

## Table list:

- I. Demographic and clinical data for 22 renal transplant patients
- II. Performance analysis of ML models
- III. Performances of the generalized linear regression algorithms in the training and testing datasets
- IV. Performances of the generalized linear regression algorithms in the simulated validation datasets
- V. Impact of imperfect sampling times on the performance of XGBoost and MAP-BE models in the simulated patients of the validation set
- VI. Training of XGBoost models separately on simulations from each popPK model (Vezina et al. and Lalagkas et al.) and comparison of their performance in predicting AUCss in the validation simulated patient's dataset

Table S1: Demographic and clinical data for 22 renal transplant patients

| Age (years), median (range)                      | 57.5 (25 – 76)                             |                             |
|--------------------------------------------------|--------------------------------------------|-----------------------------|
| Male/female, n/n                                 | 18/4                                       |                             |
| Mycophenolate mofetil yes/no, n/n                | 12/10                                      |                             |
| Creatinine clearance (mL/min), median (range)    | 45.7 (18.8 – 95.7)                         |                             |
| CMV serology +/-2, n/n                           | 10/12                                      |                             |
|                                                  | First blood sample after drug introduction | Blood sample at PK sampling |
| Hemoglobin concentration (g/L), median (range)   | 11.6 (7.85 – 13.7)                         | 12.1 (9.29 – 15.6)          |
| Platelets (10 <sup>9</sup> /L), median (range)   | 198 (117 – 532)                            | 204 (139 – 337)             |
| Lymphocytes (10 <sup>9</sup> /L), median (range) | 1.63 (0.35 – 5.44)                         | 1.51 (0.28 – 3.59)          |
| Neutrophils (10 <sup>9</sup> /L), median (range) | 5.31 (1.78 – 8.88)                         | 3.85 (0.60 – 8.04)          |

Table S2: Performance analysis of ML models

|                       | ML algorithms | RMSE, mg.h/L | R <sup>2</sup> | Relative MPE, % | Relative RMSE, % |
|-----------------------|---------------|--------------|----------------|-----------------|------------------|
| 900 mg/24h<br>of VGCV | XGBoost       | 0.83         | 0.99           | 0.032           | 1.30             |
|                       | Random forest | 1.56         | 0.99           | 3.510           | 2.18             |
|                       | glmnet        | 4.23         | 0.96           | -0.434          | 7.97             |
|                       | MARS          | 3.89         | 0.97           | 0.325           | 6.56             |
|                       | SVM           | 0.77         | 0.99           | 0.063           | 1.48             |
| 450 mg/24h<br>of VGCV | XGBoost       | 0.58         | 0.99           | 0.017           | 1.13             |
|                       | Random forest | 1.00         | 0.99           | 0.004           | 1.89             |
|                       | glmnet        | 3.88         | 0.95           | -0.369          | 9.65             |
|                       | MARS          | 2.83         | 0.97           | 0.268           | 6.55             |
|                       | SVM           | 0.21         | 0.99           | 0.077           | 0.54             |
| 450 mg/48h<br>of VGCV | XGBoost       | 1.25         | 0.99           | 0.031           | 1.71             |
|                       | Random forest | 1.98         | 0.99           | 0.039           | 2.78             |
|                       | glmnet        | 5.51         | 0.96           | -0.809          | 10.7             |
|                       | MARS          | 4.73         | 0.97           | 0.905           | 9.64             |
|                       | SVM           | 0.63         | 0.99           | -0.040          | 0.90             |
| 450 mg/72h<br>of VGCV | XGBoost       | 3.72         | 0.99           | 0.014           | 2.85             |
|                       | Random forest | 5.49         | 0.99           | 0.341           | 4.45             |
|                       | glmnet        | 10.8         | 0.96           | -0.652          | 10.3             |
|                       | MARS          | 8.98         | 0.97           | 0.205           | 8.48             |
|                       | SVM           | 1.91         | 0.99           | 0.080           | 1.50             |

Values obtained after 10-fold cross-validation on the training set

Table S3: Performances of the generalized linear regression algorithms in the training and testing datasets

| <b>VGCV dosage</b> | <b>Data set</b>          | <b>RMSE, mg.h/L</b> | <b>R<sup>2</sup></b> | <b>Relative MPE, %</b> | <b>Relative RMSE, %</b> | <b>Number of MPE of the <math>\pm 20\%</math> interval, <i>n</i> (%)</b> |
|--------------------|--------------------------|---------------------|----------------------|------------------------|-------------------------|--------------------------------------------------------------------------|
| 900 mg/24h         | Training set (n = 3,671) | 8.86 <sup>a</sup>   | 0.84 <sup>a</sup>    | -                      | -                       | -                                                                        |
|                    | Testing set (n = 1,224)  | 8.86                | 0.84                 | 1.43                   | 14.5                    | 200 (16.3 %)                                                             |
| 450 mg/24h         | Training set (n = 3,660) | 5.85                | 0.89 <sup>a</sup>    | -                      | -                       | -                                                                        |
|                    | Testing set (n = 1,223)  | 6.33                | 0.88                 | 0.05                   | 13.4                    | 162 (13.2 %)                                                             |
| 450 mg/48h         | Training set (n = 3,672) | 11.1 <sup>a</sup>   | 0.82 <sup>a</sup>    | -                      | -                       | -                                                                        |
|                    | Testing set (n = 1,225)  | 11.0                | 0.82                 | 2.22                   | 17.4                    | 223 (18.2 %)                                                             |
| 450 mg/72h         | Training set (n = 1,835) | 25.1 <sup>a</sup>   | 0.76 <sup>a</sup>    | -                      | -                       | -                                                                        |
|                    | Testing set (n = 612)    | 25.0                | 0.77                 | 3.53                   | 20.1                    | 213 (34.8 %)                                                             |

<sup>a</sup> Values obtained after 10-fold cross-validation

Table S4: Performances of the generalized linear regression algorithms in the simulated validation datasets

| VGCV dosage | Performances                                | Patients from                           |                                       |
|-------------|---------------------------------------------|-----------------------------------------|---------------------------------------|
|             |                                             | Caldés et al.<br>model ( <i>n</i> = 98) | Chen et al.<br>model ( <i>n</i> = 98) |
| 900 mg/24h  | Relative MPE, (%)                           | -16.9                                   | -13.8                                 |
|             | Relative RMSE, (%)                          | 19.3                                    | 24.0                                  |
|             | Number of MPE of the<br>$\pm 20\%$ interval | 31 (31.6 %)                             | 31 (31.6 %)                           |
| 450 mg/24h  | Relative MPE, (%)                           | -14.1                                   | -16.8                                 |
|             | Relative RMSE, (%)                          | 17.4                                    | 32.7                                  |
|             | Number of MPE of the<br>$\pm 20\%$ interval | 23 (23.5 %)                             | 37 (37.8 %)                           |
| 450 mg/48h  | Relative MPE, (%)                           | -25.7                                   | 0.05                                  |
|             | Relative RMSE, (%)                          | 27.7                                    | 37.5                                  |
|             | Number of MPE of the<br>$\pm 20\%$ interval | 69 (70.4 %)                             | 71 (72.4 %)                           |
| 450 mg/72h  | Relative MPE, (%)                           | -28.4                                   | -                                     |
|             | Relative RMSE, (%)                          | 31.1                                    | -                                     |
|             | Number of MPE of the<br>$\pm 20\%$ interval | 73 (74.5 %)                             | -                                     |

Table S5: Impact of imperfect sampling times on the performance of XGBoost and MAP-BE models in the simulated patients of the validation set

| AUCss estimation method |                        | Performances                             | Patients from                |                            |
|-------------------------|------------------------|------------------------------------------|------------------------------|----------------------------|
|                         |                        |                                          | Caldés et al. model (n = 98) | Chen et al. model (n = 98) |
| 2 samples (C4 & C10)    |                        |                                          |                              |                            |
| 900 mg/24h of VGCV      | XGBoost (C4 & C12)     | Relative MPE, (%)                        | 7.79                         | 27.1                       |
|                         |                        | Relative RMSE, (%)                       | 17.5                         | 29.3                       |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 30 (30.6%)                   | 74 (75.5 %)                |
|                         | MAP-BE Vezina et al.   | Relative MPE, (%)                        | -12.5                        | 19.1                       |
|                         |                        | Relative RMSE, (%)                       | 14.3                         | 23.5                       |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 15 (15.3 %)                  | 41 (41.9 %)                |
|                         | MAP-BE Lalagkas et al. | Relative MPE, (%)                        | 42.7                         | 78.4                       |
|                         |                        | Relative RMSE, (%)                       | 48.0                         | 82.9                       |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 83 (84.7 %)                  | 97 (99.0 %)                |
| 2 samples (C4 & C10)    |                        |                                          |                              |                            |
| 450 mg/24h of VGCV      | XGBoost (C4 & C12)     | Relative MPE, (%)                        | 9.68                         | 25.7                       |
|                         |                        | Relative RMSE, (%)                       | 15.7                         | 28.0                       |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 17 (17.3%)                   | 71 (72.5 %)                |
|                         | MAP-BE Vezina et al.   | Relative MPE, (%)                        | 17.2                         | 14.1                       |
|                         |                        | Relative RMSE, (%)                       | 11.0                         | 18.3                       |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 7 (7.14 %)                   | 29 (29.6 %)                |
|                         | MAP-BE Lalagkas et al. | Relative MPE, (%)                        | 34.8                         | 76.5                       |
|                         |                        | Relative RMSE, (%)                       | 43.4                         | 80.4                       |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 71 (72.4%)                   | 98 (100 %)                 |
| 2 samples (C0 & C10)    |                        |                                          |                              |                            |
| 450 mg/48h of VGCV      | XGBoost (C0 & C12)     | Relative MPE, (%)                        | 4.46                         | 29.7                       |
|                         |                        | Relative RMSE, (%)                       | 15.6                         | 36.4                       |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 17 (17.3 %)                  | 57 (58.2 %)                |
|                         | MAP-BE Vezina et al.   | Relative MPE, (%)                        | -16.0                        | 24.3                       |
|                         |                        | Relative RMSE, (%)                       | 19.0                         | 32.7                       |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 32 (32.6 %)                  | 52 (53.1 %)                |
|                         | MAP-BE Lalagkas et al. | Relative MPE, (%)                        | 24.4                         | 76.1                       |
|                         |                        | Relative RMSE, (%)                       | 29.4                         | 79.3                       |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 57 (58.2 %)                  | 98 (100 %)                 |
| 2 samples (C0 & C6)     |                        |                                          |                              |                            |
| 450 mg/72h of VGCV      | XGBoost (C0 & C7)      | Relative MPE, (%)                        | 12.0                         | -                          |
|                         |                        | Relative RMSE, (%)                       | 14.7                         | -                          |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 14 (14.3 %)                  | -                          |
|                         | MAP-BE Lalagkas et al. | Relative MPE, (%)                        | 31.1                         | -                          |
|                         |                        | Relative RMSE, (%)                       | 37.1                         | -                          |
|                         |                        | Number of MPE of the $\pm 20\%$ interval | 80 (81.6%)                   | -                          |

Table S6: Training of XGBoost models separately on simulations from each popPK model (Vezina et al. and Lalagkas et al.) and comparison of their performance in predicting AUCss in the validation simulated patient's dataset. This comparison was conducted exclusively on simulated patients receiving 900 mg/24h of valganciclovir.

| AUCss estimation method |                           | Performances                             | Patients from                |                            |                              |                            |                              |                            |
|-------------------------|---------------------------|------------------------------------------|------------------------------|----------------------------|------------------------------|----------------------------|------------------------------|----------------------------|
|                         |                           |                                          | Caldés et al. model (n = 98) | Chen et al. model (n = 98) | Caldés et al. model (n = 98) | Chen et al. model (n = 98) | Caldés et al. model (n = 98) | Chen et al. model (n = 98) |
|                         |                           |                                          | 2 samples (C4 and C12)       |                            | 3 samples (C3, C4 and C12)   |                            | 3 samples (C3, C4 and C6)    |                            |
| 900 mg/24h of VGCV      | XGBoost                   | Relative MPE, (%)                        | - 0.10                       | 15.0                       | 1.54                         | 8.89                       | 5.16                         | 18.1                       |
|                         |                           | Relative RMSE, (%)                       | 13.6                         | 17.5                       | 10.6                         | 10.7                       | 15.5                         | 23.2                       |
|                         |                           | Number of MPE of the $\pm 20\%$ interval | 12 (12.2%)                   | 33 (33.7 %)                | 5 (5.10 %)                   | 2 (2.04 %)                 | 14 (14.3 %)                  | 46 (46.9%)                 |
|                         | XGBoost (Lalagkas et al.) | Relative MPE, (%)                        | 0.08                         | 16.4                       | 1.06                         | 7.72                       | 8.09                         | 19.9                       |
|                         |                           | Relative RMSE, (%)                       | 13.3                         | 18.8                       | 9.78                         | 9.61                       | 16.0                         | 25.0                       |
|                         |                           | Number of MPE of the $\pm 20\%$ interval | 12 (12.2 %)                  | 36 (36.7 %)                | 4 (4.08 %)                   | 1 (1.02 %)                 | 19 (19.4 %)                  | 54 (55.1 %)                |
|                         | XGBoost (Vezina et al.)   | Relative MPE, (%)                        | - 2.93                       | 13.5                       | 2.23                         | 12.5                       | - 2.93                       | 19.7                       |
|                         |                           | Relative RMSE, (%)                       | 15.8                         | 16.6                       | 14.6                         | 14.8                       | 17.7                         | 25.8                       |
|                         |                           | Number of MPE of the $\pm 20\%$ interval | 15 (15.3 %)                  | 26 (26.5 %)                | 11 (11.2 %)                  | 14 (14.3 %)                | 25 (25.5 %)                  | 49 (50 %)                  |

## Figure list :

- I. Scatterplots of biases as a function of reference AUCss in the validation set

Figure S1: Scatterplots of biases as a function of reference AUCss in the validation set

