Title: Genetic determinants of the pharmacokinetic variability of rifampicin in Malawian adults with pulmonary tuberculosis

Authors: Derek J Sloan 1,2,3,4,5, Andrew D McCallum 1,3,5,6, Alessandro Schipani 5, Deirdre Egan 5, Henry C Mwandumba 1,3,4, Steve A Ward 3, David Waterhouse 3, Gertrude Banda 1, Theresa J Allain 4, Andrew Owen 6, Saye H Khoo 5,6, Geraint R Davies 5,6,7

Affiliations:
1 Malawi Liverpool Wellcome Trust Clinical Research Programme, College of Medicine, University of Malawi, Blantyre, Malawi
2 School of Medicine, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9TF
3 Liverpool School of Tropical Medicine, Liverpool, United Kingdom
4 Department of Medicine, College of Medicine, University of Medicine, Blantyre, Malawi
5 Royal Liverpool University Hospital, Liverpool, United Kingdom
6 Institute of Translational Medicine, University of Liverpool, Liverpool, United Kingdom
7 Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom

Corresponding author: Dr Derek J Sloan, School of Medicine, University of St Andrews, North Haugh, St Andrews, KY16 9TF, djs26@st-andrews.ac.uk

Alternative corresponding author: Dr Geraint R Davies, Institute of Infection and Global Health, The Ronald Ross Building, University of Liverpool, Liverpool, L69 7BE, gerrydavies@doctors.org.uk

Number of references: 38

Number of figures: 1

Number of tables: 1

Keywords: Tuberculosis, pharmacokinetics, single-nucleotide polymorphism, SLCO1B1, AADAC, CES-1
Variable exposure to anti-tuberculosis (TB) drugs, partially driven by genetic factors, may be associated with poor clinical outcomes. Previous studies have suggested an influence of the SLCO1B1 locus on the plasma area under the concentration-time curve (AUC) of rifampicin. We evaluated the contribution of Single Nucleotide Polymorphisms (SNPs) in SLCO1B1 and other candidate genes (AADAC, CES-1) to inter-individual pharmacokinetic variability in Malawi. 174 adults with pulmonary TB underwent sampling of plasma rifampicin concentrations at 2- and 6-hours post-dose. Data from a prior cohort of 47 intensively sampled, similar patients from the same setting were available to support population pharmacokinetic model development in NONMEM v7.2, using a two-stage strategy to improve information during the absorption phase. Contrary to recent studies in South Africa and Uganda, SNPs in SLCO1B1 did not explain variability in AUC0-∞ of rifampicin. No pharmacokinetic associations were identified with AADAC or CES-1 SNPs, which were rare in the Malawian population. Pharmacogenetic determinants of rifampicin exposure may vary between African populations. SLCO1B1 and other novel candidate genes, as well as non-genetic sources of inter-individual variability, should be further explored in geographically diverse, adequately powered cohorts.

Words: 186 (up to 250)
Although effective chemotherapy for tuberculosis (TB) has been available for several decades, cure rates are variable, remaining at 34–76% in many countries (1-6). The traditional view that treatment failure, relapse and emergence of antimicrobial resistance are predominantly driven by poor adherence has been challenged by reports that inter-individual variability in pharmacokinetic (PK) exposure to rifampicin (RIF) accounts for some unfavourable outcomes, even amongst patients who do not miss doses of medication (7).

Existing clinical data show up to 10-fold inter-individual variability in the plasma PK indices of key drugs, especially RIF (8-11). Maximum plasma concentration (Cmax) and area under the concentration-time curve (AUC) measurements are often reported as “low” suggesting that some patients are under-dosed according to currently clinical recommendations. Patient physiology, comorbidities, concomitant medications and dietary intake all influence drug exposure. However, genetic polymorphisms in drug metabolising enzymes and transporters may explain up to 30% of PK variability for all drugs (12-14). Although Africans have the highest degree of genetic diversity worldwide (15) and sub-Saharan Africa accounts for a large proportion of global TB incidence and mortality, data on the pharmacogenetic determinants of anti-tuberculosis drug exposure amongst TB-endemic African populations are sparse.

RIF is believed to be the decisive drug that enables short-course chemotherapy. Metabolism by hepatic esterases and biliary excretion of RIF occurs after first-pass metabolism and hepatocellular uptake, which may be primarily mediated by organic anion-transporting polypeptide 1B1 (OATP1B1), the product of the gene SLCO1B1 (16-17). Reports from South Africa and Uganda suggested that two SLCO1B1 Single Nucleotide Polymorphisms (SNPs), rs11045819, rs4149032, are common in African patients and associated with reduced RIF plasma exposure (18-20). However, a recent study from a
TB-endemic population in southern India reported lower variant allele frequency of these SNPs than the prior African work and no genotypic effect on the pharmacokinetics of RIF (21). More data from a broader range of global settings are required to assess the effect of SLCO1B1 genetics on rifampicin exposure.

Within the hepatocyte, RIF is thought to be metabolised by microsomal hepatic esterases, which have been incompletely characterised to date. In vitro data have suggested that the serine esterase arylacetamide deacetylase (AADAC) may mediate 25-deacetylation of rifamycins (22) and that SNPs in the genes encoding hepatic microsomal AADAC may alter RIF clearance (23). Alternative routes to breaking ester linkages in RIF metabolism may involve the carboxyesterase (e.g. CES-1 and -2) enzymes which are relatively abundant in the liver and gut. There are no clinical data to define the effect of SNPs in AADAC or CES genes on plasma RIF exposure.

To assess previously described pharmacogenetic effects in a new population, and evaluate the contribution of unexplored polymorphisms in other key metabolic processes, we assessed the impact of critical SNPs in the candidate genes SLCO1B1, AADAC and CES-1 on plasma exposure to RIF amongst Malawian adults with smear-positive pulmonary TB. To maximise our study size with limited resources in a low income setting, participants underwent sparse PK sampling at two time points. Prior intensive PK data from the same population facilitated use of two-stage population pharmacokinetic methods to improve information on RIF exposure, particularly during the absorption phase.

**Results**

The sparsely sampled dataset comprised 174 participants. 121 (69.5%) were male, median age was 30 (range: 17-61) years and median weight was 52.0 (range: 34-74) kg. 98 (56.3%) were HIV-infected...
with a median CD4 count of 174 (range: 6-783) cells/µl. 28/98 (28.6%) HIV-infected patients were on Antiretroviral Therapy (ART) at the time of recruitment. Full details are provided in Table 1.

The intensely sampled dataset used for the first stage of pharmacokinetic model-fitting comprised 47 participants. 24 (52%) were male, median age was 34 (range: 16-60) years and median weight was 52.5 (range: 35.8-74.3) kg. 30 (65%) were HIV-infected and 13/47 (27.7%) were on ART at the time of recruitment. All participants from both cohorts were black Africans newly diagnosed with tuberculosis at Queen Elizabeth Central Hospital (QECH), Blantyre, Malawi.

In the first stage of the pharmacokinetic analysis, performed with the intensively sampled dataset alone, a one-compartment model appeared most appropriate with a transit compartment model best describing the absorption phase (Δ Objective Function Value (OFV) = 111). Inter-individual random effects (IIV) were supported for apparent clearance (CL/F) apparent volume of distribution (V/F) and absorption mean transit time (MMT).

The second stage of the analysis utilised only the sparsely-sampled dataset. Due to the relative lack of information on the absorption phase from this dataset, model parameters relating to absorption were fixed at the values estimated in the first stage of the analysis. Model refinement involved a comprehensive search of the covariates available. The inclusion of an allometric scaling weight model decreased the OFV significantly (ΔOFV = 7). The only demographic covariate which significantly affected the model fit was sex, with a drop of OF of 10.5 (P<0.01). The final model showed an increase of clearance of 17% in male patients, resulting in a decreased of average AUC of 3% in males. Full PK parameters of the final two stage RIF model are included in Tables 2 and 3.

A Visual Predictive Check (VPC) of 1000 simulated datasets indicated that the final model performed adequately (Figure 1). The values of RIF AUC$_{0-\infty}$ and C$_{Max}$ were obtained using empirical Bayesian
estimates of the individual parameters. A predicted RIF AUC$_{0-\infty}$ was generated for each subject using their estimated CL/F from the final model ($\text{Dose}/[\text{CL/F}] = \text{AUC}$). The median predicted AUC$_{0-\infty}$ was 29.9 $\mu$g.hr/ml, with a range of 19.7 $\mu$g.hr/ml to 63.4 $\mu$g.hr/ml. The median CMax was 4.8 $\mu$g/ml with a range of 1.4 $\mu$g/ml to 10.9 $\mu$g/ml.

Candidate SNPS in the SLCO1B1, AADAC and CES-1 loci were assessed. The distribution of alleles at these loci is presented in Table 1. The rs11045819 SLCO1B1 SNP was rare, existing at an overall frequency of 0.07 whilst the rs4149032 SLCO1B1 SNP had a minor allele frequency of 0.32. No subjects possessed the variant rs61733692 AADAC SNP whilst the rs1803155 AADAC SNP had a minor allele frequency of 0.25. Only one patient was heterozygote for the variant rs121493368 CES-1 SNP. Since the metabolic action of these genes is believed to be exerted largely during first-pass metabolism, their impact on relative bioavailability (F) was tested as well as CL/F. Additive, dominant and recessive models of effect were tested for each SNP. Inclusion of SLCO1B1 genotypes did not significantly improve the model fit. The AADAC SNPs and haplotype, or the CES-1 SNP did not significantly alter the RIF model fit; as would be anticipated from the very low variant allele frequency in our study population.

Discussion

This study is the first pharmacogenetic analysis of anti-tuberculosis therapy from Malawi, the largest such study to date using a population modelling approach to relate pharmacogenetic polymorphisms to plasma AUC and the first to evaluate the effect of SNPs in AADAC in this context. The results of our analysis draw attention the importance of: fully considering SNP diversity within Africa; the complexity of competing metabolic processes; and recognising the limitations of observational pharmacogenetic designs.
A two-stage population PK modelling approach was used to quantify RIF exposure amongst adult patients on treatment for pulmonary TB, benefiting from the availability of data from a pre-existing, exchangeable population of intensively sampled patients (24). A sequential rather than simultaneous approach was used to handle all the data because pharmacogenetic information was unavailable for the intensive dataset, reducing the value of combined modelling with incomplete covariates.

The two-step approach enabled plausible and efficient estimation of plasma AUC from sparsely sampled participants, which is a more reliable measure of total drug exposure than a single “peak” concentration at a fixed time after dosing. Plasma AUC demonstrated up to 6-fold variability between patients, like previous African cohorts. RIF is a crucial component of the current first-line anti-tuberculosis regimen. In general, AUC/MIC (minimum inhibitory concentration) is the PK index most closely related to the bactericidal activity of RIF (25), and a recent South African study reported a threshold of AUC of RIF (≤ 13 µg.hr/ml) to be an independent predictor of poor outcome (26). Therefore, identifying host pharmacogenetic factors which influence the AUC of this key drug may help identify individuals at risk of treatment failure, and explain differences in treatment outcomes between populations.

The SLCO1B1 locus encodes an organic anion transporter (OATP1B1) implicated in the hepatic uptake of several drugs and contains at least 17 non-synonymous SNPs. Much interest in Europe has focused on the role of rs4149056 (521T>C, *5) in toxicity of statin therapy (27), whilst studies in South Africa and Uganda have implicated two SLCO1B1 SNPs (rs4149032 and rs11045819) in reduced RIF exposure (18-20). However, in keeping with a recent report from southern India (21), inclusion of these genotypes did not significantly improve the PK model fit for RIF in our study.

There are potential explanations for the discrepant results between SLCO1B1 studies. The rs4149032 variant allele frequency was higher in South Africa (0.70-0.76), where an association with RIF was reported, than our Malawian study (0.32) or southern India (0.46) where the results were negative.
Similarly, the rs11045819 variant allele frequency was higher in Uganda (0.15) than Malawi (0.07) or India (0.01). Therefore, it is possible that populations with lower allele frequencies require a larger sample size for adequate power to detect a pharmacokinetic effect. Additionally, whilst rs11045819 is believed to be functional (463C>A, *4), rs4149032 is an intronic SNP of unclear functional status which, in European populations, is in strong linkage disequilibrium with rs11045819 (28). However, these SNPs had a low level of linkage disequilibrium (D’ 0.16) in Malawi, echoing findings from South Africa (18) and southern India (21). There are large differences in linkage disequilibrium worldwide and within Africa (15), so it is possible that previously recognised SNPs may tag functional genes in some populations but not others.

AADAC, a microsomal serine deacetylase expressed mainly in the liver and gastrointestinal tract, is responsible for 25-deacetylation of rifamycins in vitro (23). Previous pre-clinical studies identified that expression of the AADAC*3 allele (rs1803155/rs61733692) significantly reduced RIF clearance (22). No subjects in our study possessed the AADAC*3 allele, and inclusion of AADAC haplotype did not significantly improve the RIF model fit. Whilst the low incidence of these AADAC SNPs excluded their role in explaining inter-individual PK variability in this Malawian population, these genes are understudied. Similarly, the low incidence of one CES-1 SNP in our population does not exclude the possibility that polymorphisms in this gene are relevant in other populations, or that other mutations are important. Future work may describe higher SNP incidence in other settings, or identify alternative SNPs of relevance to drug exposure.

The gene loci assessed here focussed on hepatic drug uptake and metabolism. Additional processes, including widely distributed mucosal P-glycoprotein transporters (29) and orphan nuclear receptor regulatory elements which influence gene induction (30) may also influence RIF metabolism. Investigation of these broader pharmacogenetic factors may be studied separately.
These data indicate that inter-individual variability in RIF PK amongst Malawian adults was not explained by polymorphisms in the candidate genes studied. They illustrate the importance of local ethnic background in the context of high SNP diversity within Africa and emphasise the need for caution in extrapolating findings across the continent. Growing recognition of the extent and clinical consequences of variable RIF PK (31-37) coupled with recent evidence that dose escalation is safe and tolerable (38) highlights the need for ongoing work to better define the relative importance of polymorphisms and non-genetic risk factors for low antibiotic exposure in a range of populations.

There were several limitations to our study. Standard Fixed Dose Combination (FDC) tablets were used in treatment of the participants under field conditions, but we have not been able to account for the effect of quality of the potentially varying drug formulation on PK variability between the two cohorts. This reflects the reality of the TB drug supply chain in most countries. The use of sparse PK sampling at three fixed time points over six hours limited the precision of the estimates in each model, particularly for the absorption parameters. The use of data from a similar, intensively sampled cohort, to develop a population PK model mitigated this to some extent, achieving reasonable precision for the key PK parameter, AUC. Finally, the size of the dataset was not predicated on this secondary pharmacogenetic analysis and the negative findings could be due to a relative lack of power. However, our study was larger than any reported African cohort to date and there are limitations on the number of PK profiles, even based on sparse sampling protocols, that can practically be obtained in the field situation. On the other hand, it is widely recognised that candidate gene studies can be prone to chance findings which may not be subsequently replicated.

In conclusion, high inter-individual variability in plasma exposure to RIF amongst Malawian adults with pulmonary TB cannot be explained by genetic heterogeneity in SLCO1B1, as suggested from other African populations. Similarly, the variability cannot be explained by the novel candidate gene SNPs in AADAC, CES-1 that were evaluated but the low frequency of variant genes in our population
does not exclude an association elsewhere. The true significance of pharmacogenetic influences on
disposition of anti-tuberculosis drugs by comparison with non-genetic factors may only be
established through consistent observation across geographically diverse and adequately powered
cohort studies.

**Ethics**

Ethical approval was obtained from the College of Medicine Research Ethics Committee, University
of Malawi, and the Research Ethics Committee of Liverpool School of Tropical Medicine.

**Materials and methods**

**Study participants**

The sparse pharmacokinetic data sampling was conducted within a prospective cohort study
carried out in Blantyre, Malawi from 2010-2012. Consenting adults aged 16-65 years with
sputum smear-positive pulmonary TB were eligible. Exclusion criteria included haemoglobin<6g/dL,
creatinine>177μmol/l, total bilirubin>51μmol/l, alanine transaminase >200IU/l, clinical status
suggestive of imminent mortality (WHO Performance Score 4 (39)), pregnancy, TB treatment within
five years, corticosteroid therapy or baseline resistance to rifampicin and isoniazid using the
Genotype MTBDRplus 2.0 line probe assay (LPA, Hain Life Sciences). Patient characteristics have
been reported previously (40). Participants received daily FDC tablets according to a WHO-approved
weight-adjusted regimen (including rifampicin 8-12mg/kg) and standard National Tuberculosis
Programme guidelines (41). Adherence was monitored by direct questioning and pill counts. All
patients had point of care HIV serology. ART was provided per national protocols.

The intensively sampled pharmacokinetic data used for the first stage of model construction was
obtained from a prior study of adult patients with sputum smear-positive pulmonary TB at the same
hospital, conducted from 2007-2008. Recruitment criteria, patient characteristics and treatment protocols have also been reported previously (24), and were similar to the sparse sampling dataset.

Genotyping

Genomic DNA extraction and genotyping were performed from whole blood samples collected at baseline from each patient, as previously described (18). TaqMan real-time PCR using fluorescent probes for allelic discrimination was used to detect SNPs of two RIF transporting or metabolising genes: SLCO1B1 (rs11045819, rs4149032) and AADAC (rs1803155, rs61733693). These SNPs are previously reported to have functional significance for RIF PK (18,22,23). An exploratory SNP (rs12149368) in CES-1 was also assessed.

Drug plasma concentration determination

For the sparse sampling cohort, blood collections to measure steady-state RIF concentrations were undertaken on day 14 or 21 after TB treatment initiation. Patients attended the study clinic after an overnight fast at 0730 hours. Samples were collected pre-dose, then 2 and 6 hours after medications. Plasma was separated by centrifugation and stored at -70°C until analysis.

RIF concentrations were determined using a liquid chromatographic/tandem mass spectrometry method (24) using appropriate internal standards validated to internationally recognised acceptance criteria as previously described (42). The lower limit of quantification (LLQ) was 0.5 µg/ml for RIF. 5% of samples had RIF results below the LLQ. These data-points were handled by imputing a value which was 50% of the LLQ.

Population pharmacokinetic analysis
A population PK model for RIF was developed using NONMEM® (version 7.2.0, ICON Development Solutions). Since the sparse sampling data contained little information on RIF absorption, a two-stage model-building strategy was used.

In the first stage, a dataset previously obtained from 47 adult tuberculosis patients in Blantyre employing intensive sampling and the same drug assay (24) was used to characterise the absorption phase. No covariates were included. One- and two-compartment models with alternative models of absorption were fitted to the data using the First Order Conditional method of estimation with interaction. Among the models explored were simple first-order absorption or a sequence of zero- and first-order absorption incorporating either lag times or transit compartment absorption. Proportional, additive and combined proportional and additive error models were considered to describe residual variability. The minimal OFV (equal to -2 log likelihood) was used as a goodness-of-fit metric with a decrease of 3.84 corresponding to a statistically significant difference between models (P=0.05, χ2 distribution, one degree of freedom). Residual plots were also examined.

Once the appropriate structural model was established, the values of the absorption parameters were fixed and a second-stage analysis performed using data from the sparsely-sampled participants of the current study. The following covariates were explored: body weight, age, gender, HIV status and SNP genotypes.

Exponential errors following a log-normal distribution were assumed for the description of inter-individual variability in pharmacokinetic parameters, as shown in the following equation:

\[ \theta_i = \theta_i \cdot \exp(\eta_i) \]

where \( \theta_i \) is pharmacokinetic parameter “x” of the \( i \)th individual; \( \theta_i \) is the population parameter estimate; and \( \eta_i \) is the log inter-individual variability for parameter “x” drawn from a normal distribution with a mean of zero and variance \( \omega^2 \). RIF was administered orally, CL and V represent...
apparent values (CL/F and V/F, respectively, where F is the oral bioavailability, which was fixed to 1).

An allometric weight model was applied to standardize the pharmacokinetic parameters using a standard weight (WTstd) of 70 kg. An allometric weight model for clearance parameters is given by

\[ \text{CL/F}_{\text{wt}} = \left( \frac{\text{WT}}{\text{WTstd}} \right)^{3/4} \]

and for volume parameters is given \[ \text{V/F}_{\text{wt}} = \left( \frac{\text{WT}}{\text{WTstd}} \right)^{1} \]

where CL/Fwt and Vwt are the weight functions for clearance parameters and volume of distribution parameters, respectively, and WT is the individual weight value.

Dichotomous covariates were introduced as a power model and continuous variables were modelled using a power model with normalized covariate:

\[ \theta_i = \theta_1 \times \theta_{\text{cov}}^X \]

\[ \theta_i = \theta_1 \times (\text{COV}_i / \text{COV}_{\text{median}})^{\theta_{\text{cov}}} \]

where \( \theta \) is the pharmacokinetic parameter of the \( i \)th individual; \( \theta_1 \) is the population parameter estimate; in equation 2 (dichotomous covariates) \( \theta_{\text{cov}} \) is the ratio value of \( \theta \) for the individuals \( X = 0 \) or \( 1 \). In equation 3 (continuous covariates) COV is the value of the covariate for the \( i \)th individual, \( \text{COV}_{\text{median}} \) is the median value, \( \theta_{\text{cov}} \) is the fraction of \( \theta \) relative to the consider covariate effect.

Genotype information was coded as an index variable, shown below for CL/F:

\[ TVCL = \theta_0 + \theta_1 \times X_1 + \theta_2 \times X_2 \]

where \( \theta_0 \) is the typical value of CL/F for individuals with homozygosity, \( \theta_1 \) is the relative difference in CL/F for heterozygous for the mutant allele when \( X_1 = 1 \), and \( \theta_2 \) is the relative difference in CL/F for patients homozygous for the mutant allele when \( X_2 = 1 \).

Graphical methods were used to explore the relationship of covariates versus individual predicted pharmacokinetic parameters. Each covariate was introduced separately into the model and only
retained if inclusion in the model produced a statistically significant decrease in OFV of 3.84 (P≤0.05). A backwards elimination step was then carried out once all relevant covariates were incorporated and covariates were retained if their removal from the model produced a significant increase in OFV (>6.63 points; P≤0.01, χ2 distribution, one degree of freedom).

To perform a VPC using Perl-speaks-NONMEM (PsN), 1000 datasets were simulated using the parameter estimates defined by the final model with the SIMULATION SUBPROBLEMS option of NONMEM7. From the simulated data, 90% prediction intervals (P5–P95) were constructed. Observed data from the original dataset were superimposed for both regimens. PsN was used to run a nonparametric bootstrap of 200 iterations to provide unbiased estimates of the standard errors and the 95% confidence intervals of the estimated parameters.

Estimates of AUC<sub>0-∞</sub> for RIF were calculated from simulated values of CL/F using the equation:

\[
AUC_{0-\infty} = \frac{\text{Dose in milligrams}}{(CL/F)}
\]

Acknowledgements: D.J.S. was supported by a Wellcome Trust Clinical PhD Fellowship (086757/Z/08/A to D.J.S.). A.D.M. was supported by a National Institute for Health Research Integrated Clinical Academic Training Fellowship and a Wellcome Trust Clinical PhD Fellowship (105/392/B/14/Z). The Malawi Liverpool Wellcome Trust Clinical Research Programme is supported by a strategic award from the Wellcome Trust. We also acknowledge infrastructural support for bioanalysis from the Liverpool Biomedical Research Centre funded by Liverpool Health Partners. Lisa Stone was integral to some of the laboratory work described in this manuscript but sadly passed away before it was completed; we would like to gratefully acknowledge her contribution. We thank the patients who participated in this study, and their families.

Conflicts of interest/disclosures: All authors: No potential conflicts of interest.
Author contributions: D.J.S., A.D.M. and G.R.D wrote the article; D.J.S., H.C.M., S.H.K. and G.R.D.
performed the research, D.J.S., A.D.M., A.S. and G.R.D. analysed the data.
References


Pharmacol. 72(1), 51-62.

US. 2008. Population pharmacokinetics of rifampin in pulmonary tuberculosis patients, including a 
semimechanistic model to describe variable absorption. Antimicrob. Agents Chemother. 52(6), 2138- 
48.

12. Evans WE, McLeod HL. 2003. Pharmacogenomics - drug disposition, drug targets, and side 


and pharmacogenetic issues in tuberculosis drug development and clinical research. J. Infect. Dis. 
211(Suppl 3), S115-25.

324(5930): 1035-44.


Figure legends

Figure 1: Visual predictive check for the final rifampicin model. The lower, middle and upper lines are the 5\textsuperscript{th} percentile, median and 95\textsuperscript{th} percentile of the observed data, respectively. The shaded areas are the 95\% CIs for the 5\textsuperscript{th} percentile, median and 95\textsuperscript{th} percentile of the simulated data.
Table 1: Demographic and clinical description of cohort

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (n=174)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, median (range)</td>
<td>30 (17-61)</td>
</tr>
<tr>
<td>Male sex, n (%)</td>
<td>121 (69.5)</td>
</tr>
<tr>
<td>HIV infected, n (%)</td>
<td>98 (56.3)</td>
</tr>
<tr>
<td>CD4 in HIV infected in cells/μl, median (range)</td>
<td>174 (6-783)</td>
</tr>
<tr>
<td>On ART at baseline if HIV infected, n (%)</td>
<td>28 (28.6)</td>
</tr>
<tr>
<td>Weight in kg, median (range)</td>
<td>52 (34-74)</td>
</tr>
<tr>
<td>Adherence, n (%)*</td>
<td></td>
</tr>
<tr>
<td>Missed no doses</td>
<td>143 (82.2)</td>
</tr>
<tr>
<td>Missed &lt;2 doses</td>
<td>10 (5.7)</td>
</tr>
<tr>
<td>Missed &gt;2 doses</td>
<td>2 (1.1)</td>
</tr>
<tr>
<td>Rifampicin / isoniazid dose (mg), n (%)</td>
<td></td>
</tr>
<tr>
<td>300 / 150</td>
<td>2 (1.1)</td>
</tr>
<tr>
<td>450 / 225</td>
<td>113 (64.9)</td>
</tr>
<tr>
<td>600 / 300</td>
<td>59 (33.9)</td>
</tr>
<tr>
<td>SLCO1B1 mutations</td>
<td></td>
</tr>
<tr>
<td>rs11045819</td>
<td></td>
</tr>
<tr>
<td>wild (CC)</td>
<td>150 (86.2)</td>
</tr>
<tr>
<td>heterozygote (AC)</td>
<td>24 (13.8)</td>
</tr>
<tr>
<td>variant (AA)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>rs4149032</td>
<td></td>
</tr>
<tr>
<td>wild (TT)</td>
<td>89 (51.1)</td>
</tr>
<tr>
<td>heterozygote (CT)</td>
<td>59 (33.9)</td>
</tr>
<tr>
<td>variant (CC)</td>
<td>26 (14.9)</td>
</tr>
<tr>
<td>AADAC mutations</td>
<td></td>
</tr>
<tr>
<td>rs1803155</td>
<td></td>
</tr>
<tr>
<td>wild (GG)</td>
<td>3 (1.7)</td>
</tr>
<tr>
<td>heterozygote (CG)</td>
<td>80 (46.0)</td>
</tr>
<tr>
<td>variant (CC)</td>
<td>91 (52.3)</td>
</tr>
<tr>
<td>rs61733692</td>
<td></td>
</tr>
<tr>
<td>wild (TT)</td>
<td>174 (100)</td>
</tr>
<tr>
<td>heterozygote (CT)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>variant (CC)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>AADAC haplotype</td>
<td></td>
</tr>
<tr>
<td>AADAC*1/*1</td>
<td>3 (1.7)</td>
</tr>
<tr>
<td>AADAC*1/*2</td>
<td>80 (46.0)</td>
</tr>
<tr>
<td>AADAC*2/*2</td>
<td>91 (52.3)</td>
</tr>
<tr>
<td>CES1 mutations</td>
<td></td>
</tr>
<tr>
<td>rs12149368</td>
<td></td>
</tr>
<tr>
<td>wild (GG)</td>
<td>173 (99.4)</td>
</tr>
<tr>
<td>heterozygote (CG)</td>
<td>1 (0.6)</td>
</tr>
<tr>
<td>variant (CC)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

ART, anti-retroviral therapy; SNP, single nucleotide polymorphism; NAT2, N-acetyltransferase 2; SLCO1B1, solute carrier organic anion transporter family member 1B1; AADAC, arylacetamide deacetylase; CES1, carboxylesterase 1.

*Adherence data only available for 155 patients.
Table 2: parameter value estimates for the base model stage 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Typical value</th>
<th>%RSE</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL/F (l/h)</td>
<td>19.5</td>
<td>3.9</td>
<td>18.2-20.8</td>
</tr>
<tr>
<td>V/F (l)</td>
<td>27.1</td>
<td>13.9</td>
<td>20.9-33.3</td>
</tr>
<tr>
<td>(K_{a}) (h(^{-1}))</td>
<td>0.277</td>
<td>9.8</td>
<td>0.23-0.32</td>
</tr>
<tr>
<td>NN</td>
<td>1.5</td>
<td>41</td>
<td>0.09-3.9</td>
</tr>
<tr>
<td>MMT (h)</td>
<td>0.326</td>
<td>35</td>
<td>0.05-1.0</td>
</tr>
</tbody>
</table>

Random effects

<table>
<thead>
<tr>
<th></th>
<th>IIVCL</th>
<th>IIVV</th>
<th>IIVMMT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.165</td>
<td>0.425</td>
<td>0.0706</td>
</tr>
<tr>
<td></td>
<td>15.3</td>
<td>27.1</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>0.12-0.20</td>
<td>0.25-0.59</td>
<td>0.01-1.79</td>
</tr>
</tbody>
</table>

Residual variability:

| Proportional error (%) | 0.23 | 7.8  | 0.20-0.26 |

CL/F, clearance (\(F\) is unknown bioavailability); V/F apparent volume of distribution; \(K_{a}\) absorption rate constant; NN, number of transit compartment; MMT, absorption mean transit time; IIVCL, inter individual variability on clearance; IIVV, inter individual variability on volume; IIVMMT, inter individual variability on absorption mean transit time. a RSE, relative standard error reported. b CI, confidence interval.
Table 3: parameter value estimates for the final model stage 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Typical value (Shrinkage %)</th>
<th>%RSE a</th>
<th>95% CI b</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL/F (l/h)</td>
<td>19.6</td>
<td>12</td>
<td>16.7-22.5</td>
</tr>
<tr>
<td>V/F (l)</td>
<td>23.6</td>
<td>9</td>
<td>17.1-30.1</td>
</tr>
<tr>
<td>Kc (h⁻¹)</td>
<td>0.277 (FIX)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NN</td>
<td>1.5 (FIX)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMT (h)</td>
<td>0.326 (FIX)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>THETA sex male</td>
<td>1.2</td>
<td>13</td>
<td>1.0-1.3</td>
</tr>
</tbody>
</table>

Random effects

| IIVCL              | 0.076 (22)                  | 29     | 0.033-0.11|
| IIVV               | 0.397 (28)                  | 29     | 0.17-0.63 |
| IIVMMT             | 0.0706 (FIX)                |        |          |

Residual variability:

| Proportional error (%) | 0.22 | 12 | 0.19-0.26 |

CL/F, clearance (F is unknown bioavailability); V/F apparent volume of distribution; NN, number of transit compartment; MMT, absorption mean transit time; THETA sex, fractional change in clearance for males; IIVCL, inter individual variability on clearance; IIVV, inter individual variability on volume; IIVMMT, inter individual variability on absorption mean transit time. * RSE, relative standard error reported. ** CI, confidence interval.