*The AAPS Journal (2019) 21: 40*
DOI: 10.1208/s12248-019-0299-9

*Research Article*

# A Receiver Operating Characteristic Framework for Non-adherence Detection Using Drug Concentration Thresholds—Application to Simulated Risperidone Data in Schizophrenic Patients

**Carlos Pérez-Ruixo,[1,3] Bart Remmerie,[1] Juan José Peréz-Ruixo,[1] and An Vermeulen[2]**

*Received 19 September 2018; accepted 12 January 2019; published online 14 March 2019*

***Abstract.*** Non-adherence to antipsychotic medication is a primary factor in disease relapse in schizophrenic patients. We sought to evaluate if plasma concentrations of the antipsychotic risperidone can be used as a predictor of treatment adherence and to identify the optimal plasma concentration threshold to reliably distinguish between adherent and non-adherent patients. A population pharmacokinetic model was used to simulate plasma risperidone steady-state trough concentrations in 1000 virtual patients, where 60% of the patients were 100% adherent to their medication, while 40% of the patients were non-adherent to their medication. The probability of adherence was assessed by receiver operating characteristic (ROC) analysis on $C_{trough}$. The area under the ROC curve ($AUC_{ROC}$) was used to identify the optimal $C_{trough}$ threshold. Single *vs* multiple $C_{trough}$ at steady state was also evaluated. After a single risperidone $C_{trough}$ measurement, the $AUC_{ROC}$ (95% CI) was estimated to be 0.71 (0.69–0.72) and the optimal $C_{trough}$ threshold accounting for the lowest number of adherent and non-adherent misclassifications was estimated to be 11.9 ng/mL. After multiple $C_{trough}$ measurements, the $AUC_{ROC}$ (95% CI) increased up to 0.85 (0.84–0.87) for three $C_{trough}$ measurements. The optimal probability threshold to reliably discriminate between adherent and non-adherent patients was estimated to be 0.51. Using this model which is reflective of typical adherence to antipsychotic medication, we found that three consecutive steady-state $C_{trough}$ measurements are needed for an accurate and precise diagnostic test to discriminate between patients who are adherent or non-adherent to treatment.

**KEY WORDS:** diagnostic test; drug adherence; drug plasma concentrations; population PK model–based simulation; ROC analysis.

## INTRODUCTION

Schizophrenia is a chronic, relapsing, serious mental disorder which affects more than 21 million people worldwide (1). Antipsychotic medication is a key component of the treatment to promote recovery and improve the quality of life and functional outcomes in patients with schizophrenia (2). Non-adherence to prescribed antipsychotic treatments has been recognized as a health-system problem worldwide since it is the primary factor in disease relapse, including the exacerbation of psychosis, increased clinic and emergency room visits, and rehospitalization (3) that results in high health care and societal costs (4,5).

Non-adherence is possibly the most challenging aspect of antipsychotic treatment (6) since it affects approximately 40% of the patients (3,7). Non-adherent patients are those considered to miss more than 20% of the dose intakes (7). This cutoff has validity in predicting subsequent hospitalization across several chronic conditions (8). However, for individual patients, the degree of non-adherence that affects health outcomes will vary and depend on multiple factors including the condition, its severity, the risk of recurrence, the relative effectiveness of the medication, and its dose and frequency of administration.

There is a critical need for accurate adherence assessments of orally administered antipsychotic agents. The most commonly used methods are either subjective measures like medication self-reporting and physician or treatment provider reports, or objective measures such as pill counts and electronic monitors (9). Each method to assess adherence has its own drawbacks, and adherence assessments in themselves can influence adherence behavior. Antipsychotic plasma levels may be used to monitor treatment compliance

[1] Janssen Research & Development, Clinical Pharmacology & Pharmacometrics, Beerse, Belgium.
[2] Janssen Research & Development, Quantitative Sciences Consulting, Beerse, Belgium.
[3] To whom correspondence should be addressed. (e–mail: cperezru@ITS.JNJ.com)

when adherence is in doubt or response is inadequate (10). Because antipsychotics show a large variability in response, and improvement of symptoms is not linked to a well-defined therapeutic window, antipsychotic concentrations are challenging to interpret. To this aim, we developed an empirical approach using reference ranges to validate plasma drug concentration thresholds which are predictive of non-adherence to risperidone in patients with schizophrenia (11).

The lack of standard methods to reliably discriminate between adherent and non-adherent patients with schizophrenia also warrants the development of scientifically based quantitative approaches. Several applications of receiver operating characteristic (ROC) curves in the field of pharmacokinetics (PK) have been published. Suzuki Y *et al.* analyzed the pharmacokinetic-pharmacodynamic relationship of vancomycin and used the ROC curves to determine if the drug exposure parameters correlate with the efficacy and nephrotoxicity of vancomycin (12), Olivares-Morales A *et al.* developed a model for the categorical prediction of human oral bioavailability from animal data by employing a threshold decision tool based on ROC analysis (13), and most recently, Clements JD *et al.* used a combination of ROC analysis together with clinical trial simulations to optimize drug titration in clinical trials by maximizing efficacy while minimizing safety concerns (14).

Population adherence modeling approaches have been used in a variety of therapeutic areas. Hiligsmann *et al.* (15) used a series of adherence models based on Markov microsimulation to evaluate the potential clinical and economic implications of non-adherence to bisphosphonate therapy by mixing and testing different scenarios for both compliance and persistence. A predictive model for non-adherence to glaucoma medications was developed using stepwise logistic regression and was validated with a separate cohort of patients using the criteria of equality of regression coefficients, discrimination, and calibration (16).

Here, we propose for the first time the use of ROC analyses (17,18) in combination with population PK model–based simulation techniques to evaluate if measuring antipsychotic plasma concentrations can be used as a predictor of treatment adherence. The methodology described here can be applied to other antipsychotic treatments as well as to other treatments of different diseases where identifying the lack of adherence could lead to successful clinical outcomes.

## METHODS

A previous population PK model developed using data from four Phase 1 and three Phase 3 studies was used to perform the model-based simulation scenarios (19). The model incorporated a two-compartment disposition model for both risperidone and paliperidone, with first-order elimination from each of the central compartments (Fig. 1). Input was described as a zero-order first-order process after an absorption lag time (ALAG1), with a fraction of the dose from the depot compartment being converted to the paliperidone metabolite via first-pass metabolism (FP). Apparent central volumes of distribution for the parent and metabolite ($V_{c,PAR}/F$ and $V_{c,MET}/F$, respectively) were assumed to be the same, as were the absorption rate constants (KA) from the depot compartment to $V_{c,PAR}/F$ and $V_{c,MET}/F$. As the formation of paliperidone is primarily mediated by CYP2D6, a mixture model for poor, intermediate, or extensive metabolizing status was included to describe variability in apparent clearance from parent to metabolite ($CL_{PM}/F$). Concomitant carbamazepine (CBZ) was included as a covariate on apparent clearance of both the parent and metabolite ($CL_{PAR}/F$ and $CL_{MET}/F$). Between patient variability (BSV) was included on absorption terms ALAG1, KA, FP, and duration of input (D1), in addition to $CL_{PAR}/F$, $CL_{PM}/F$, $CL_{MET}/F$, intercompartmental clearance of the parent drug ($Q_{PAR}/F$), and $V_{c,PAR}/F$ and $V_{c,MET}/F$. In addition, between occasion variability (BOV) was incorporated on apparent bioavailability (F1), D1, and KA. The model was developed using a log transformation of both sides

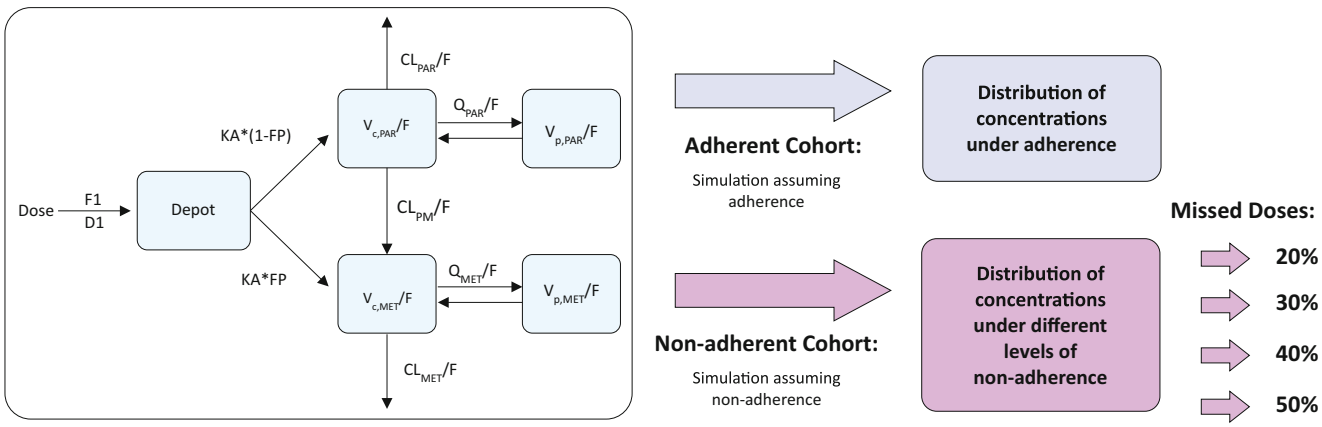


**Fig. 1.** Schematic of the structural population PK model for risperidone and the simulated scenarios. Among the 1000 virtual patients, 60% were adherent ($N = 600$) and the remaining 400 patients were non-adherent. From the non-adherent cohort ($N = 400$), there were different levels of non-compliance, meaning that each 100 individuals from the non-adherent cohort will miss 20% ($N = 100$), 30% ($N = 100$), 40% ($N = 100$), and 50% ($N = 100$) of the doses respectively. Abbreviations: *F1* bioavailability, *D1* duration of zero-order input into the depot compartment, *KA* absorption rate constant, *FP* fraction of dose converted to 9-OH (first-pass metabolism), $V_{c,PAR}/F$ apparent central volume of distribution of parent, $V_{c,MET}/F$ apparent central volume of distribution of metabolite, $V_{p,PAR}/F$ apparent peripheral volume of distribution of parent, $V_{p,MET}/F$ apparent peripheral volume of distribution of metabolite, $CL_{PAR}/F$ apparent clearance of parent, $CL_{PM}/F$ apparent clearance from parent to metabolite, $CL_{MET}/F$ apparent clearance of metabolite, $Q_{PAR}/F$ apparent intercompartmental clearance of parent, $Q_{MET}/F$ apparent intercompartmental clearance of metabolite

approach and the residual unexplained variability (RUV) for risperidone and the metabolite was additive in the log domain. The final model parameter estimates, shown in Table I with a schematic diagram presented in Fig. 1, were obtained with the first-order estimation method in NONMEM®.

A systematic review of 39 studies reported a mean prevalence rate of drug treatment adherence in schizophrenia patients of approximately 60% (3). Therefore, the described population PK model of risperidone was used to simulate steady-state plasma risperidone active moiety (risperidone + active metabolite 9-hydroxy-risperidone) at 24 h after the last dose ($C_{trough}$). From the 1000 virtual patients simulated, a total of 60% were adherent ($N = 600$) while the remaining were non-adherent. From the non-adherent cohort ($N = 400$), there were different levels of non-compliance meaning that each 100 individuals from the non-adherent cohort will miss 20% ($N = 100$), 30% ($N = 100$), 40% ($N = 100$), and 50% ($N = 100$) of the doses respectively. The diagram presented in Fig. 1 illustrates the described simulated cohorts.

The $C_{trough}$ was assessed as a predictor of risperidone adherence using a ROC analysis among the two cohorts of patients simulated (20). ROC analysis is used in clinical epidemiology to quantify how accurate medical diagnostic tests can discriminate between two patient states and to identify the best predictive threshold or cutoff value that leads to the minimal number of misclassifications between states (21,22). The ROC curve was based on the notion of a "separator" scale, where a predictor from two different populations forms a pair of overlapping distributions. In this case, the overlapping distributions of $C_{trough}$ come from the adherent and non-adherent cohorts. The complete separation of the two underlying distributions implies a perfectly discriminating test while a complete overlap implies no discrimination (20,23). In a ROC curve, the true positive rate (sensitivity) is plotted as the function of the false positive rate

**Table I.** Risperidone-Paliperidone Population PK Model Parameter Estimates

| Parameter | Parameter estimate SE (%CV) | BSV estimate SE (%CV) | BOV estimate SE (%CV) |
|---|---|---|---|
| Apparent bioavailability (F1) | | – | 46.5 (24.4) |
| Absorption lag time (ALAG1, h) | 0.165 (9.4) | 41.0 (69.0) | – |
| Duration of input (D1, h) | 0.458 (7.9) | 113 (38.3) | 158 (25.9) |
| Absorption rate constant (KA, /h) | 2.34 (16.9) | 149 (32.5) | 108 (33.6) |
| Fraction of dose converted to paliperidone (%) | | 117[a] (17.9) | – |
| Poor metabolizers | 1.75 (47.3) | | |
| Intermediate metabolizers | 10.9 (38.0) | | |
| Extensive metabolizers | 41.3 (7.0) | | |
| Apparent intercompartmental clearance of parent ($Q_{PAR}$/F, L/h) | 3.65 (14.8) | 215 (131) | – |
| Apparent clearance of parent ($CL_{PAR}$/F, L/h) | | 184 (72.0) | – |
| No concomitant CBZ (L/h) | 2.84 (28.1) | | |
| With concomitant CBZ (L/h) | 6.49 (33.0) | | |
| Apparent clearance from parent to metabolite ($CL_{PM}$/F, L/h) | | 33.3 (27.7) | – |
| Poor metabolizers | 1.18 (16.5) | | |
| Intermediate metabolizers | 4.37 (22.4) | | |
| Extensive metabolizers | 19.6 (10.8) | | |
| Apparent central volume of distribution of parent ($V_{c,PAR}$/F, L) | 137 (8.3) | 30.0 (61.5) | – |
| Apparent peripheral volume of distribution of parent ($V_{p,PAR}$/F, L) | 100 (23.6) | 53.9 (91.8) | – |
| Apparent intercompartmental clearance of metabolite ($Q_{MET}$/F, L/h) | 1.67 (20.3) | 0 FIXED | – |
| Apparent clearance of metabolite ($CL_{MET}$/F, L/h) | | 20.4 (65.5) | – |
| No concomitant CBZ (L/h) | 5.99 (5.4) | | – |
| With concomitant CBZ (L/h) | 6.22 (10.2) | | – |
| Apparent central volume of distribution of metabolite ($V_{c,MET}$/F, L) | 137 (8.3) | 30.0 (61.5) | – |
| Apparent peripheral volume of distribution of metabolite ($V_{p,MET}$/F, L) | 91.8 (16.0) | 80.7 (68.0) | – |
| Mixture for metabolizing status (%) | | – | |
| Poor metabolizers | 4.93 (34.5) | | |
| Intermediate metabolizers | | | |
| Single-dose trials | 23.5 (32.3) | | |
| All other trials | 8.26 (30.5) | | |
| Extensive metabolizers | | | |
| Single-dose trials | 71.6 | | |
| All other trials | 86.8 | | |
| RUV risperidone (%CV) | 30.5 (18.4) | | |
| RUV paliperidone (%CV) | 36.2 (13.4) | | |

*%CV* percent coefficient of variation, *BOV* between occasion variability, *BSV* between patient variability, *CBZ* carbamazepine, *h* hour, *L* liter, *RUV* residual unexplained variability, *SD* standard deviation, *SE* standard error

[a] %SD in the logit domain

[b] Same as $V_{c,PAR}$

(1-specificity) for different cutoff points of a given predictor, meaning that each point on the ROC curve represents a sensitivity/specificity pair corresponding to a decision threshold. The sensitivity is inversely related to specificity, in the sense that sensitivity increases as specificity decreases across various thresholds; hence, the ROC curve graphically displays a full picture of the trade-off between the sensitivity ($y$-axis) and 1-specificity ($x$-axis) across a series of thresholds or cutoff levels (22,24). For each simulated $C_{trough}$, the sensitivity represents the proportion of patients who are adherent to risperidone and were identified correctly and was calculated according to Eq. 1:

$$\text{Sensitivity} = (\text{TP})/(\text{TP} + \text{FN}) \quad (1)$$

where, TP is the rate of true positive and FN the rate of false negative, understanding TP as a risperidone adherent patient who has been correctly identified as being adherent, while FN is considered to represent an adherent patient who has been incorrectly identified as being non-adherent. Similarly, for each simulated $C_{trough}$, the specificity represents the proportion of patients who are non-adherent and were identified correctly and was calculated according to Eq. 2:

$$\text{Specificity} = (\text{TN})/(\text{TN} + \text{FP}) \quad (2)$$

where, TN is the rate of true negative and FP the rate of false positive, understanding TN as a non-adherent patient who has been identified as being non-adherent, while FP is a non-adherent patient who has been identified as being adherent. The 95% confidence intervals (95% CI) of the sensitivity and specificity were computed using the Wilson exact approach (25).

The area under the ROC curve ($AUC_{ROC}$) is an effective measure of the accuracy of a diagnostic test, or in other words, how well the predictor (in this case, $C_{trough}$) can distinguish between two diagnostic states (in this case, adherent and non-adherent). The larger the $AUC_{ROC}$, the better the overall performance of the diagnostic test to correctly identify between two patients' states. An $AUC_{ROC}$ of 1 represents a perfect test, while an $AUC_{ROC}$ of 0.50 reflects that the performance of a diagnostic test is no better than chance (i.e., tossing a fair coin (23)). The $AUC_{ROC}$ was calculated using the trapezoidal rule (20). The 95% CI of the $AUC_{ROC}$ was calculated using the Hanley and McNeil method (23). It was assumed that $AUC_{ROC}$ values higher than 0.80 represent an accurate diagnostic test (20,23).

To use the ROC analysis for patient management, a decision threshold or cutoff must be identified. Two major elements determine which of the possible sensitivity/specificity combinations (and the corresponding decision threshold) is the most appropriate for a particular application of the test (26): (a) The relative cost or undesirability of errors (FP and FN classifications), and the value or benefits of correct classifications (27) (Eq. 3); and (b) the relative proportions of the two states that the test is intended to discriminate between, and which are related to the prevalence (Eq. 4).

Assigning cost to FP or FN classifications is complex. This can be expressed in terms of health costs or financial costs and can be viewed from the perspective of the patient, the care providers, the insurers, dependents, and society. Nonetheless, some judgment about the relative costs of false results should be made when selecting rationally an operating decision threshold. In order to define the optimal predictive cutoff value that will allow to reliably distinguish between adherent and non-adherent patients with schizophrenia, the efficiency criterion was calculated for each potential cutoff value (20) as described in Eq. 3:

$$\text{Efficiency} = \text{Sensitivity} - \text{Cost} \times (100 - \text{Specificity}) \quad (3)$$

where, sensitivity and specificity were expressed as percentages, and the overall cost of misclassification was calculated as indicated in Eq. 4:

$$\text{Cost} = (\text{FPC})/(\text{FNC}) \times (1 - \text{Prev})/\text{Prev} \quad (4)$$

where Prev is the prevalence of patients being fully compliant with risperidone (60%) while the ratio between the false positive cost (FPC) and false negative cost (FNC) represents the relative cost of misclassification. When this ratio is 1, meaning that the cost of identifying a non-adherent patient as being adherent (FP) is equal to the cost of identifying an adherent patient as being non-adherent (FN), the cost in Eq. 4 becomes a slope function within what is otherwise known as the Youden's J statistic, a tool used to assess the performance of a diagnostic test for a given cutoff value as defined in Eq. 5:

$$\text{J} = \text{Sensitivity} + \text{Specificity} - 100 \quad (5)$$

Based on a Youden's J index, the best predictive cutoff value should be the one that maximizes the sum of sensitivity + specificity and corresponds to the maximum vertical distance between the ROC curve and the diagonal line. Youden's J index is used to assess the performance of a diagnostic test and represents an optimal trade-off between sensitivity and specificity taking into account an equal assigned weight cost to the sensitivity and the specificity (28). In this analysis, the optimal cutoff was calculated based on Youden's J index. Furthermore, the impact of the FPC to FNC ratio on the cutoff value was evaluated by giving a three-fold higher cost to the FP, since the goal was to minimize the FP rate. Based on this assumption, the diagnostic test parameters $AUC_{ROC}$, sensitivity, specificity, PPV and NPV, and optimal cutoff values were calculated and their clinical implications assessed. This method thus weighs the relative costs of the different predictions in the adherence diagnosis.

Even though sensitivity and specificity are considered the fundamental operating characteristics to assess the accuracy of a diagnostic test, in practice, their capacity for quantifying medical uncertainty is limited. Sensitivity and specificity do not help clinicians to estimate the probability of adherence in individual patients, especially when the patient comes from a population where the prevalence of adherence is different from that in the population used to develop the diagnostic test. Therefore, it is essential to know the probability that the diagnostic test result will give the correct patient diagnosis, regardless of the prevalence of the diagnosis. To evaluate this,

the data must be approached from the direction of the test results, using predictive values (29). Predictive values may be used to estimate probability of disease but both positive predictive value (PPV) and negative predictive value (NPV) vary according to disease prevalence. The PPV and NPV describe patients' probability of being truly adherent to their treatment once the results of his or her tests are known. The PPV of a test is defined as the proportion of patients with a positive test for treatment adherence who actually are adherent to risperidone as indicated in Eq. 6:

$$\mathrm{PPV} = (\mathrm{TP})/(\mathrm{TP} + \mathrm{FP}) \tag{6}$$

In addition, the NPV of a test is the proportion of patients with a negative test for treatment adherence who actually are non-adherent to risperidone as indicated in Eq. 7:

$$\mathrm{NPV} = (\mathrm{TN})/(\mathrm{TN} + \mathrm{FN}) \tag{7}$$

The ROC analysis described above was performed to evaluate if measuring a single *vs* multiple active moiety $C_{\mathrm{trough}}$ accurately predict risperidone treatment adherence and if it allows defining the best predictive cutoff value to discriminate between adherent and non-adherent patients with schizophrenia with the least number of misclassifications. For the assessment of the predictive performance of multiple active moiety $C_{\mathrm{trough}}$, a multivariate logistic regression model was built to evaluate if multiple (up to three) $C_{\mathrm{trough}}$ measured at different days at a steady state improve the prediction of being adherent or non-adherent to the treatment. The logistic regression model provides a link between multiple independent variables ($C_{\mathrm{trough}}$) and a binary dependent variable (treatment adherence *vs* treatment non-adherence), and allowed the calculation of probabilities of treatment adherence (p) according to the following equation:

$$\ln\left(\frac{p}{(1-p)}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \tag{8}$$

where $\alpha$ represents the intercept of the logistic regression model, $X$ corresponds to the $C_{\mathrm{trough}}$ measurement at the $i$th visit (for $i = 1$, 2, or 3), and $\beta_i$ represents the coefficient that quantifies the relationship between the $C_{\mathrm{trough}}$ at $i$th visit and the probability of treatment adherence, which ranges from 0 to 1. A $p$ value close to zero represents a high probability that the patient is considered non-adherent with the drug treatment, while a $p$ close to 1 will be a patient with a high probability to be adherent with the drug treatment. The $p$ obtained from the multivariate logistic regression model will be used as classifier of drug treatment adherence using the ROC methodology; therefore, the sensitivity, specificity, $\mathrm{AUC_{ROC}}$, PPV, and NPV as well as the optimal cutoff value were calculated.

Since the analysis was performed using $C_{\mathrm{trough}}$ as predictor of treatment adherence, a sensitivity analysis was performed to evaluate if the time of risperidone active moiety plasma measurement relative to risperidone dosing could potentially influence the accuracy, precision, and the optimal cutoff threshold of the diagnostic test across time. A total of five bins were generated after 24 h of receiving 4-mg risperidone dose at a steady state. The bins were established as 0–4, 5–9, 10–14, 15–19, and 20–24 h post-dose. For each bin, the average concentration ($C_{\mathrm{ave}}$) for the adherent and non-adherent population was calculated, as well as the $\mathrm{AUC_{ROC}}$ with its 95% CI, and the optimal cutoff threshold for each bin was calculated and compared across time.

The plasma concentration-time data were simulated using non-linear mixed effects modeling using NONMEM® version 7.1.0 or higher (ICON plc) (30). The Fortran compiler was Intel(R) Fortran 64 Compiler Professional, Version 11.1. The data management, exploratory analysis, diagnostic graphics, and the statistical analysis were carried out using R Project for Statistical Computing, Version 3.4.1 or higher for Windows (31).

## RESULTS

The simulated distribution of the active moiety $C_{\mathrm{trough}}$ at a steady state for both adherent and non-adherent patients is presented in Fig. 2a. The median (CV%) $C_{\mathrm{trough}}$ for the adherent cohort was 16.3 ng/mL (64.1%) while the median (CV%) $C_{\mathrm{trough}}$ for the non-adherent cohort was 9.5 ng/mL (81.7%). Figure 2c depicts the ROC curve representing the trade-off between sensitivity and specificity when a single $C_{\mathrm{trough}}$ is used as a predictor of treatment adherence. Figure 3a represents the sensitivity, specificity, and efficiency as a function of different $C_{\mathrm{trough}}$ cutoff values. The optimal $C_{\mathrm{trough}}$ threshold was estimated to be 11.9 ng/mL, which coincided with the intersection of the curves in Fig. 2a and provides the lowest number of misclassifications of treatment adherence and non-adherence. A summary of the diagnostic test parameters, $\mathrm{AUC_{ROC}}$, as well as sensitivity, specificity, PPV, and NPV, at the optimal cutoff, is presented in Table II.

Figure 2b shows the distribution of the predicted probability of adherence for both adherent and non-adherent patients after fitting a multivariate logistic regression model to multiple (in this case 3) simulated $C_{\mathrm{trough}}$ measurements at different visit days. The logistic regression model improved as more drug concentrations were included, the median (CV%) predicted probability of adherence for the adherent cohort was 0.79 (23.4%) while the median predicted probability of adherence (CV%) for the non-adherent cohort was 0.34 (76.3%). The intersection of the curves shows the threshold representing the optimal cutoff value where the lowest number of misclassifications is achieved when using multiple simulated $C_{\mathrm{trough}}$ measures at different visit days as predictors of drug treatment adherence.

The ROC curve representing the trade-off between sensitivity and specificity is shown in Fig. 2d. The sensitivity, specificity, and efficiency as a function of the predicted probability of adherence threshold, which ranges from 0 (non-adherent patient) to 1 (adherent patient), are depicted in Fig. 3b. The optimal predicted probability of adherence threshold, which accounts for the lowest number of misclassifications, was estimated to be 0.51 (patients scoring below this threshold will be declared non-adherent whereas patients scoring above 0.51 will be declared adherent to their treatment).

When additional concentration measurements are included as predictors of treatment adherence, the variability of

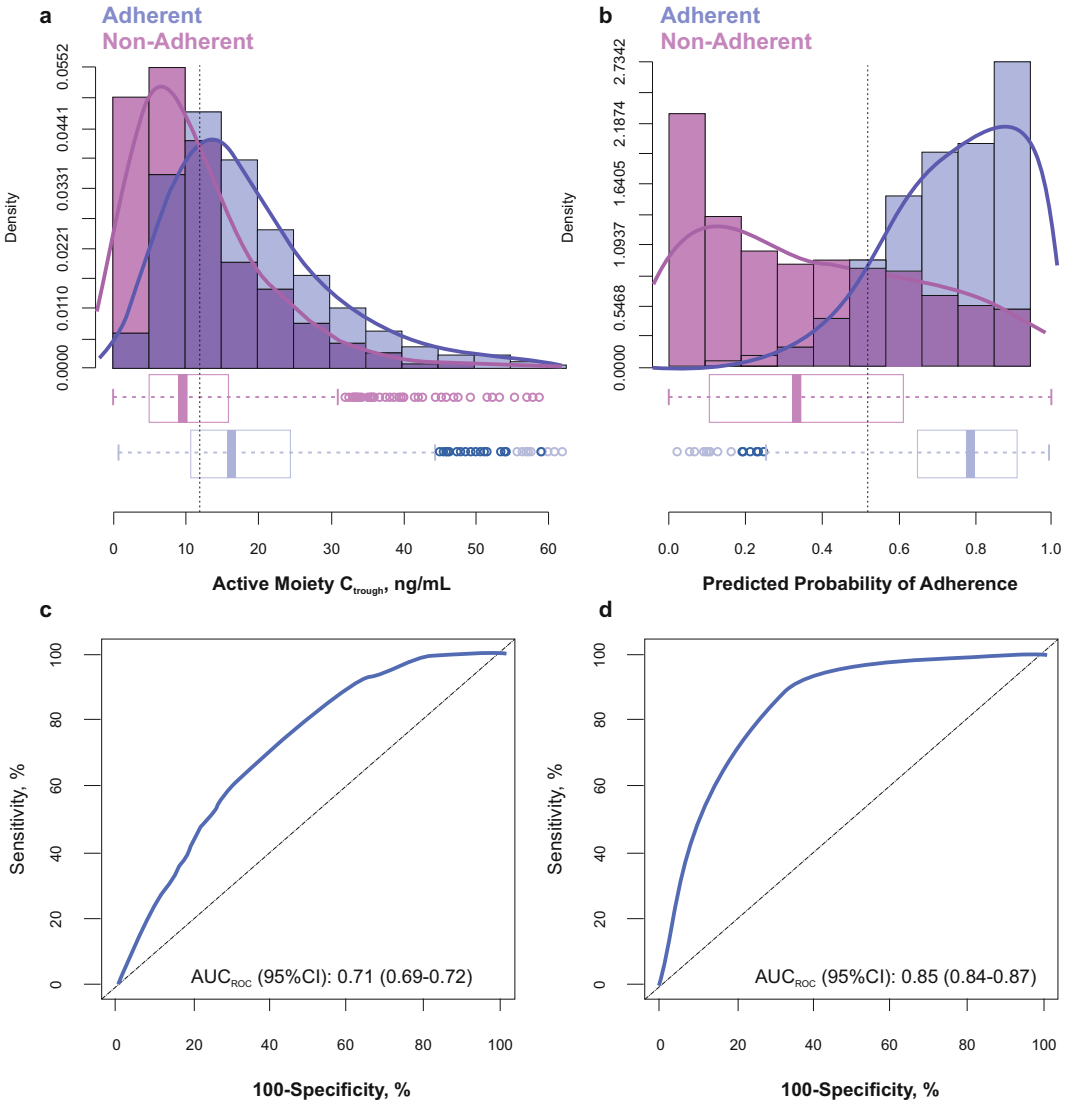


**Fig. 2.** The active moiety $C_{trough}$ (**a**) and the model predicted probability of adherence (**b**) for adherent and non-adherent patients. The dashed vertical line represents the most efficient cutoff value where the lowest number of misclassifications is achieved. ROC curves represent the trade-off between sensitivity and specificity when single (**c**) or multiple (**d**) drug concentration measurements are used as predictors of drug treatment adherence

the simulated distribution of the active moiety $C_{trough}$ at a steady state between adherent and non-adherent patients becomes lower, meaning that the intersection of the distribution curves becomes narrower. After the inclusion of two additional measurements, a reduction in the CV% from 64.1 to 23.4% in the adherent cohort was seen, while this reduction was from 81.7 to 76.3% in the non-adherent cohort. This reduction in the variability translated into a higher $AUC_{ROC}$ (95% CI), increasing from 0.71 (0.69–0.72) to 0.85 (0.84–0.87) from 1 to 3 $C_{trough}$ measurements at a steady state. The area under the $AUC_{ROC}$ results is considered excellent for $AUC_{ROC}$ values between 0.90 and 1.0, good for AUC values between 0.80 and 0.90, fair for AUC values between 0.70 and 0.80, poor for AUC values between 0.60 and 0.70, and failed for AUC values between 0.50 and 0.60 (32,33). For all the rest of the diagnostic test performance parameters (sensitivity, specificity, PPV, and NPV), the estimated probability for predictors of treatment adherence is shown to be higher for the multiple $C_{trough}$ than the single $C_{trough}$ measurements.

We estimated the sensitivity, specificity PPV, and NPV and found that, although the $AUC_{ROC}$ (95 CI%) provides a fair diagnostic test performance (0.71 [0.69–0.72]), the specificity (95% CI) or true negative rate was estimated to be 0.60 (0.58–0.63), meaning the percentage of non-adherent patients who are correctly identified as non-adherent is low (60%). However, after including two additional concentration measurements (patient follow-up), all the diagnostic test parameters (sensitivity, specificity, PPV, and NPV) improved as shown in Table II.

A summary of the diagnostic test parameters ($AUC_{ROC}$, sensitivity, specificity, PPV, and NPV) when the FPC is three times higher than the FNC is presented in Table II. Figure 3c shows the sensitivity, specificity, and efficiency for each $C_{trough}$, and predicted probability of adherence threshold assuming a three-fold higher FPC relative to FNC is shown in

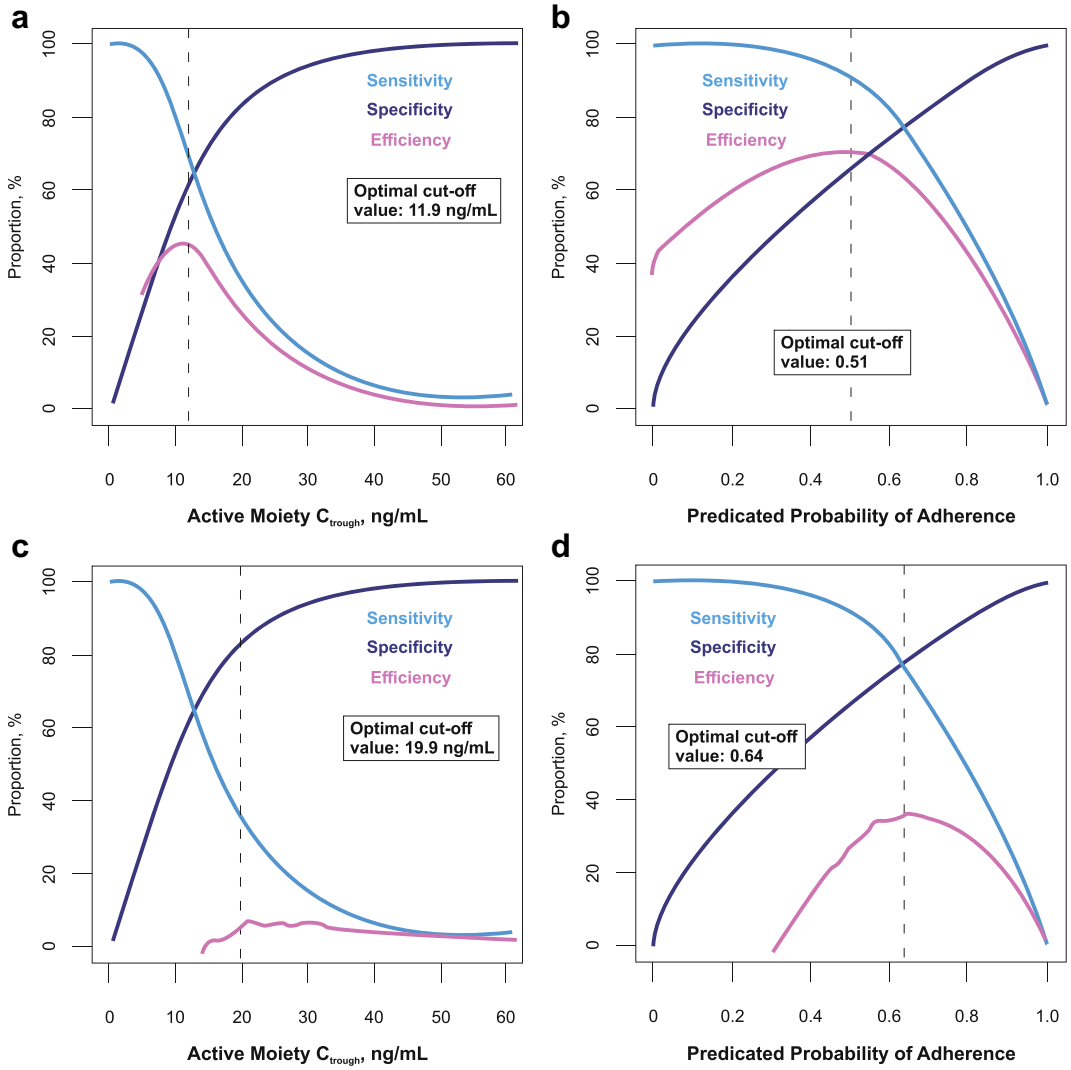


**Fig. 3.** Sensitivity, specificity, and efficiency for each $C_{trough}$ (**a**) and predicted probability of adherence (**b**) threshold assuming an equal false positive cost and false negative cost (**a**, **b**), and a three-fold higher cost to the false positive patients than to the false negative patients (**c**, **d**). The dashed vertical line represents the most efficient cutoff value where the lowest level of misclassification is achieved

Fig. 3d. The optimal predictive $C_{trough}$ threshold was estimated to be 19.9 ng/mL, whereas the optimal predicted probability of adherence threshold was 0.64.

The $AUC_{ROC}$ and the corresponding optimal cutoff values to detect the treatment adherence are shown across five different bins within the 24 h dosing interval after receiving 4-mg risperidone dose at a steady state (Fig. 4). A consistent $AUC_{ROC}$ was observed across the different bins evaluated, whereas the optimal cutoff threshold for each bin decreased according to the time course of risperidone's plasma concentrations (Fig. 4a). After normalizing drug concentrations in each bin to trough levels, the corresponding optimal cutoff values to detect treatment adherence become consistent across the different bins (Fig. 4b). The normalization to trough levels was done by dividing drug concentrations in adherent and non-adherent patients by the ratio of the average concentration in each time bin and the average $C_{trough}$ in the adherent population. Consequently, the optimal cutoff threshold identified in this analysis can be extended to different sampling time points by normalizing by the different ratios due to the time course of drug concentrations and also to different doses by applying the linearity principle (34). Finally, the relationship between PPV and NPV as a function of prevalence is shown in Fig. 5. The dashed vertical line which depicts the mean prevalence related to drug treatment adherence in individuals with schizophrenia is estimated to be approximately 60% (3).

## DISCUSSION

In this analysis, we propose the use of ROC analyses, multivariate logistic regression, and population PK model–based simulation techniques to evaluate plasma drug concentrations as a predictor of treatment adherence. While this is indeed the case, the principles as well as the specific methodology developed in this manuscript can be used for other compounds with different half-lives to obtain an early

**Table II.** Diagnostic Test Parameters for Drug Concentration Used as Predictors of Treatment Adherence

| Parameter | Single drug concentration estimate (95% CI) | Multiple ($n = 3$) drug concentration estimate (95% CI) |
|---|---|---|
| Equal false positive and false negative misclassification cost | | |
| $AUC_{ROC}$ | 0.71 (0.69–0.72) | 0.85 (0.84–0.87) |
| Sensitivity | 0.71 (0.69–0.73) | 0.92 (0.90–0.93) |
| Specificity | 0.60 (0.58–0.63) | 0.66 (0.63–0.68) |
| Positive predictive value | 0.74 (0.72–0.76) | 0.81 (0.79–0.83) |
| Negative predictive value | 0.56 (0.54–0.59) | 0.83 (0.81–0.85) |
| Optimal cutoff value | 11.9 ng/mL | 0.51 |
| False positive cost 3 times higher than the false negative cost | | |
| $AUC_{ROC}$ | 0.71 (0.69–0.72) | 0.85 (0.84–0.87) |
| Sensitivity | 0.37 (0.35–0.39) | 0.76 (0.74–0.78) |
| Specificity | 0.84 (0.82–0.86) | 0.79 (0.77–0.81) |
| Positive predictive value | 0.78 (0.76–0.80) | 0.85 (0.83–0.86) |
| Negative predictive value | 0.45 (0.43–0.49) | 0.67 (0.65–0.70) |
| Optimal cutoff value | 19.9 ng/mL | 0.64 |

*$AUC_{ROC}$* area under the receiver operating characteristic curve, *CI* confidence interval

assessment about the potential implications of a lack of adherence on clinical efficacy and/or safety outcomes. The simulation exercise would help clinicians to know the plasma concentration cutoff levels beyond which a patient is at risk to be non-adherent for other drugs.

Perfect adherence was assumed when defining the reference ranges (11). Due to naturally occurring variation, 10% of patients were expected to be below the lower limit of the 80% reference range for antipsychotic blood levels, despite perfect adherence. Recently, Melkote *et al.* found that 18% of patients from the evaluation dataset had exposures below the reference range antipsychotic blood levels, pointing to a lower level of adherence, and dropping below the reference range was significantly linked to treatment failure (35). In addition, the reference ranges have been validated, i.e., proven to be predictive for the PK of the drug. Only through the use of modeling and simulation, the PK for dosing scenarios that were not tested *in vivo* can be evaluated, and important conclusions derived. Drug forgiveness, conventionally defined as the ability to maintain a therapeutic effect despite missing one or more doses, is dependent on the pharmacokinetic/pharmacodynamic properties of the specific drug (36). Using the active moiety of risperidone with a longer duration of action, due to its extended half-life, i.e., more "forgiving" in our model, allows an objective prediction of the adherence threshold to be made more accurately.

The main advantage of using a multivariate logistic regression model is that additional patient information can be included into the model; such as patients' covariates or clinician assessment scales such as the Clinical Global Impression–Schizophrenia scale or Brief Psychiatric Rating Scale (37), which can measure negative symptom severity and track treatment response, allowing us to accurately discriminate between the adherent and non-adherent populations.

We evaluated the impact of having single *vs* multiple risperidone active moiety $C_{trough}$ to predict patient adherence. Furthermore, we assessed the relative benefit/cost of each type of decision when declaring a patient to be adherent or non-adherent to their treatment and selecting the optimal cutoff value that will lead to the lowest number of misclassifications.

The optimal predictive risperidone active moiety plasma concentration value for 4 mg QD risperidone that allows identification of the lack of treatment adherence was 11.9 ng/mL. This estimate coincides with the lower cut-point value of the previously developed reference ranges for risperidone which captured the plausible plasma concentration range in 80% of patients within a fully adherent population (11). These reference ranges were derived from the expected range of plasma concentrations for risperidone in a fully adherent patient population, taking formulation, dose, time after dose, and expected variability between the patients into account, and its aim was to allow clinicians to interpret individual concentration values more easily given adherence of the patient.

For single $C_{trough}$, the NPV was estimated to be 56%, meaning that in only 56% of the cases that a patient has a negative test for adherence (scores below 11.9 ng/mL), the patient will be a truly non-adherent patient. By including a total of three concentration measurements, the NPV increased up to 83%, which represents the probability that the clinician will be able to correctly identify a truly non-adherent patient. This indicates that multiple $C_{trough}$ provide greater accuracy than a single $C_{trough}$ measurement. On the other hand, we assessed the impact of increasing the FPC to three times higher with respect to the FNC on the diagnostic test parameters and the optimal cutoff value. For both single and multiple drug concentration measurements (Table II), when increasing the FPC, we are substantially increasing the specificity of the diagnostic test at the cost of reducing the sensitivity. In contrast, the NPV is decreasing, since we are giving more weight to the FPC, meaning that by increasing the FPC/FNC ratio, we will have a higher amount of FN with respect to the FP. In our study, we are interested in increasing the specificity of the diagnostic test and therefore this means that the optimal cutoff value will be increased from 11.9 to 19.9 ng/mL when using a single $C_{trough}$, and from 0.51 to 0.64 when using multiple $C_{trough}$ (Fig. 3).

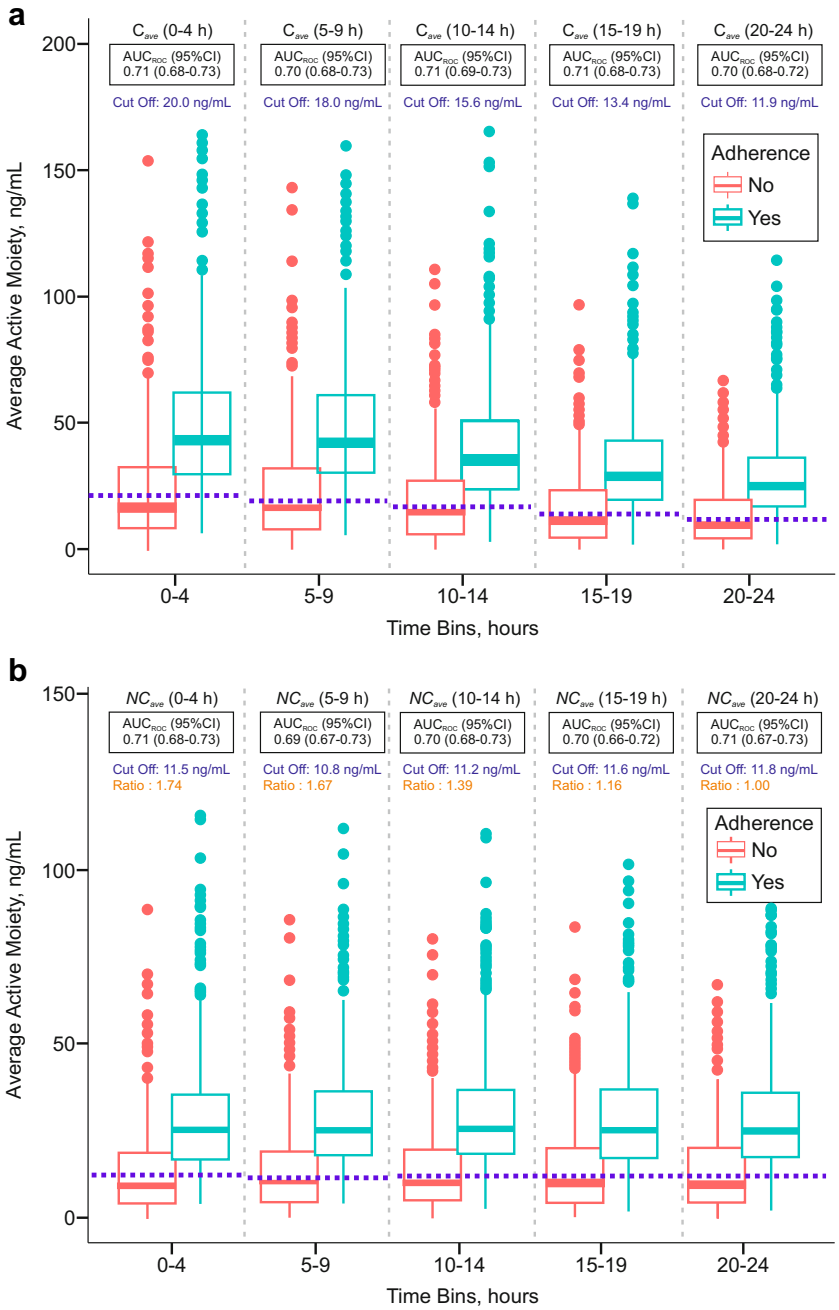


**Fig. 4.** AUC$_{ROC}$ and optimal cutoff concentration-time sensitivity analysis. Abbreviations: *AUC$_{ROC}$* area under the ROC curve, *$C_{ave}$* average concentration, *CI* confidence interval, *ROC* receiver operating characteristic

While sensitivity and specificity are intrinsic attributes of the test being evaluated (given similar patient and specimen characteristics) and are independent of the prevalence of disease in the population being tested, PPV and NPV are highly dependent on the population prevalence of the disease (29). This means that the same diagnostic test will have a different predictive accuracy according to the clinical setting in which it is applied. Similar as other diagnostic tests, if the test is applied to a population where the prevalence of adherence is low (e.g., 10%), the PPV will decrease as a consequence of an increase in the false positive rate as the diagnostic test is applied to a population which is mainly non-adherent. On the other hand, if the diagnostic test is conducted in a hospital setting where most of the patients are adherent to the medication, the test will increase PPV values. In contrast, the NPV will decrease since a higher number of adherent patients will be available in the population (Fig. 5). Due to this nature of the diagnostic test, it is important to know the prevalence of adherence of a given population since different adherence prevalence will raise different PPV and NPV values. The simulation presented in Fig. 5 shows the importance of the likelihood of the treatment adherence in the individual in whom the test has been performed. Thus, a test with good sensitivity and specificity may have a low PPV when it is used in a population where the likelihood of adherence is low. In view of this

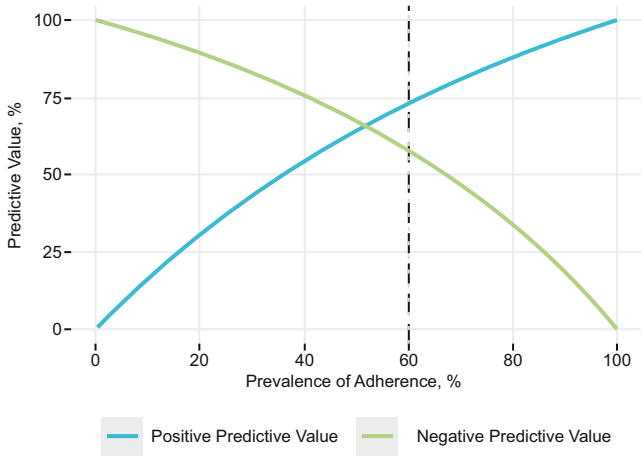


**Fig. 5.** Relationship between PPV and NPV as function of the prevalence of adherence. The mean prevalence related to treatment adherence was estimated to be approximately 60% as represented by the dashed vertical line

phenomenon, it is prudent to apply a diagnostic test only in those patients with a high probability of being non-adherent and this could be assessed by the corresponding symptoms and signs of non-adherence to treatment. In addition, clinicians should apply published predictive values of a test to their own populations with caution, especially when the prevalence of adherence in their population is different to the one reported by Lacro *et al.* (3), since this was the selected treatment adherence prevalence (60%) in our population and in which this analysis was carried out.

Finally, Fig. 4 shows a consistent $AUC_{ROC}$ across the different bins evaluated meaning that the same predictive performance of the diagnostic test will be achieved regardless of when the plasma concentration was measured post-dose. The optimal cutoff threshold for each bin decreased according to the risperidone plasma concentration-time profile; therefore, essential information for the clinician would be to know the time when the patient has taken his/her medication before applying the described diagnostic test.

Although the purpose of the present work was to provide an overall framework to conduct model-based assessments, a limitation of the presented methodology is that the compliance model was assumed to follow a Bernoulli distribution whereas implementing Markov-chains as described by Fellows *et al.* (38) would have allowed incorporation of different patient behaviors such as a 'patient cooperativity index' and may have been more predictive of real-world's patient behavior.

Plasma concentration measurements of antipsychotics should be considered as a complementary method to the current clinical practice for assessing treatment adherence, especially in the case of central nervous system diseases. Measurements with point-of-care devices would be of benefit as it is less invasive and provides rapid results, and the technology is available. Plasma concentrations alone are not sufficient; this methodology needs to be combined with an evaluation of the clinical status of the patient. Monitoring on more than one occasion, which additionally supports building a relationship between physician and patient, may help identify specific patterns of non-adherence. However, we do not propose this methodology to be a kind of therapeutic drug monitoring.

## CONCLUSIONS

The methodology proposed is appropriate to identify thresholds of drug concentration that enable an accurate and precise identification of non-adherent patients to drug treatment. The optimal predictive risperidone active moiety $C_{trough}$ value for 4-mg QD risperidone allowing identification of the lack of treatment adherence was 11.9 ng/mL which is consistent with the lower limit of the 80% reference range for adherent patients. Three consecutive steady-state $C_{trough}$ measurements are the minimum number of plasma concentrations needed to have an accurate and precise diagnostic test to properly discriminate between adherent and non-adherent patients, if the 40% of non-adherent patients are missing at least 20% of the dose intakes. Although an antipsychotic was used in this model for patients with schizophrenia, the analysis has to be expanded to more drugs than risperidone only. Along with a full clinical evaluation of the patient, this diagnostic test can assist the clinician to differentiate lack of efficacy from lack of adherence to antipsychotic treatment.

## ACKNOWLEDGEMENTS

Stacey E. Shehin, Ph.D. (PRA Health Sciences) provided medical writing assistance, which was funded by Janssen Research & Development. Harry Ma, Ph.D. (Janssen Global Services) provided additional editorial support. Portions of this study have been previously presented at the Population Approach Group of Europe's 26th Annual Scientific Meeting, June 6–9, 2017, Budapest, Hungary. All authors meet ICMJE criteria, had full access to the study data, and take responsibility for integrity of the data.

## FUNDING INFORMATION

This study was funded by Janssen Research & Development, LLC.

## COMPLIANCE WITH ETHICAL STANDARDS

**Conflict of interest** All authors are employees of Janssen Research & Development, LLC, and may hold stock options or shares in the company.

## REFERENCES

1. World Health Organization. Mental Health, Schizophrenia. 2017. Available from: http://www.who.int/mental_health/management/schizophrenia/en/. Accessed 13 Nov 2018.
2. Keating D, McWilliams S, Schneider I, Hynes C, Cousins G, Strawbridge J, et al. Pharmacological guidelines for schizophrenia: a systematic review and comparison of recommendations for the first episode. BMJ Open. 2017;7(1):e013881.
3. Lacro JP, Dunn LB, Dolder CR, Leckband SG, Jeste DV. Prevalence of and risk factors for medication nonadherence in patients with schizophrenia: a comprehensive review of recent literature. J Clin Psychiatry. 2002;63(10):892–909.

4. Acosta FJ, Hernandez JL, Pereira J, Herrera J, Rodriguez CJ. Medication adherence in schizophrenia. World J Psychiatry. 2012;2(5):74–82.
5. Novick D, Haro JM, Suarez D, Perez V, Dittmann RW, Haddad PM. Predictors and clinical consequences of non-adherence with antipsychotic medication in the outpatient treatment of schizophrenia. Psychiatry Res. 2010;176(2–3):109–13.
6. Higashi K, Medic G, Littlewood KJ, Diez T, Granstrom O, De Hert M. Medication adherence in schizophrenia: factors influencing adherence and consequences of nonadherence, a systematic literature review. Ther Adv Psychopharmacol. 2013;3(4):200–18.
7. Velligan DI, Weiden PJ, Sajatovic M, Scott J, Carpenter D, Ross R, et al. The expert consensus guideline series: adherence problems in patients with serious and persistent mental illness. J Clin Psychiatry. 2009;70(Suppl 4):1–46 quiz 7-8.
8. Karve S, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. Good and poor adherence: optimal cut-point for adherence measures using administrative claims data. Curr Med Res Opin. 2009;25(9):2303–10.
9. Velligan DI, Lam Y-WF, Glahn DC, Barrett JA, Maples NJ, Ereshefsky L, et al. Defining and assessing adherence to oral antipsychotics: a review of the literature. Schizophr Bull. 2006;32(4):724–42.
10. Horvitz-Lennon M, Mattke S, Predmore Z, Howes OD. The role of antipsychotic plasma levels in the treatment of schizophrenia. Am J Psychiatry. 2017;174(5):421–6.
11. Korell J, Green B, Remmerie B, Vermeulen A. Determination of plasma concentration reference ranges for risperidone and paliperidone. CPT Pharmacometrics Syst Pharmacol. 2017;6(9):589–95.
12. Suzuki Y, Kawasaki K, Sato Y, Tokimatsu I, Itoh H, Hiramatsu K, et al. Is peak concentration needed in therapeutic drug monitoring of vancomycin? A pharmacokinetic-pharmacodynamic analysis in patients with methicillin-resistant Staphylococcus aureus pneumonia. Chemotherapy. 2012;58(4):308–12.
13. Olivares-Morales A, Hatley OJ, Turner D, Galetin A, Aarons L, Rostami-Hodjegan A. The use of ROC analysis for the qualitative prediction of human oral bioavailability from animal data. Pharm Res. 2014;31(3):720–30.
14. Clements JD, Perez Ruixo JJ, Gibbs JP, Doshi S, Perez Ruixo C, Melhem M. Receiver operating characteristic analysis and clinical trial simulation to inform dose titration cecisions. CPT: Pharmacometrics Systems Pharmacol. 2018;7(11):771–9.
15. Hiligsmann M, Rabenda V, Gathon H-J, Ethgen O, Reginster J-Y. Potential clinical and economic impact of nonadherence with osteoporosis medications. Calcif Tissue Int. 2010;86(3):202–10.
16. Chang DS, Friedman DS, Frazier T, Plyler R, Boland MV. Development and validation of a predictive model for nonadherence with once-daily glaucoma medications. Ophthalmology. 2013;120(7):1396–402.
17. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. Clin Biochem Rev. 2008;29(Suppl 1):S83–7.
18. Park SH, Goo JM, Jo CH. Receiver operating characteristic (ROC) curve: practical review for radiologists. Korean J Radiol. 2004;5(1):11–8.
19. Vermeulen A, Piotrovsky V, Ludwig EA. Population pharmacokinetics of risperidone and 9-hydroxyrisperidone in patients with acute episodes associated with bipolar I disorder. J Pharmacokinet Pharmacodyn. 2007;34(2):183–206.
20. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem. 1993;39(4):561–77.
21. Metz CE. Basic principles of ROC analysis. Semin Nucl Med. 1978;8(4):283–98.
22. Swets JA. Indices of discrimination or diagnostic accuracy: their ROCs and implied models. Psychol Bull. 1986;99(1):100–17.
23. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29–36.
24. Boyd JC. Statistical analysis and presentation of data. In: Price CP, editor. Evidence-based laboratory medicine; principles, practice and outcomes. 2nd ed. Washington DC, USA: AACC Press; 2007. p. 113–40.
25. Wilson EB. Probable inference, the law of succession, and statistical inference. J Am Stat Assoc. 1927;22(158):209–12.
26. McNeil BJ, Keller E, Adelstein SJ. Primer on certain elements of medical decision making. N Engl J Med. 1975;293(5):211–5.
27. Lusted LB. Signal detectability and medical decision-making. Science. 1971;171(3977):1217–9.
28. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3(1):32–5.
29. Altman DG, Bland JM. Diagnostic tests 2: predictive values. BMJ. 1994;309(6947):102.
30. NONMEM. 7.1.0 Users Guides. Icon Development Solutions: Ellicott City, MD; 1989-2009.
31. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2012.
32. El Khouli RH, Macura KJ, Barker PB, Habba MR, Jacobs MA, Bluemke DA. Relationship of temporal resolution to diagnostic performance for dynamic contrast enhanced MRI of the breast. J Magn Reson Imaging. 2009;30(5):999–1004.
33. Ludemann L, Grieger W, Wurm R, Wust P, Zimmer C. Glioma assessment using quantitative blood volume maps generated by T1-weighted dynamic contrast-enhanced magnetic resonance imaging: a receiver operating characteristic study. Acta Radiol. 2006;47(3):303–10.
34. Mannaert E, Vermeulen A, Remmerie B, Bouhours P, Levron JC. Pharmacokinetic profile of long-acting injectable risperidone at steady-state: comparison with oral administration. Encéphale. 2005;31(5 Pt 1):609–15.
35. Melkote R, Singh A, Vermeulen A, Remmerie B, Savitz A. Relationship between antipsychotic blood levels and treatment failure during the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study. Schizophr Res. 2018;201:324–8.
36. Morrison A, Stauffer ME, Kaufman AS. Relationship between adherence rate threshold and drug 'forgiveness'. Clin Pharmacokinet. 2017;56(12):1435–40.
37. Kane JM. Tools to assess negative symptoms in schizophrenia. J Clin Psychiatry. 2013;74(6):e12-e.
38. Fellows K, Stoneking CJ, Ramanathan M. Bayesian population modeling of drug dosing adherence. J Pharmacokinet Pharmacodyn. 2015;42(5):515–25.