Original article

Improving the prediction of virological response to tipranavir: the development and validation of a tipranavir-weighted mutation score

*Jonathan M Schapiro[1], Joseph Scherer[2], Charles A Boucher[3], John D Baxter[4], Clemens Tilke[5], Carlo F Perno[6], Franco Maggiolo[7], Maria M Santoro[6], David B Hall[2]**

[1]National Hemophilia Center, Tel Hashomer, Israel
[2]Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT, USA
[3]Erasmus Medical Center, Rotterdam, the Netherlands
[4]Cooper University Hospital/Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Camden, NJ, USA
[5]Boehringer Ingelheim, Ingelheim, Germany
[6]Department of Experimental Medicine and Biochemical Sciences, University of Rome Tor Vergata, Rome, Italy
[7]Division of Infectious Diseases, Unit of Antiviral Therapy, Ospedali Riuniti, Bergamo, Italy

*Corresponding author e-mail: david.hall@boehringer-ingelheim.com

Background: The purpose of this study was to develop a tipranavir-weighted mutation score that provides guidance to treating physicians on the relative effect of specific protease mutations on tipranavir activity.
Methods: Weights were developed using data from RESIST tipranavir-treated patients based on regressions of virological response at weeks 8 and 24, accounting for baseline CD4$^+$ T-cell count and background regimen activity. The resulting weighted score and cutoffs were validated using a set of cohort patients external to the tipranavir development programme. Response rates were tabulated for the new weighted score and compared with other tipranavir mutation scores used in clinical practice.
Results: The final weights were 74P, 82L/T, 83D and 47V (+4), 58E and 84V (+3), 36I, 43T and 54A/M/V (+2), 10V, 33F and 46L (+1), 24I and 76V (-2), 50L/V (-4), and 54L (-6). Tipranavir-weighted score susceptibility categories were susceptible ≤3, partially susceptible >3 but ≤10, and resistant ≥11. Week 48 response rates for RESIST patients were 34.6%, 15.9% and 5.9%, respectively. Using the external cohort data (*n*=150), the weighted score was highly associated with week 8 viral load reduction (*P*=0.0027). Only one other score achieved statistical significance.
Conclusions: The tipranavir-weighted score developed and externally validated here, in three datasets representing a broad population of treatment-experienced patients, can be used to make clinical decisions about whether to consider tipranavir in a treatment-experienced patient who has limited treatment options.

Introduction

Guidance documents recommend resistance testing for HIV type-1 (HIV-1) patients whenever changes in antiretroviral regimen are being considered as a result of virological failure [1–3]. Almost all testing is population genotypic-based, for reasons of availability and cost. Interpretation is complicated, especially in patients with extensive treatment experience where there might be evidence of reduced susceptibility to many drugs. Of the six reported algorithms used to assess resistance to ritonavir-boosted tipranavir (TPV/r) on the basis of a population genotypic sequence [4–10], all have different lists of mutations, different weights assigned to mutations and, frequently, different interpretations regarding whether or not TPV/r is suitable for a particular patient. The algorithms based on analyses of large datasets are primarily developed to predict *in vitro* phenotypic resistance. The relationship between phenotypic resistance and response is then subsequently analysed. Other algorithms are based on response in modest-sized patient populations, with limited heterogeneity of prior antiretroviral experience. Finally, some algorithms are based on expert interpretations of all available evidence.

Important factors for deriving an accurate and robust mutation score are the use of a large database of mutation

profiles for score development, adjusting for background regimen support to more accurately assess the drug's antiviral activity and making it clinically relevant by demonstrating a strong relationship with virological response. We used the data from the Phase III TPV development programme to develop a weighted mutation score for TPV/r and from a cohort-based group of TPV/r-treated patients to check validity. The ultimate goal is to provide clinicians who prescribe TPV on the basis of genotypic data the best tools with which to effectively construct an optimal antiretroviral regimen for each patient and to predict each patient's virological response.

Methods

The sequence of methods used in developing, improving and validating the TPV-weighted score (TWS) are described later in detail and summarized in Table 1 and Figure 1.

Table 1. Steps, datasets and analyses to derive and validate the tipranavir-weighted score

Background activity score

RESIST patients evaluable for 8-week response (*n*=1,015)

- 8-Week change from baseline HIV-1 RNA $\log_{10}$ copies/ml
- Multiple linear regression estimation of drug effect

Initial weights

RESIST development set of TPV/r patients selected randomly (*n*=566)

Weights normalized, median across six models selected

- 8 Weeks: change from baseline HIV-1 RNA $\log_{10}$ copies/ml (ANOVA); response >1 $\log_{10}$ copies/ml decrease in HIV-1 RNA (logistic regression); response HIV-1 RNA<400 copies/ml (logistic regression)
- 24 Weeks: change from baseline HIV-1 RNA $\log_{10}$ copies/ml (ANOVA); response >1 $\log_{10}$ copies/ml decrease in HIV-1 RNA (logistic regression); response HIV-1 RNA<400 copies/ml (logistic regression)

Clinical cutoffs

RESIST TPV/r patients (*n*=745)

Weighted score used to select upper and lower cutoffs maximizing AUROC

Adjustment of weights

RESIST independent set of TPV/r patients selected randomly (*n*=179)

Tr51 TPV/r patients with mutations found for ≥3 of the codons 33, 82, 84 and 90

- Weighted score and clinical cutoffs used to identify mutations disproportionately in non-responders below lower clinical cutoff and disproportionately in responders above the upper clinical cutoff
- Optimal weight for improving classification determined by systematic search among 0.5 increment changes in weights

External validation

EX cohort TPV/r experience

- Classification based on adjusted weights evaluated
- Adjustments between original and adjusted weights considered
- Final weighted score compared with other published scores

AUROC, area under the receiver operator curve; EX, protease-inhibitor-experienced patients from Italy; HIV-1, HIV type-1; TPV/r, ritonavir-boosted tipranavir; Tr51, clinical trial 1182.51.

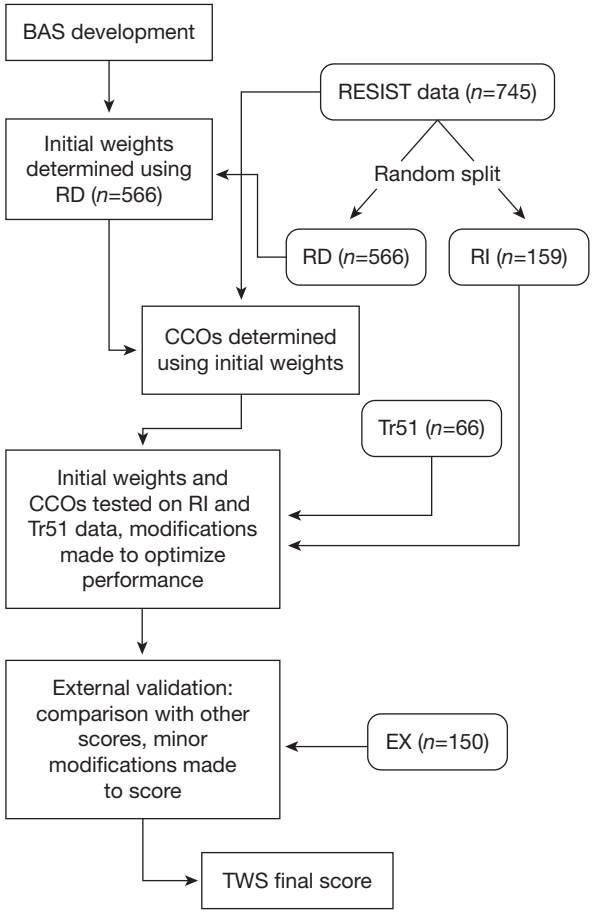
Study patients

There were three sources of patients used for the development and validation of the TWS.

RESIST studies

The RESIST studies were multinational, open-label Phase III trials conducted at 171 sites in Europe and in North and South America, in which 1,483 HIV-1-infected adults were randomized to receive TPV/r (500/200 mg twice daily) or an investigator-selected, ritonavir-boosted comparator protease inhibitor (PI; CPI/r) selected from amprenavir/ritonavir, indinavir/ritonavir, lopinavir/ritonavir and saquinavir/ritonavir The most appropriate CPI/r and optimized background regimen (OBR), which included ≥2 reverse transcriptase inhibitors, with or without the HIV-1 fusion inhibitor enfuvirtide, were selected for each

Figure 1. Flow chart for analyses and datasets



BAS, background activity score; CCOs, clinical cutoffs; EX, protease-inhibitor-experienced patients from Italy; RD, development set; RI; independent set; Tr51, clinical trial 1182.51; TWS, tipranavir-weighted score.

patient. There were 745 patients randomized to TPV/r, of which 566 were randomly selected to be included in the initial score development set (RD) and 179 were selected to be included in the independent set (RI) to be used for assessing the validity of the initial score. Patients were excluded from RESIST if their HIV-1 had ≥3 mutations of the codons 33, 82, 84 and 90.

Clinical trial 1182.51

The clinical trial 1182.51 (Tr51) was primarily a pharmacokinetic trial, studying the coadministration of different PIs with TPV/r. The trial was designed to enrol those patients who were excluded from RESIST because their virus had mutations that had been associated with poor response rates in the dose-ranging trial, 1182.52 [11]. The inclusion of the control arm TPV/r patients from Tr51 in the adjustment of the initial weighted score developed from RESIST patients, was based on the recognition that RESIST entry criteria could have biased the weighting of mutations at positions 33, 82, 84 or 90.

Italian cohort of patients

A group of PI-experienced patients from Italy (EX) were initially administered TPV/r in combination with an OBR from 2004 to 2007. This surveillance-based treatment cohort of patients was included to provide a source of data external from the TPV development programme to validate the weighted score. The prespecified criteria to be included in the cohort sent to the collaborating investigators, which resulted in 150 patients being included in the cohort, were the following: previous experience with at least two PIs; availability of a genotypic resistance test performed within 3 months prior to starting TPV/r; availability of a viral load measurement within a month prior to starting TPV/r and at least one measurement after starting TPV/r; baseline viral load >1,000 copies/ml; and availability of information on which drugs were included in the patient's background regimen.

Adjustments for background regimen support

The development of the weighted score and (model-based) comparisons with other scores were all adjusted for the activity of the OBR. Adjustment for background regimen activity is usually approached using a genotypic susceptibility score (GSS) or a phenotypic susceptibility score [12] designed to predict the support provided by a given regimen when the circulating virus displays the measured level of resistance to the drugs in the regimen. In the RESIST study there was sufficient variability in treatment regimens and resistance status to allow the adjustment to be based on estimation of effect rather than a projection from other results or expert opinion.

Model-building for effect estimation began with indicator variables for each antiretroviral, with three levels of genotypically determined baseline resistance (susceptible [S], partially susceptible [PS] or resistant [R]) and three levels of prior experience (never used, used historically but not in current regimen, or used in the current regimen). Response was measured as change from baseline in $\log_{10}$ copies/ml plasma HIV-1 RNA after 8 weeks of treatment. The results were examined to identify opportunities to reduce dimensionality by grouping drugs that were in the same class or subclasses (for example, non-nucleoside reverse transcriptase inhibitors [NNRTIs] or emtricitabine and lamivudine) and that showed no evidence of differences in effect. This examination revealed that never-used and historically-used could be treated as one category (that is, not in the current regimen). All of the contributions to virological response were estimated in a linear model. Further details of methods and results are provided in Additional file 1.

Analyses were conducted on 1,015 patients who did not deviate from treatment in any way that could make assessment of response to treatment questionable (for example, a changed component of OBR). Contributions to virological response were rounded to the nearest 0.25 $\log_{10}$ copies/ml and were then added for the drugs in each regimen to arrive at a single estimate of OBR antiviral activity for each patient, which we refer to as the background activity score (BAS; Additional file 1) [13]. The BAS was then used as a covariate in all analyses of the relationship between genotype and virological response to TPV/r.

The Tr51 and EX datasets were smaller and with less information for each patient. A GSS based on the Stanford HIVdb [8] for each drug in the respective OBRs was developed using the following algorithm: R (Stanford score >60), weight of 0 for both nucleoside reverse transcriptase inhibitors (NRTIs) and NNRTIs; S (Stanford score <30), weight of 0.5 for NRTIs and 1 for NNRTIs to adjust for potency; and PS (Stanford score 30–60), weight linearly transformed from the Stanford scale to the NRTI scale (0–0.5, that is, weight =1-score/60) or the NNRTI scale (0–1, that is, weight =2-score/30) to weight according to distance from S score. For drugs from new classes (for example, raltegravir, enfuvirtide or maraviroc), if previously used, then a weight of 0 was assigned. If not previously used, then a weight of 1 was assigned. The GSS was calculated by adding the weights from all the drugs included in the patient's OBR.

Score development

Mutations

The initial list of mutations for consideration was restricted to the mutations in the existing TPV

Table 2. Summary of tipranavir mutation scores

Score	Abbreviation	Score range[a]	Clinical cutoffs[b]	Reference
BI tipranavir-weighted score	TWS	-5–18	3 and 10	[2]
BI tipranavir-unweighted score	TUS	0–9	3 and 7	[1]
Agence Nationale de Recherches sur le SIDA	ANRS	-1–3	1 and 2	[3]
Stanford HIVdb	STAN	0–102	30 and 60	[4]
Rega Institute version 7.1.1	REGA	-0.25–10.25	2 and 4	[5]
Monogram Biosciences	MB	0–15	2 and 7	[6]

[a]Observed range of the respective scores. [b]Susceptible defined as less than or equal to the lower limit, partially susceptible as greater than or equal to the lower limit and less than or equal to the upper limit, and resistant defined as greater than or equal to the upper limit. BI, Boehringer Ingelheim.

mutation score (10V, 13V, 20M/R/V, 33F, 35G, 36I, 43T, 46L, 47V, 54A/M/V, 58E, 69K, 74P, 82L/T, 83D and 84V) plus five mutations associated with increased susceptibility to TPV (24I, 30N, 50L/V, 54L and 76V) [6].

Determination of initial weights

Weights were based on multiple linear regression estimates of the above mutations as predictors of virological response (change from baseline as a continuous response variable and 1 $\log_{10}$ copies/ml reduction, viral load [VL]<400 copies/ml as dichotomous response variables) at weeks 8 and 24. Initial weights were created using the RD by normalizing the regression estimates for each model to a -10–10 range and then taking the integer-truncated median of the weights across the six models. The models used are described in Table 1.

Adjustment of weights

The initial weights based on the RD, as described earlier, were then tested on the RESIST RI and Tr51 to assess the consistency of the weights in a set of patients external to the RD data and a set of patients more experienced than RESIST with potentially different mutational patterns because of the inclusion of patients with ≥3 mutations at positions 33, 82, 84 and 90. Mutations that were over-represented in patients with discordance between prediction of response (based on the initial weights) and actual virological response (TPV/r-attributable response, defined as a VL reduction of at least 0.5 $\log_{10}$ copies/ml over BAS support) in the RESIST RI and/or Tr51 data were considered candidates for weight modification. Modifications were determined by averaging all weights that resulted in a better prediction than the initial weights in the RI and Tr51 datasets. The resulting score was tested using the EX dataset with further modifications made if the results suggested the initial weighting was more predictive. The result is the final TWS.

Calculation of clinical cutoffs

The clinical cutoffs (CCOs) for TPV/r based on TWS were calculated using all TPV/r-treated patients in RESIST by fitting a logistic regression model with week 48 VL<50 copies/ml as the response variable and baseline CD4$^+$ T-cell count, BAS and TWS (categorized into S, PS and R) as the independent predictors of response. All integer-valued lower and upper cutoffs of TWS (with the requirement that the response rate for those predicted resistant was <5%) were included in separate regressions. The values that resulted in the largest area under the receiver operator curve (AUROC) were determined to be the best cutoffs for predicting week 48 response after adjusting for baseline CD4$^+$ T-cell count and BAS contribution.

External validation

To assess the validity of the new weighted score, regressions of week 8 VL reduction and week 48 VL<50 copies/ml were run with baseline VL, GSS and commonly used TPV genotypic algorithms as the predictor variables to assess the strength of each score's association with virological response. To provide a more direct comparison of the scores, pairwise combinations of scores were included in the same regression models. The score that explains more of the variability of virological response will tend to retain its statistical significance, whereas the other score will lose most or all of its association with virological response. The genotypic algorithms included in the comparisons were the TPV/r-unweighted score (TUS) [6], Agence Nationale de Recherches sur le SIDA (ANRS) [7], the Stanford HIVdb (STAN) [8], Rega Institute version 7.1.1 [10] and Monogram Biosciences [5] (Table 2).

Results

Baseline characteristics of patients

The characteristics of patients in each of the four datasets used to derive and validate the TWS are displayed in Table 3. There were no major differences in baseline characteristics between the RD and RI datasets. The Tr51 population was more experienced and, thus, had more baseline resistance, whereas the EX population had less.

Model fitting and determination of weights

Initial weights

The resulting weights from the model fitting and initial weighting strategy are presented in Table 4 as are the prevalences of the mutations considered. Mutations/polymorphisms at codons 13, 20, 30, 33, 35 and 69 were assigned initial weights of 0. At codon 54, leucine (L) received weight -7 and alanine (A), methionine (M) or valine (V) received weight 3.

Adjustment of weights

The regression models used to derive the initial weighted score were repeated using other sources of data to assess the consistency of the initial weighting and to help determine if weight modifications should be considered. For those mutations with relatively high prevalences, the model estimates were consistent across datasets at weeks 8 and 48. Notably, 33F showed little association with response in RESIST and EX populations, but a strong relationship with decreased response at week 8 in the Tr51 population.

Based on prevalence in patients who were misclassified as responders or non-responders, 36I, 47V, 54A/M/V, 58E, 74P and 82L/T were selected as candidates for a weight decrease and 33F, 54L and 84V were selected as candidates for a weight increase. The results of the search for optimal weight modifications for the above mutations and final weights are shown in Table 4.

Calculation of clinical cutoffs

The lower and upper CCOs that resulted in the largest AUROC were 3 and 10, respectively, with an AUROC of 76.4%; thus, TWS values for classifying susceptibility were defined as S, TWS≤3; PS, TWS>3 but ≤10; and R, TWS≥11.

Performance of the tipranavir-weighted score

Based on all RESIST TPV/r-treated patients, the week 8 response rates (at least a 1 $\log_{10}$ copies/ml decrease in VL) by CCOs were 69.6%, 48.0% and 35.3% for TWS=S, PS and R, respectively (P<0.0001). The response rates for patients with (and without) an active OBR (defined as BAS≥0.5) were 79.4% (54.2%), 60.9% (30.8%) and 63.2% (0.0%) for S, PS and R patients, respectively; thus, the effect of TWS on week 8 response can largely be attributed to patients with little support. The large response rate for TWS=R patients with an active OBR is reasonable because response is to the entire regimen and not just TPV/r.

Based on the same RESIST TPV/r-treated patients, the week 48 response rates (VL<50 copies/ml) were 34.6%, 15.9% and 5.9% for S, PS and R patients, respectively (P<0.0001). The response rates for patients with (and without) an active OBR were 45.5% (17.5%), 23.5% (5.8%) and 10.5% (0.0%) for S, PS and R patients, respectively. Although having an active OBR remains a large predictor of success, unlike week 8 response, the TWS predicts week 48 response well in both patients with and without an active OBR.

There was also a strong relationship between TWS and virological response in Tr51 with median week 8 VL reductions of -1.4, -1.0 and -0.7 for TWS=S, PS and R, respectively (P=0.0092). The response rates by TWS classification were similar between the RD and the RI.

Table 3. Summary of demographics and HIV type-1 characteristics of patients across datasets

Variable	RD	RI	Tr51	EX
Patients, *n*	566	179	67	150
Female gender, %	16.1	14	7.5	20.7
Median age, years	43.0	43.0	44.0	44.0
Baseline VL				
Median copies/ml	4.78	4.81	4.78	4.63
>100,000 copies/ml, %	37.1	39.1	40.3	26.0
Baseline CD4$^+$ T-cell count				
Median cells/mm^3	165	143	181	148
<50 cells/mm^3, %	20.2	21.8	20.0	20.6
ENF-treated, %	22.1	23.5	29.9	49.3[a]
BAS				
Median	0.5	0.5	0.5	0.75
≥0.5, %	58.7	59.2	55.2	68.7
ANRS				
Median	1.0	1.0	1.0	1.0
S, %	78.6	74.9	73.1	80.7
R, %	6.0	3.9	10.4	3.3
MB				
Median	4.5	4.5	6.0	3.75
S, %	19.8	15.1	0.0	28.0
R, %	10.4	11.7	34.3	13.3
Rega				
Median	3.5	3.75	5.0	3.25
S, %	16.4	12.3	0.0	27.3
R, %	33.2	35.8	85.1	32.7
STAN				
Median	39.0	41.0	57.0	34.0
S, %	26.0	23.5	1.5	37.3
R, %	5.8	6.1	38.8	13.3
TUS				
Median	3.0	4.0	5.0	3.0
S, %	55.7	47.5	32.8	60.7
R, %	0.2	0.0	9.0	0.7
TWS				
Median	4.0	4.0	7.0	3.5
S, %	42.8	37.4	23.9	50.0
R, %	4.8	3.9	23.9	7.3

[a]Includes two patients who took raltegravir. ANRS, Agence Nationale de Recherches sur le SIDA; BAS, background activity score; ENF, enfuvirtide; EX, protease-inhibitor-experienced patients from Italy; MB, Monogram Biosciences; R, resistant; RD, development set; Rega, Rega Institute version 7.1.1; RI, independent set; S, susceptible; STAN, Stanford HIVdb; Tr51, clinical trial 1182.51; TUS, tipranavir-unweighted score; TWS, tipranavir-weighted score; VL, viral load.

Table 4. Summary of prevalence, weightings and adjustments across datasets for candidate HIV type-1 protease mutations

Mutation	Prevalence[a]	Initial weight[b]	RI plus Tr51 adjustment[c]	EX adjustment[d]	TWS weight[e]
10V	12.9	1	–	–	1
13V	34.8	0	–	–	0
20M/R/V	34.4	0	–	–	0
24I	15.5	-2	–	–	-2
30N	3.8	0	–	–	0
33F	28.7	0	+2	-1	1
35G	1.0	0	–	–	0
36I	52.2	2	-1	+1	2
43T	13.6	2	–	–	2
46L	20.8	1	–	–	1
47V	14.0	6	-2	–	4
50L/V	7.1	-4	–	–	-4
54A/M/V	63.7	3	-1	–	2
54L	7.1	-7	+1	–	-6
58E	15.7	5	-2	–	3
69K	5.1	0	–	–	0
74P	3.4	6	-2	–	4
76V	8.4	-2	–	–	-2
82L/V	4.7	5	-2	+1	4
83D	1.4	4	–	–	4
84V	30.1	2	+1	–	3

[a]Prevalence of the mutation in the entire RESIST clinical development database (pre-tipranavir [TPV] treatment only). [b]Initial weight assigned to each mutation based on the mean of the transformed regression estimates across models obtained using the development set as source data. [c]Adjustment made to mutations that were most prevalent in those patients in the independent set (RI) and/or clinical trial 1182.51 (Tr51) with incorrect resistant prediction (lower weight) or incorrect susceptible prediction (increase weight). For assessing accuracy of prediction, ritonavir-boosted-TPV-attributable response used, defined as at least a 0.5 $\log_{10}$ copies/ml greater decrease in viral load (VL) over background activity score (BAS) support (BAS support maximized at 1 $\log_{10}$ copies/ml), that is, if BAS=0.5 then any VL decrease of at least 1 $\log_{10}$ copies/ml will be considered a response. [d]The protease-inhibitor-experienced patients from Italy (EX) dataset was used to confirm the weight modifications suggested by the RI and Tr51 data. For those with lack of confirmation that the adjustments helped prediction, the weight was adjusted back to the initial weighting. [e]Final TPV-weighted score (TWS) weight assigned to each mutation.

Virological responses across datasets by TWS CCOs are shown in Table 5.

External validation

All scores (Table 2) were included in regression models with week 8 VL reduction and logistic regression models with week 48 VL<50 copies/ml virological response, adjusted for GSS and baseline VL in all models. In predicting week 8 VL reductions, TWS was highly associated with response (P=0.0027). The only other scores to approach statistical significance were TUS (P=0.0509) and MB (P=0.0332).

This pattern of results was similar for week 48 VL<50 copies/ml with only TWS approaching statistical significance (P=0.0792). No other score showed any association with week 48 VL<50 copies/ml. The virological responses across scores for weeks 8 and 48 with associated significance levels are shown in Table 6.

The attained significance levels when each pairwise combination of scores were included in the same regression models are shown in Table 7. When each score was included in a regression model with TWS, TWS maintained statistical significance and all other scores lost any association with virological response to TPV, demonstrating the increased prediction of TWS compared with the other scores. This result was true for both week 8 and week 48 responses.

Discussion

We developed a weighted genotypic score that accurately predicts TPV/r activity. We subsequently validated the TWS on an external cohort and demonstrated that it compares favourably to other scores.

Weighted genotypic scores are used to predict resistance and thereby response, using expert systems (Stanford) or statistical modelling (Monogram, VIRCO; Mechelen, Belgium). Superiority of weighted algorithms has been demonstrated in the course of development [14]. The objective of this exercise was to analytically define a weighted genotypic score to directly predict response to TPV/r.

It is clear that in determining a mutation score, a two-step approach is needed that first reduces the total number of mutations in the model by a univariate procedure and then applies a multiple selection procedure to arrive at a combination that best predicts response [12]. The derivation of the TUW represented the first

Table 5. Summary of virological response for datasets used to develop the TWS by TWS CCOs

Time point[a]	Dataset	Classification			*P*-value[b]
		S (TWS<3)	PS (3<TWS≤10)	R (TWS>10)	
Week 8 VR	RESIST	215/309 (69.6)	193/402 (48.0)	12/34 (35.3)	<0.0001
	RD	169/242 (69.8)	143/297 (48.1)	10/27 (37.0)	<0.0001
	RI	46/67 (68.7)	50/105 (47.6)	2/7 (28.6)	0.0059
Week 48 VR	RESIST	107/309 (34.6)	64/402 (15.9)	2/34 (5.9)	<0.0001
	RD	87/242 (36.0)	48/297 (16.2)	2/27 (7.4)	<0.0001
	RI	20/67 (29.9)	16/105 (15.2)	0/7 (0.0)	0.0057
Median week 8 VLR	Tr51	-1.43 (*n*=16)	-1.00 (*n*=34)	-0.73 (*n*=16)	0.0092

Values presented are *n*/total *n* (%) unless otherwise indicated. [a]Week 8 virological response (VR)>1 $\log_{10}$ copies/ml reduction from baseline, week 48 VR= viral load (VL)<50 copies/ml and week 8 VL reduction (VLR)=VL change from baseline 1 $\log_{10}$ copies/ml. [b]*P*-values are based on a regression of the week 8 VLR or week 48 VL<50 copies/ml with tipranavir-weighted score (TWS), baseline VL and background activity score (genotypic susceptibility score for clinical trial 1182.51 [Tr51]). CCO, clinical cutoff; PS, partially susceptible; R, resistant; RD, development set; RI, independent set; S, susceptible.

Table 6. Summary of virological response in EX for all score classifications

Time point[a]	Score	Classification			*P*-value[b]
		S	PS	R	
Week 8 VR	TWS	55/70 (78.6)	29/62 (46.8)	4/11 (36.4)	0.0027
	TUS	62/87 (71.3)	26/55 (47.3)	0/1 (0.0)	0.0509
	STAN	38/54 (70.4)	42/69 (60.9)	8/20 (40.0)	0.2867
	REGA	29/41 (70.7)	37/56 (66.1)	22/46 (47.8)	0.2641
	ANRS	72/116 (62.1)	14/22 (63.6)	2/5 (40.0)	0.1021
	MB	32/42 (76.2)	49/82 (59.8)	7/19 (36.8)	0.0332
Week 48 VR	TWS	29/75 (38.7)	20/64 (31.3)	1/11 (9.1)	0.0792
	TUS	35/91 (38.5)	15/58 (25.9)	0/1 (0.0)	0.5379
	STAN	20/56 (35.7)	26/74 (35.1)	4/20 (20.0)	0.8520
	REGA	16/41 (39.0)	19/60 (31.7)	15/49 (30.6)	0.8366
	ANRS	39/121 (32.2)	9/24 (37.5)	2/5 (40.0)	0.8523
	MB	18/42 (42.9)	28/88 (31.8)	4/20 (20.0)	0.2241

Values presented are *n*/total *n* (%) unless otherwise indicated. [a]Week 8 virological response (VR)>1 $\log_{10}$ copies/ml reduction from baseline and week 48 VR= viral load (VL)<50 copies/ml. [b]*P*-values are based on a regression of the week 8 VL reduction or week 48 VL<50 copies/ml with the specified score, baseline VL and genotypic susceptibility score as predictors using the protease-inhibitor-experienced patients from Italy (EX) dataset. ANRS, Agence Nationale de Recherches sur le SIDA; MB, Monogram Biosciences; PS, partially susceptible; R, resistant; Rega, Rega Institute version 7.1.1; S, susceptible; STAN, Stanford HIVdb; TUS, tipranavir-unweighted score; TWS, tipranavir-weighted score.

step in this process, which included mutations associated with reduced response to TPV/r if present before treatment, and also those mutations that emerged upon failure with TPV/r [6]. We began the derivation of the TWS with the predefined set of TPV score mutations plus a small number of mutations associated with increased susceptibility to TPV in the initial analysis [6]. The analyses presented here result in a weighting of mutations based on multivariable parametric or non-parametric methodology with adequate adjustments for background activity and proved to be an important advance for determining weighted mutation scores for PIs. The predictive ability of the score and better performance compared with other commonly used algorithms was demonstrated using an independent set of patients external to the TPV development programme. Clinical cutoffs were determined using the RESIST data. It is important to note that the cutoffs were established using the same set of patients used to develop the score; thus, they might not prove to be adequate when the score is used in other populations of patients. It is, therefore, advisable to keep the actual value of the weighted score in mind and not rely completely on the interpretation.

A score based entirely on mutations that are associated with reduced susceptibility will not predict response well for PI-experienced patients. The mutations 24I, 50L/V, 54L and 76V, selected by other PIs, predicted increased response to TPV/r and thus remained in the final score with large negative weights; however, the effect of these mutations on long-term response has not yet been fully described and needs further study. One criticism of the initial TPV score proposed by Baxter *et al.* [6] was that a number of the score mutations had been previously identified as common polymorphisms in non-B HIV-1 viruses [15]. The TWS reduces this

Table 7. Summary of regressions for pairwise combinations of scores in the same model using EX validation dataset

Score	Week 8 VL reduction		Week 48 VL<50 copies/ml	
	TWS[a]	Score[b]	TWS[a]	Score[b]
TUS	0.0114	0.7517	0.0643	WD
STAN	0.0240	WD	0.1644	WD
REGA	0.0025	WD	0.0157	WD
ANRS	0.0011	WD	0.0074	WD
MB	0.0193	WD	0.0470	WD

[a]Tipranavir-weighted score (TWS) indicates significance level for TWS when included in the regression model with baseline viral load (VL), background activity score (BAS) and the specified tipranavir mutation score. [b]Score indicates significance level for the specified tipranavir mutation score when included in the regression model with baseline VL, BAS and TWS. A regression estimate in the wrong direction (that is, increasing score associated with better response) is denoted by WD, as the *P*-value is not meaningful. ANRS, Agence Nationale de Recherches sur le SIDA; EX, protease-inhibitor-experienced patients from Italy; MB, Monogram Biosciences; REGA, Rega Institute version 7.1.1; STAN, Stanford HIVdb; TUS, tipranavir-unweighted score.

concern by giving no weight to the polymorphisms/mutations 13V, 20M/R/V, 35G and 69K.

It is both an advantage and a limitation of this study that the new drugs, such as maraviroc and raltegravir, now available as companions to TPV, were not available in the RESIST study. The limited support for TPV enabled us to see clearly the influence of mutations on response. At the same time, only a fraction of the results are directly relevant to the effect of a potent regimen, with all antiretrovirals active against the patient's HIV-1 population. The availability of a reliable score, such as TWS, for predicting the efficacy/inactivity of TPV, can help identify patients who can take advantage of TPV. This is particularly true in cases where other PI options are no longer available, either because of issues regarding resistance or tolerance.

In conclusion, the TWS, developed and externally validated here in three datasets representing a broad population of treatment experienced patients, can be used to make clinical decisions about whether to consider TPV in a treatment-experienced patient who has limited treatment options. These data demonstrate the importance of large datasets for development as well as external validation of mutational scores, in order to increase the applicability of the score in clinical practice.

Disclosure statement

JMS has recently received research support, honorarium or consulting fees from Boehringer Ingelheim, Abbott Laboratories, MSD, Roche, Gilead Sciences, GlaxoSmithKline, Pfizer, Monogram Biosciences, Tibotec, Siemens, Ambrillia and Bristol–Myers Squibb. CB has recently received research support, honorarium or consulting fees from Abbott Laboratories, Merck, GlaxoSmithKline, Roche, Pfizer and Viiv. JB is a consultant to Boehringer Ingelheim and is on the speakers' bureau for Pfizer, Boehringer Ingelheim and Tibotec. JS, CT and DH are full-time employees of Boehringer Ingelheim. FM has served as a consultant on advisory boards for Boehringer Ingelheim, Bristol–Myers Squibb, Gilead Sciences, GlaxoSmithKline, Roche and Tibotec; he has received lecture fees from Abbott Laboratories, Bayer, Bristol–Myers Squibb, Gilead Sciences, GlaxoSmithKline, MSD, Pfizer and Roche, and has received research and educational grants from Boehringer Ingelheim, Bristol–Myers Squibb, GlaxoSmithKline, Janssen–Cilag and Roche. CP and MS declare no competing interests.

Additional file

Additional file 1: BAS development and regression estimates for each mutation can be found at http://www.intmedpress.com/uploads/documents/AVT-10-OA-1653_Schapiro_Add_file.pdf

References

1. Hirsch MS, Gunthard HF, Schapiro JM, *et al*. Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society–USA panel. *Clin Infect Dis* 2008; **47:**266–285.
2. Vandamme AM, Sönnerborg A, Ait-Khaled M, *et al*. Updated European recommendations for the clinical use of HIV drug resistance testing. *Antivir Ther* 2004; **9:**829–848.
3. United States Department of Health and Human Services Panel on Clinical Practices for Treatment of HIV Infection. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents 2008. (Updated 1 December 2009. Accessed 21 October 2009.) Available from http://aidsinfo.nih.gov/contentfiles/AdultandAdolescentGL.pdf
4. Vermeiren H, Van Craenenbroeck E, Alen P, *et al*. Prediction of HIV-1 drug susceptibility phenotype from the viral genotype using linear regression modeling. *J Virol Methods* 2007; **145:**47–55.
5. Parkin N, Chappey C. Protease mutations associated with higher or lower than expected tipranavir (TPV) susceptibility based on the TPV mutation score. *13th Conference on Retroviruses and Opportunistic Infections*. 5–8 February 2006, Denver, CO, USA. Poster 637.
6. Baxter JD, Schapiro JM, Boucher CA, *et al*. Genotypic changes in human immunodeficiency virus type 1 protease associated with reduced susceptibility and virologic response to the protease inhibitor tipranavir. *J Virol* 2006; **80:**10794–10801.

7. Marcelin AG, Masquelier B, Descamps D, *et al*. Tipranavir-ritonavir genotypic resistance score in protease inhibitor-experienced patients. *Antimicrob Agents Chemother* 2008; **52:**3237–3243.
8. Stanford University HIV Drug Resistance Database. (Updated 24 August 2010. Accessed 15 July 2008.) Available from http://hivdb.stanford.edu/pages/algs/HIValg.html
9. Van Laethem K, De Luca A, Antinori A, Cingolani A, Perna CF, Vandamme AM. A genotypic drug resistance interpretation algorithm that significantly predicts therapy response in HIV-1-infected patients. *Antivir Ther* 2002; **7:**123–129.
10. Rega Institute for Medical Research. Algorithm for the use of genotypic HIV-1 resistance data, Rega v7.1.1©, 5 July 2007. (Updated 10 June 2008. Accessed 15 July 2008.) Available from http://www.rega.kuleuven.be/cev/index.php?id=30
11. Walmsley SL, Katlama C, Lazzarin A, *et al*. Pharmacokinetics, safety, and efficacy of tipranavir boosted with ritonavir alone and in combination with other boosted protease inhibitors as part of optimized combination antiretroviral therapy in highly treatment-experienced patients (BI Study 1182.51). *J Acquir Immune Defic Syndr* 2008; **47:**429–440.
12. Brun-Vézinet F, Costagliola D, Khaled MA, *et al*. Clinically validated genotype analysis: guiding principles and statistical concerns. *Antivir Ther* 2004; **9:**465–478.
13. Hall DB, Schapiro JM, Boucher CAB, Baxter JD, Piliero P, Scherer J. Development and performance of a trial-specific genotypic sensitivity score based on RESIST data that incorporates treatment history in addition to predicted susceptibility. *Reviews in Antiviral Therapy* 2008; **2:**106.
14. Parkin NT, Chappey C, Petropoulos CJ. Improving lopinavir genotype algorithm through phenotype correlations: novel mutation patterns and amprenavir cross-resistance. *AIDS* 2003; **17:**955–961.
15. Poveda E, de Mendoza C, Parkin N, *et al*. Evidence for different susceptibility to tipranavir and darunavir in patients infected with distinct HIV-1 subtypes. *AIDS* 2008; **22:**611–616.

Accepted for publication 30 June 2010