AUC and patient outcome than any single concentration-time point or dose . While the last conference consensus recommended MPA therapeutic drug monitoring (TDM) based on the AUC as the best exposure marker ², TDM is still not performed everywhere or for all patients. This is probably due to the poor predictive value of the C₀ and the difficulty to estimate MPA AUC without the use of population pharmacokinetic models and maximum a posteriori Bayesian estimators based on limited sampling strategies. Indeed, this approach is not easy to implement in routine practice and needs some pharmacokinetic background and computer skills.

This led us to launch in 2005 the Immunosuppressant Bayesian Dose Adjustment (ISBA) expert system and website ³ to share tools able to estimate the inter-dose AUC of immunosuppressants using population pharmacokinetics (POPPK) models and maximum a posteriori Bayesian estimation (MAP-BE) on the basis of 3 blood samples and patient characteristics (type of graft, age, post transplantation period, drug measurement assay) and the expertise required to interpret the data and make dose recommendations. On this website, each request posted is validated in less than 48h by a trained pharmacologist, representing a huge workload due to the large number of requests received from transplant centers worldwide (>120,000 since 2005). The POPPK models used in the ISBA back-office for MMF have all been developed using an in-house PK modelling package (ITSIM) that employs the iterative two-stage Bayesian estimation method (ITSB), as previously reported⁴. To describe the complex MPA profiles, we used one-compartment models with linear elimination and absorption modeled using two gamma distributions⁵. No covariate is used in the models but several

priors (“models”) have been built based on a few characteristics: cyclosporine co-treatment, adult/children, type of transplantation, drug assay and post-transplantation period. These models have been largely and successfully used, as reported in several papers⁶⁻¹¹.

Machine learning is widely used in numerous application including pharmacology (Pubmed Entry “Machine learning” & “Pharmacology”= 600 in 2018, 846 in 2019) especially in structure/activity predictions^{12,13} or drug discovery¹⁴ but only a few applications to predict drug exposure, PK parameters or optimal dose exist¹⁵⁻¹⁹. Recently, we successfully applied a machine learning approach for tacrolimus AUC estimation that yielded better performance in terms of relative bias or imprecision vs reference trapezoidal rule AUC than maximum a posteriori Bayesian estimation (MAP-BE), even with only 2 samples¹⁹. We used extreme gradient boosting (Xgboost R package) where simple regression trees are iteratively built by finding among all the input variables, split values that minimize the prediction error. The iterative process constructs an additional regression tree of the same structure that minimizes the residual errors of the previous regression tree²⁰.

A major difference between the population pharmacokinetics approach and machine learning is that POPPK involves results from specifying a mechanistic model, physiologically interpretable to describe the observed data while ML aims to find the best algorithmic model that minimizes prediction errors based on the data and results. ML exhibits excellent performances while losing in interpretability (“black box” phenomenon). The more the algorithm used to predict a result is complex, the more the performances are improved but the interpretability decreases²¹. Currently, the algorithms associated with the highest number of wins in the Kaggle competition are deep neural networks and Xgboost methods²².

The objective of this study was to develop Xgboost models to estimate mycophenolic acid (MPA) inter-dose AUC_{0-12h}, based on a limited number of blood concentrations (2 or 3) and predictors, and to compare their performance to that of MAP-BE in external validation datasets.

Methods

Patients and data

The MPA AUC₀₋₁₂ estimation and dose recommendation requests received on our ISBA website since 2007, whatever the type of transplantation, were extracted and cleaned using the tidyverse framework²³. We selected the requests with: an interdose of 12h; MPA plasma levels measured using HPLC; and at least 3 sampling time points at approximately 20 min (10 to 30 min, C20), 60 min (45 to 75 min, C1) and 180 min (160 to 200 min, C3) after drug intake. The 20, 60, 180 min limited sampling strategy (LSS) was previously reported as the optimal sampling schedule in most situations for MMF AUC₀₋₁₂ estimation based on samples taken in the 4h post dosing². The other predictors available were the morning dose of MMF, the time elapsed between transplantation and sampling, the indication of MMF (in decreasing order: kidney, heart transplantation, nephrotic syndrome, liver, lung transplantation, lupus in adult patients, lupus in pediatric patients, hematopoietic stem cell transplantation and all the remaining cases gathered as “others”), the associated immunosuppressant (cyclosporine, tacrolimus, sirolimus/everolimus or none of these) and patients' age.

Plan of the study

In this study, supervised learning was used to predict AUC_{0-12h}, whose reference value had been obtained by our ISBA expert system using MAP-BE and at least 3 concentrations. Two machine learning (ML) models were developed in parallel, one including only 2 concentrations and the other including 3. The data were split into a training set, used to build the model, tune the hyperparameters and evaluate model performance by cross-validation and a test set used to evaluate the performance of the models in an independent set of data. The performances were evaluated in the test set only after the best model had been optimized in the training set, by calculating the root mean square error (RMSE, expressed in mg*h/L) between the estimated and reference AUCs. Because the reference (MAP BE) AUC_{0-12h} can be considered as an imprecise and

potentially biased estimate of the “true” AUC, the results of our Xgboost models based on 2 or 3 concentrations to the “true” AUCs obtained from full PK profiles using MAP BE applied to all the samples available were also compared. For that, we used 4 different datasets from adult renal transplant patients (i) on MMF + tacrolimus; PCCP²⁴, (ii) on MMF + cyclosporine (Stablocine²⁵) and (iii) on MMF + cyclosporine or sirolimus (CONCEPT²⁶) and one from adult heart transplant patients on MPA + cyclosporine or tacrolimus (PIGREC¹⁰). The performances of the Xgboost models in these confirmation datasets were also compared to that of the MAP-BE based on the 3 time-points limited sampling strategies available on ISBA.

Feature engineering

The MPA concentrations were binned into 3 theoretical time classes (concentrations at 20min (C20 sampled between 10 and 30min), at 1h (C1 sampled between 45 and 75 min) and at 3h (C3 sampled between 160 and 200 min)), leading to 3 predictors per patient. The relative deviation with respect to the theoretical times was also considered. For instance, if the actual sampling time was 1.06h for the theoretical time 1h, the relative time difference was $(1.06 - 1)/1 = 0.06$. The differences between the concentrations C1 – C20, C1 – C3 and C3 – C20 were also considered as potential predictors and, to add information about potentially delayed absorption peaks, 3 dummy variables corresponding to C3>C1, C3>C20 and C3>C20&C1 were created. Finally, the set of features used to predict AUC0-12h were: the categorized indication of MMF, the categorized associated immunosuppressant, patient age, time elapsed between transplantation and MPA plasma sampling, MMF morning dose, MPA concentrations at times 20min, 1h and 3h, relative deviation from the theoretical times, differences between concentrations, and dummy variables for delayed absorption peak. For the models based on 2 concentrations (C1 and C3), the relative time difference for C20 and the concentration differences or dummy delayed absorption including C20 were removed.

Exploratory data analyses

A correlation matrix and scatterplots were drawn to explore the correlation between AUC0-12h and predictors using the GGally package²⁷. Boxplots were also built to investigate AUC0-12h/dose in the subgroup with cyclosporine vs other immunosuppressants, split by MMF indications. Indeed, cyclosporine is known to be associated with lower AUC/dose values due to inhibition of MPA enterohepatic recycling²⁸.

Pre-processing of the data

For all the ML analyses, the tidymodels framework was used²⁹. The two categorical variables “indication for MMF” and “associated immunosuppressants” were one hot encoded (i.e., transformed in 0/1 dummy variables). No other pre-processing was applied to the data, as Xgboost does not require normalization. There was no missing data in the predictors. Data splitting was performed by random selection of patients in a training (75%) and a test set (25%).

Development of the Xgboost models

A tuning step was done by searching, in a 40X40 grid, the parameter combination associated with the lowest RMSE and highest r^2 between estimated and reference AUC0-12 values, using a 10-fold cross-validation (for which the training dataset was randomly split into 10 parts). In brief, the best combination of parameters was investigated in 90% of the training set (called the analysis subset) and evaluated in the remaining 10% (called the assessment subset) and this process was repeated 10 times by circular permutation. The parameters tuned were: the number of predictors randomly sampled at each split (mtry, between 1 and 28), the minimum number of data points in a node required for the node to be split further (min_n between 1 and 40), the maximum depth of the tree (tree_depth, between 1 and 15), the rate at which the boosting algorithm adapts from iteration-to-iteration (learn_rate, between 0 and 0.08) and the amount of data randomly exposed to the fitting routine at each tree (sample_size, between 0.1 and 1). The relative importance of the predictors was then evaluated by random permutations and variable importance plots were drawn. Once the best

parameter combinations had been selected, the model was evaluated using 10-fold cross-validations to assess the mean RMSE and r^2 and their standard deviations in the training set and scatter plots of estimated and residuals vs. reference AUC0-12 were drawn. In a last step, the AUC0-12 prediction was performed in the test set and the performances were evaluated by calculation of the RMSE, r^2 , the relative mean prediction error (MPE), relative RMSE (imprecision) and the number and proportion of estimations with an MPE value out of the $\pm 20\%$ interval. Scatter plots of predicted vs. reference AUC0-12, and of residuals vs. reference AUC0-12 were also drawn.

External evaluation vs full PK profiles and comparison with POPPK

Concentrations at 20min, 1h and 3h as well as dose, sampling times and time elapsed between transplantation and MPA plasma sampling were extracted from the PK databases to predict the AUC0-12 using the ML models and the MAP-BE based on 3 concentrations, as used in ISBA. The full concentration profiles were also used to calculate the “true” AUC0-12 using MAP-BE with all the available samples.

Performances of the ML models using 2 or 3 concentrations and of MAP-BE based on 3 concentrations were compared to the reference AUC0-12 in terms of the relative MPE and relative RMSE and the proportion of MPE out of the $\pm 10\%$ or the $\pm 20\%$ interval. Additionally, scatter plots of predicted AUCs or residuals vs. the reference AUC0-12 were drawn to compare the different approaches.

The PCCP study²⁴ included 128 PK profiles of 11 samples (0, 0.33, 0.66, 1, 1.5, 2, 3, 4, 6, 8, 12h post dosing) collected at 7 and 14 days, 1, 3 and 6 months after renal transplantation. The Stablocine study²⁵ included 20 PK profiles of 10 or 11 samples (0, 0.33, 0.66, 1, 1.5, 2, 3, 4, 6, 9h and (not present in every profile) 12h post dosing) in stable renal transplant patients. The CONCEPT study²⁶ included 67 PK profiles of 10 or 11 samples (0, 0.33, 0.66, 1, 1.5, 2, 3, 4, 6, 9h and (not present in every profile) 12h post dosing) collected at weeks 12, 16 and 26 after renal transplantation. The

PIGREC study¹⁰ included 75 PK profiles of 10 samples (0, 0.33, 0.66, 1, 1.5, 2, 3, 4, 6 and 9h post dosing) collected between day 7 and day 14, and at 1,3 and 12 months after transplantation.

Results

Patients and data

A total of 12,877 MPA AUC_{0-12h} requests from 6884 patients fulfilling the inclusion criteria were available in the cleaned datasets extracted from ISBA (out of 12,890 requests before cleaning). The characteristics of the training and test sets are reported in Table 1. The AUC₀₋₁₂ ranged between 1 and 181 mg*h/L and 3 and 164 mg*h/L in the training and test sets, respectively. Boxplots of AUC₀₋₁₂/dose in the different subgroups defined by cyclosporine vs other immunosuppressants and indication of MMF are presented in Figure S1.

Exploratory data analyses

The correlation matrix between AUC₀₋₁₂ and individual predictors is presented in Figure S2 and shows rather poor correlation in general, the best being with C1h ($r=0.7$) and C3h ($r=0.66$).

Xgboost model

The best-tuned parameter values for each model are presented in Supplemental Table S1. Excellent validation results were obtained in the training set (using 10-fold cross-validation) and in the test set (Table 2). Interestingly, there was no difference between the two validation steps, suggesting that there was no overfitting in the training set, thanks to the cross-validation. AUC₀₋₁₂ ML estimation was approximately twice better when using 3 concentrations rather than 2. The relative MPE was close to 0 and the relative RMSE <20% in both the training and test sets. The scatter plots and residual plots obtained in the test sets are presented in Figure 1. The variable importance plot of each model is presented in Figure S3 and shows that concentrations at time 3 and 1h were the most important variables.

External evaluation vs the reference AUC0-12 and comparison with POPPK estimates

The results of external validation in full-PK databases are presented in Table 3 and show that only the Xgboost model with 3 concentrations led to better results than with 2 points or the 3-point MAP-BE, in comparison to the reference AUC. The prediction performance in the early period post-transplantation (<1 month) was similar to that in the whole validation datasets.

Figure 2 presents the scatter plots and residual plots of estimated vs. reference AUC0-12 in these final validation studies. Figure 3 presents for each individual the MPA AUC0-12 estimated using ML with 2 or 3 samples, MAP-BE with 3 samples and the reference AUC0-12.

Discussion

In this work, we developed Xgboost ML models based on 2 or 3 samples to estimate the AUC0-12h of MPA in transplant or auto-immune indications. The models were developed using the standard ML procedure with a development in a training set using cross-validation for tuning of the parameters and estimation of the performances and once the best model was selected, validation in the test set. The results were then compared on an external dataset with those obtained by extensive-sampling AUC0-12 estimated using MAP-BE using all the available samples (reference AUC0-12) and to MAP-BE estimates based on a 3-point LSS, as employed by the ISBA expert system³ in renal and heart transplant patients. The performances of the ML were excellent as compared to the reference AUC0-12 when using 3 concentrations, and numerically better than those of MAP-BE based on the 3-point LSS.

The performances obtained with the 3-point ML model were better in all the external validation datasets, in terms of imprecision and number of profiles out of the $\pm 20\%$ acceptability interval, than the 3-point MAP-BE. Additionally, the MAP-BEs available in ISBA and used in the present study for comparison to ML were initially developed using the datasets used here for validation, probably

advantaging them. However, in line with our previous results for tacrolimus, ML yields better results. The ML model using only 2 samples showed acceptable bias values but imprecision around 25%, which is not so high in comparison to literature results but very close to the threshold of 25% considered as the limit for accurate estimation in the last MPA consensus². This 2-point ML estimator might still be used in cases where only 2 samples are available, but the prescriber has to be informed of the degraded performance in comparison to the 3-point model.

In addition to better performances than the 3 samples MAP-BE used in ISBA, the ML approach has the advantage of practicability, as our single ML model can accommodate all associated immunosuppressants or MMF indications, when different MAP-BEs are necessary for each subgroup in ISBA.

There are some differences between this study and our previous one with tacrolimus¹⁹. Indeed, tacrolimus POPPK is not easy to model, but it has obvious associations with covariates that explain part of the inter-individual variability, the dose-exposure relationship is good and the AUC is fairly correlated with single time-points (the first being the trough level, used in routine practice)³⁰. All these elements may explain why our first ML algorithms were successful. The situation is very different with MMF, which can be regarded as the worst-case scenario: MPA exhibits very complex pharmacokinetics, the very diverse (and sometimes chaotic) individual profiles are much more difficult to model and there is only a weak correlation between the dose and the AUC (as shown here), and between the AUC and any single time point². Therefore, it was much more challenging to develop accurate ML algorithms for MPA, and this achievement suggests it should be possible for all the other drugs for which TDM is performed. The progress when compared to MAP-BE looks less impressive than with tacrolimus, but this highlights the fact that learning from a less precise reference (here MAP-BE imprecision is ~20% for MMF) results in somewhat less precise prediction. If the ML models were trained in a large, full-profile dataset they would have been more accurate, but we did not have such a database. In our hands, trading rich profiles for a lower number of patients

(i.e. training the ML algorithms in the validation datasets) did not result in a better performance either.

The principle of ML is to build algorithms that estimate individual values with the lowest imprecision, which is perfectly adapted to TDM. However, ML algorithms cannot study the determinants of inter- or intra-individual variability or perform simulations to estimate a priori the first dose, or calculate the probability of target attainment for different dosing regimens, for instance. Therefore, we think that ML may replace POPPK for individual exposure estimation but that POPPK will remain very useful for other purposes. Additionally, although being very accurate, ML algorithms are not as flexible as POPPK models. For instance, our ML models cannot take into account more than 3 samples per patient and the sampling times have to be in the same ranges as those of the training set (since the validation for times outside these ranges has not yet been done).

Imprecision values observed in the present study (19%, 14.7%, 22.5%, in renal and 14.6 % in heart transplantation) are slightly better than those reported in the literature, in the range of 20%-25%^{5,10,31-35}. These values are higher in comparison to other immunosuppressants (e.g. tacrolimus, with imprecision ~10%) but they reflect the complexity of MPA pharmacokinetics. Some papers reported slightly better precision, but these POPPK models were either not validated in independent datasets^{7,9,36} or they compared 3-points MAP-BE AUCs to reference AUCs obtained using 6 samples and the same POPPK model, leading probably to overestimation of the MAP-BE results¹¹.

The randomized, comparative APOMYGRE study in 137 adult kidney transplant recipients on CsA showed that concentration-controlled MMF dosing using the MAP-BE implemented in ISBA had a significant lower incidence of acute rejection than patients with standard-dose MMF³⁷. In their meta-analysis, Metz et al.³⁸ concluded that when concentration-controlled dosing is performed based on pharmacokinetic calculation to a target concentration, MPA exposure is well controlled and clinical outcomes are improved, which has been stressed in the very recent consensus conference paper².

Based on the better performance of our ML estimator in comparison to ISBA MAP-BEs (among which those used in the APOMYGRE study), we propose to use it for routine calculation of MPA AUC and will make it available to the medical community through a web-interface. In the meantime, a temporary shiny application (https://jbwoillard.shinyapps.io/App-7_mmf_ml/) has been built to test the model (for research purposes only).

The added value of AUC-based dose adjustment of MMF has currently been demonstrated for the very early post-transplant phase (especially the first weeks after transplantation)². The performances of our ML models in patients transplanted for less than 1 month (only one study in kidney and one in heart transplant patients had data collected within the 1st month) were very similar to those in the whole validation population.

The use of our ML estimator is submitted to some restrictions, due to the datasets used to design it: it was developed using MPA concentrations measured using HPLC and in which the inter-dose was 12h. Additionally, the external dataset used for external validation comprised only renal and heart transplant recipients. It's performances in the other MMF indications for which the model was developed (nephrotic syndrome, liver or lung transplantation, lupus, pediatric lupus, hematopoietic stem cell transplantation and "other" mostly corresponding to auto-immune disease) are still to be evaluated.

We believe that the ML models we developed are insensitive to the achievement of the pharmacokinetic steady-state, because they do not rely on a pharmacokinetic model and they gave excellent results in the very large training and validation datasets, where part of the profiles were certainly not at steady-state (e.g., profiles collected on day 7 post-transplantation where the dose had already been adjusted based on adverse effects).

However, these ML models have some limitations. To cope with deviations from the theoretical times, we included them as new features in the model but we have no idea about the accuracy of the algorithms when samples are drawn outside the ranges actually observed in our datasets. This has to

be investigated in further studies. In the meantime, we do not recommend using the algorithms when samples are collected outside the pre-defined ranges (20 min: from 10 to 30 min; 60 min: from 45 to 75 min; and 180 min: from 160 to 200 min). Also, our ML models are less flexible than MAP-BE and cannot manage more than 3 samples. These ML models are perfectly adapted for routine use but not really for research projects or specific cases in which more than 3 samples are available.

In conclusion, an Xgboost ML model providing accurate estimation of mycophenolic acid AUC0-12h has been developed for many MMF indications, and validated in independent groups of kidney and heart transplant recipients. It will soon be implemented in a dedicated web interface for these two indications. A patent has been filed.

Study Highlights

- What is the current knowledge on the topic? Mycophenolic acid (MPA) AUC0-12h for individual dose adjustment is difficult to estimate in routine patient care.
- What question did this study address? We investigated whether machine learning models could estimate MPA AUC0-12h using a limited number of blood concentrations, as well as or even better than deterministic pharmacokinetic models with Bayesian estimation.
- What does this study add to our knowledge? We developed and validated in kidney or heart transplants Xgboost machine learning models allowing accurate estimation of MPA AUC0-12h based on 3 blood concentrations with better performances than that of the pharmacokinetic approach previously employed.
- How might this change clinical pharmacology or translational science? These models will be soon implemented in an expert system made available to the transplant community through a dedicated website.

Acknowledgment

We thank the patients whose data were used in this study and the physicians and clinical pharmacologists for their use of the ISBA expert system.

Author Contributions

Jean-Baptiste WOILLARD designed the research, Jean-Baptiste Woillard, Marc Labriffe, Jean Debord and Pierre Marquet performed the research and analyzed the data, Jean-Baptiste Woillard and Pierre Marquet wrote the manuscript.

References

1. Allison, A. C. & Eugui, E. M. Mycophenolate mofetil and its mechanisms of action. *Immunopharmacology* **47**, 85–118 (2000).
2. Bergan, S. *et al.* Mycophenolate Personalized Therapy: Consensus Report by the International Association of Therapeutic Drug Monitoring and Clinical Toxicology. *Therapeutic Drug Monitoring* **unpublished**,
3. ISBA website. Accessed 14 October 2020. at <<https://pharmaco.chu-limoges.fr/>>
4. Debord, J. *et al.* Application of a gamma model of absorption to oral cyclosporin. *Clin Pharmacokinet* **40**, 375–382 (2001).
5. Prémaud, A. *et al.* A double absorption-phase model adequately describes mycophenolic acid plasma profiles in de novo renal transplant recipients given oral mycophenolate mofetil. *Clin Pharmacokinet* **44**, 837–847 (2005).
6. Saint-Marcoux, F. *et al.* Pharmacokinetic modelling and development of Bayesian estimators for therapeutic drug monitoring of mycophenolate mofetil in reduced-intensity haematopoietic stem cell transplantation. *Clin Pharmacokinet* **48**, 667–675 (2009).
7. Zahr, N. *et al.* Pharmacokinetic study of mycophenolate mofetil in patients with systemic lupus erythematosus and design of Bayesian estimator using limited sampling strategies. *Clin Pharmacokinet* **47**, 277–284 (2008).
8. Saint-Marcoux, F. *et al.* Development of a Bayesian estimator for the therapeutic drug monitoring of mycophenolate mofetil in children with idiopathic nephrotic syndrome. *Pharmacol. Res.* **63**, 423–431 (2011).
9. Woillard, J.-B. *et al.* Pharmacokinetics of mycophenolate mofetil in children with lupus and clinical findings in favour of therapeutic drug monitoring. *Br J Clin Pharmacol* **78**, 867–876 (2014).
10. Woillard, J.-B. *et al.* Mycophenolic mofetil optimized pharmacokinetic modelling, and exposure-effect associations in adult heart transplant recipients. *Pharmacol. Res.* **99**, 308–315 (2015).
11. Labriffe, M. *et al.* Population pharmacokinetics and Bayesian estimators for intravenous mycophenolate mofetil in haematopoietic stem cell transplant patients. *Br J Clin Pharmacol* (2020).doi:10.1111/bcp.14261
12. Lima, A. N. *et al.* Use of machine learning approaches for novel drug discovery. *Expert Opin Drug Discov* **11**, 225–239 (2016).
13. Hammann, F., Schöning, V. & Drewe, J. Prediction of clinically relevant drug-induced liver injury from structure using machine learning. *J Appl Toxicol* **39**, 412–419 (2019).
14. Stokes, J. M. *et al.* A Deep Learning Approach to Antibiotic Discovery. *Cell* **180**, 688-702.e13 (2020).

15. Hirayu, K., Kawano, H., Orii, H. & Tsuji, Y. Estimation of Blood Drug Concentration by LSTM Network. In *Proceedings of The 6th IIAE International Conference on Intelligent Systems and Image Processing 2018* 222–227 (The Institute of Industrial Application Engineers, 2018).doi:10.12792/icip2018.043
16. Tang, J. *et al.* Application of Machine-Learning Models to Predict Tacrolimus Stable Dose in Renal Transplant Recipients. *Sci Rep* **7**, 42192 (2017).
17. Imai, S., Takekuma, Y., Miyai, T. & Sugawara, M. A New Algorithm Optimized for Initial Dose Settings of Vancomycin Using Machine Learning. *Biol Pharm Bull* **43**, 188–193 (2020).
18. Woillard, J.-B. *et al.* A Machine Learning Approach to Estimate the Glomerular Filtration Rate in Intensive Care Unit Patients Based on Plasma Iohexol Concentrations and Covariates. *Clin Pharmacokinet* (2020).doi:10.1007/s40262-020-00927-6
19. Woillard, J.-B., Labriffe, M., Debord, J. & Marquet, P. Tacrolimus exposure prediction using machine learning. *Clinical Pharmacology & Therapeutics n/a*, (2020).
20. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16* 785–794 (2016).doi:10.1145/2939672.2939785
21. Badillo, S. *et al.* An Introduction to Machine Learning. *Clinical Pharmacology & Therapeutics* **107**, 871–885 (2020).
22. Ranking of Kaggle algorithms by competitions won | Data Science and Machine Learning. at <a>
23. Wickham, H. *et al.* Welcome to the Tidyverse. *Journal of Open Source Software* **4**, 1686 (2019).
24. Benkali, K. *et al.* Tacrolimus population pharmacokinetic-pharmacogenetic analysis and Bayesian estimation in renal transplant recipients. *Clin Pharmacokinet* **48**, 805–816 (2009).
25. Le Guellec, C. *et al.* Simultaneous estimation of cyclosporin and mycophenolic acid areas under the curve in stable renal transplant patients using a limited sampling strategy. *Eur J Clin Pharmacol* **57**, 805–811 (2002).
26. Djebli, N., Picard, N., Rérolle, J.-P., Le Meur, Y. & Marquet, P. Influence of the UGT2B7 promoter region and exon 2 polymorphisms and comedications on Acyl-MPAG production in vitro and in adult renal transplant patients. *Pharmacogenet Genomics* **17**, 321–330 (2007).
27. Schloerke, B. *et al.* *GGally: Extension to 'ggplot2'*. (2020).at <https://CRAN.R-project.org/package=GGally>
28. Sherwin, C. M. T., Fukuda, T., Brunner, H. I., Goebel, J. & Vinks, A. A. The evolution of population pharmacokinetic models to describe the enterohepatic recycling of mycophenolic acid in solid organ transplantation and autoimmune disease. *Clin Pharmacokinet* **50**, 1–24 (2011).

29. Kuhn, M. & Wickham, H. tidymodels: Easily Install and Load the 'Tidymodels' Packages version 0.1.0 from CRAN. at <<https://rdr.io/cran/tidymodels/>>
30. Brunet, M. *et al.* Therapeutic Drug Monitoring of Tacrolimus-Personalized Therapy: Second Consensus Report. *Ther Drug Monit* **41**, 261–307 (2019).
31. Winter, B. C. M. de *et al.* Bayesian estimation of mycophenolate mofetil in lung transplantation, using a population pharmacokinetic model developed in kidney and lung transplant recipients. *Clin Pharmacokinet* **51**, 29–39 (2012).
32. Musuamba, F. T. *et al.* Limited sampling models and Bayesian estimation for mycophenolic acid area under the curve prediction in stable renal transplant patients co-medicated with ciclosporin or sirolimus. *Clin Pharmacokinet* **48**, 745–758 (2009).
33. Le Guellec, C. *et al.* Population pharmacokinetics and Bayesian estimation of mycophenolic acid concentrations in stable renal transplant patients. *Clin Pharmacokinet* **43**, 253–266 (2004).
34. Chen, B. *et al.* Population Pharmacokinetics and Bayesian Estimation of Mycophenolic Acid Exposure in Chinese Renal Allograft Recipients After Administration of EC-MPS. *J Clin Pharmacol* **59**, 578–589 (2019).
35. Langers, P. *et al.* Limited sampling model for advanced mycophenolic acid therapeutic drug monitoring after liver transplantation. *Ther Drug Monit* **36**, 141–147 (2014).
36. Zhao, W. *et al.* Population pharmacokinetics and Bayesian estimator of mycophenolic acid in children with idiopathic nephrotic syndrome. *Br J Clin Pharmacol* **69**, 358–366 (2010).
37. Le Meur, Y. *et al.* Individualized mycophenolate mofetil dosing based on drug exposure significantly improves patient outcomes after renal transplantation. *Am. J. Transplant.* **7**, 2496–2503 (2007).
38. Metz, D. K. *et al.* Optimizing Mycophenolic Acid Exposure in Kidney Transplant Recipients: Time for Target Concentration Intervention. *Transplantation* **103**, 2012–2030 (2019).

Table 1: Characteristics of the ISBA requests used for the development and validation of the models

	Train (n=9658)	Test (n=3219)
Time between transplantation and MPA blood concentrations (years)	1.07 [0.25, 5.42]	1.06 [0.25, 5.73]
AUC0-12h (mg*h/L)	39.30 [29.11, 51.20]	39.24 [28.96, 50.88]
Patient age (year)	53.74 [38.58, 64.45]	52.85 [39.01, 64.10]
Morning dose (mg)	750 [500, 1000]	750 [500, 1000]
Deviation from the 20m theoretical time	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]
Deviation from the 1h theoretical time	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]
Deviation from the 3h theoretical time	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]
Concentration difference between C1h and C20m	1.70 [-1.44, 5.61]	1.70 [-1.58, 5.50]
Concentration difference between C1h and C3h	5.20 [2.03, 9.20]	5.39 [2.10, 9.30]
Concentration difference between C20m and C3h	2.10 [-0.90, 8.00]	2.30 [-0.80, 8.38]
Delayed absorption (C3>C1)	1293 (13.4)	409 (12.7)
Delayed absorption (C3>C20)	3228 (33.4)	1035 (32.2)
Delayed absorption (C3>C20 & C3>C1)	1145 (11.9)	379 (11.8)
Associated immunosuppressant		
Tacrolimus	5265 (54.5)	1811 (56.3)
Ciclosporine	2580 (26.7)	830 (25.8)
None	1542 (16.0)	483 (15.0)
Sirolimus/everolimus	271 (2.8)	95 (3.0)
Transplant type		
Kidney	7398 (76.6)	2470 (76.7)
Heart	782 (8.1)	246 (7.6)
Nephrotic syndrome	524 (5.4)	187 (5.8)
Liver	296 (3.1)	100 (3.1)
Lung	202 (2.1)	79 (2.5)
Lupus	97 (1.0)	30 (0.9)
Pediatric lupus	45 (0.5)	9 (0.3)

Bone marrow	34 (0.4)	13 (0.4)
Other	280 (2.9)	85 (2.6)

(median [IQR]) are presented for continuous data and n(%) for categorical data.

Table 2: Performances of the models in the training and testing datasets to estimate AUC0-12 obtained from 3 samples MAP-BE from ISBA

		MPA 2 samples	MPA 3 samples
training	*RMSE \pm SD (mg*h/L)	6.78 \pm 0.08	5.59 \pm 0.07
	*R² \pm SD	0.864 \pm 0.004	0.908 \pm 0.002
	Relative MPE (%)	2.7	1.8
	Relative RMSE (%)	18.6	14.5
testing	RMSE (μg*h/L)	6.41	5.34
	R²	0.876	0.914
	Relative MPE (%)	2.1	1.4
	Relative RMSE (%)	17.0	13.2
	Number of MPE out of the \pm20% interval n(%)	560 (17.4%)	306 (9.5%)

*Values and standard deviations obtained after 10 fold cross-validation,

Table 3: Performances of ML models and MAP-BE based on 3 samples LSS to estimate reference AUC0-12s obtained from full PK profiles using the trapezoidal rule.

Study	Method	All patients				Patients in the first month post transplantation			
		Relative MPE (%)	Relative RMSE (%)	Bias out of $\pm 20\%$ n (%)	Bias out of $\pm 10\%$ n (%)	Relative MPE (%)	Relative RMSE (%)	Bias out of $\pm 20\%$ n (%)	Bias out of $\pm 10\%$ n (%)
PCCP study (n=128 / n 1 st month = 49)	Xgboost 2 concentrations	-1.2	24.2	46 (35.9)	80 (62.5)	-1.1	24.1	20 (40.8)	67.3 (33)
	Xgboost 3 concentrations	-4.0	19.0	39 (30.5)	76 (59.4)	-4.9	19.2	15 (30.6)	57.1 (28)
	MAP-BE 3 concentrations	1.0	19.9	37 (28.9)	81 (63.3)	2.4	20.0	15 (30.6)	61.1 (30)
Stablocine (n=20)	Xgboost 2 concentrations	1.5	17.8	4 (20)	10 (50)	NA	NA	NA	NA
	Xgboost 3 concentrations	2.8	14.7	2 (10)	8 (40)	NA	NA	NA	NA
	MAP-BE 3 concentrations	9.3	20.5	7 (35)	9 (45)	NA	NA	NA	NA
CONCEPT (n=67)	Xgboost 2 concentrations	-0.8	26.3	27 (40)	42 (63)	NA	NA	NA	NA
	Xgboost 3 concentrations	-1.1	22.5	14 (21)	43 (64)	NA	NA	NA	NA
	MAP-BE 3 concentrations	2.1	25.1	18 (27)	40 (60)	NA	NA	NA	NA
PIGREC (n=75/ n 1 st month = 20)	Xgboost 2 concentrations	2.5	23.1	31 (41)	44 (59)	7.7	20.4	7 (35)	12 (60)
	Xgboost 3 concentrations	-1.2	15.6	17 (23)	41 (55)	3.5	14.3	6 (30)	10 (50)
	MAP-BE 3 concentrations	3.5	17.3	17 (23)	47 (63)	4.6	19.1	5 (25)	11 (55)

Figure 1: Scatter plots and residual plots of predicted vs reference interdose AUC0-12s in the test set, for MPA using 2 or 3 MPA plasma concentrations.

Figure 2: Scatter plots and residual plots of predicted vs reference interdose AUC0-12s in the external validation studies split into renal and heart transplantation. The MPA AUC0-12 is estimated using the Xgboost model from 2 or 3 MPA plasma concentrations and MAP-BE is the maximum a posteriori Bayesian estimation based on 3 samples and the parametric POPPK models available on ISBA.

Figure 3: Comparison of the different methods of MPA AUC0-12 estimation in the external validation studies. MAP-BE is maximum a posteriori Bayesian estimation based on the 20min, 1 and 3h limited sampling strategy and the parametric POPPK models available in ISBA. The colour blue refers to kidney transplant patients and red, heart transplant patients.