

# Genotypic tipranavir scores as predictors of response

J Schapiro<sup>1</sup>, J Scherer<sup>2</sup>, R Vinisko<sup>2</sup>, VM Kohlbrenner<sup>2</sup>, JD Baxter<sup>3</sup>, CAB Boucher<sup>4</sup>, RD MacArthur<sup>5</sup>, DB Hall<sup>2</sup>

<sup>1</sup> National Hemophilia Center, Israel

<sup>2</sup> Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT, United States

<sup>3</sup> Cooper Hospital/UMDNJ-Robert Wood Johnson Medical School, New Jersey, United States

<sup>4</sup> Department of Virology, University Medical Center Utrecht, the Netherlands

<sup>5</sup> Wayne State University, Detroit, MI, USA

## Poster Number P3.4/16

11th European AIDS Conference (EACS),  
Madrid, Spain  
24–27 October 2007

Corresponding Author: Jonathan M Schapiro  
National Hemophilia Center  
Israel  
Tel: +972 3 5302409  
Email: jmschaps@zahav.net.il

## Introduction

Many genotypic scores have recently been developed to predict phenotype and/or response to treatment with tipranavir. An evaluation of how well these scores predict virologic response to tipranavir in the RESIST trials should provide some direction regarding their applicability.

## Methods

In the RESIST trials 745 patients were treated with TPV/r and had evaluable results for treatment response and baseline genotype. The TPV mutation scores compared in this study are provided below in Table 1.

Table 1: Summary of various tipranavir mutation scores compared with objective of the score and cut-offs used for clinical interpretations of resistance.

Score	Abb.	Objective	Range	CC01 <sup>1</sup>	CC02 <sup>2</sup>
MacArthur [1]	Mac	Develop a simple score to predict response	0 to 4	1	N/A
Marcelin and Calvez [2]	Mar	Develop a simple score to predict response	-1 to 4	1	N/A
Parkin [3]	P	Modified tipranavir score incorporating some weighting relating genotype to phenotype	-1.5 to 13	2	7
Stanford HIVdb [4]	ST	Weighted score that predicts virologic response	0 to 89	30	60
Virco Virtual Phenotype® [5]	VP	Weighted score that predicts tipranavir phenotype	0.1 to 31.6	1.2	5.4
Rega Institute [6]	REGA	Weighted score that predicts virologic response	0 to 9.75	2	4
BI Tipranavir Unweighted Score [7]	TUW	Develop a score to predict phenotype and virologic response	0 to 9	3	7
BI Tipranavir Weighted Score [8]	TW	Weighted score that predicts virologic response	-7 to 20	3	10

<sup>1</sup>CC0 = Clinical Cut-Off: if score < CC01, then the virus is predicted to be fully susceptible to tipranavir; if CC01 < score < CC02 then virus is predicted to be partially susceptible to tipranavir; if score > CC02 then virus is predicted to be resistant to tipranavir. Only one cut-off was suggested by MacArthur and Marcelin/Calvez due to the small range of their respective scores. So, for these scores the patient is either predicted to be susceptible or resistant depending on the score being < or > than CC01. Note also that the Stanford score provides 5 categories of resistance in intervals of 15, but these were collapsed to 3 for comparison purposes.

To compare the prediction accuracies of each score, week 8 ( $\geq 1$  log drop in HIV RNA), 24 and 48 (HIV RNA < 50 copies/mL) virologic response associations with each score were compared using logistic regression (assessed by areas under the receiver operator curves (AUROC)), adjusted for background drug activity (see below for a brief description). Spearman correlations of each score (R) with weeks 8, 24 and 48 viral load decline were also compared. These comparisons were made for all TPV/r-treated patients in RESIST and also for those with an OBR predicted to provide less than one log of activity which would exclude, among others, those patients including new enfuvirtide in their regimen.

Comparisons were also made using the interpretation algorithms shown above in Table 1. An attempt was made to determine optimal cuts using the same methodology used to arrive at the clinical cut-offs for the BI tipranavir weighted score [8]. Comparisons of the predictive accuracy based on the existing cut-offs to our proposed cut-offs are shown as well as comparisons amongst the scores using the new cut-offs.

### Background activity scores

The development of the weighted score and (model-based) comparisons to other scores were all adjusted for the activity of the background regimen. The estimated activities for each patient were based on a novel approach developed by Hall et al [9] that estimates the contribution of each drug in the Optimized Background Regimen (OBR) using a linear model-based approach accounting for the baseline resistance (using Virtual Phenotype™) and historical use of each drug. Often, the activity of the drug in the background regimen, especially with the NRTIs, is largely dependent on the previous pattern of use rather than the baseline resistance. These adjustments prove to be much better at predicting response than a simple count of the number of active drugs in the OBR.

## Results

### Comparisons of score weights and score correlations

Table 2 shows all amino acid substitutions that are given a weight in each of the 8 scores compared. Mutations that appear in at least 5 of the 8 scores are 33F, 36I, 43T, 46L, 47V, 54A/M/V, 58E, 82L/T and 84V. Some notable differences across scores are:

- Mutations at V32 receiving large weights in the Stanford score and little inclusion in other scores.
  - 54L receives a weight of 5 by Stanford but a negative weight by Parkin (-1) and large negative weight by the TW score (-7)
  - 74P and 83D, fairly rare mutations, receive much larger (relative) weights by TW than the other scores
  - 84V is strongly associated with changes in phenotype but less of a relationship is seen with virologic response, thus receives discrepant relative weights with larger weights given by P, ST and REGA compared to TW.
  - 90M is given a relatively large weight by ST but only little consideration by the other scores.
- Table 3 shows the pairwise correlations of the 8 scores. The strongest correlations were between REGA with P (0.84) and ST (0.83). The TW correlated most strongly with P (0.76) with correlations with both simple scores similar (0.39 for both) and also similar correlations with the other scores (ranging from 0.60 to 0.67). The Parkin score objective was similar to TW in that it weighted some existing TPV score mutations and added a few mutations that increased response so the fact that it correlated better with TW than TUW is reasonable.

Table 2: Mutations and weights included in each score

Site	Amino Acid	Mar	Mac	P	ST	TUW	REGA	TW
L10	I		0.5	2			0.5	
	V		0.5	3		1	0.5	1
	F			2			0.5	
V11	I			2				
	L		1				1	
I13	V		0.5	8		1	0.25	
	M			1			0.25	
K20	I		0.5	1			0.5	
	V		1	1			0.5	
L24	T				2			-2
	F		-1		1			
D30	N		-1					
	I		1		8		1	
V32	A			5				
	F			1				
L33	I		1	1	8	1	1	
	MV						0.5	
E35	G		1	2	1		0.5	
	D		1				0.25	
M36	I	1		1		1	0.25	2
	LV	1						
L38	W							1
	R41							1
K43	T		1	2	1	1	0.5	2
	I						0.5	
K45	I						0.5	
	V		0.5	8		1	0.5	1
M46	L		0.5	8			0.5	
	V		0.5	8			0.5	
I47	V	1	2	10	1	1.5	6	
	A			5				
I50	L		-1				-0.5	-4
	V		-1		-5			-4
F53	L	-1		3				
	W	-1						
I54	Y			2				
	A	1	2	15	1	1.5	3	
M			-1	5	1	0.5	-7	
	S	1		5		1	3	
T				5		0.5		
	V			5		0.5		
Q58	E	1		1	10	1	0.5	3
	K	1		2	1	0.5	5	
H69	K	1		1			0.75	
	N/Q/R/Y	1						
A71	L			1			0.5	
	V			3			0.5	
T				1			0.5	
	F			2			0.5	
G73	I				2		1	
	C		1	2			0.5	
S				2			0.5	
	A			2				
T74	P		1	2	1		6	
	V		-1				-2	
L76	V			10			0.5	
	V82	AC/M		15			0.5	
I			-1					
	L			40	1	1	5	
T		1	2	25	1	1.5	5	
	D			2	1		4	
N83	V		1	2			1	
	D		1	2			1	
I84	V	1	2	25	1	1.5	2	
	A/C			25				
N88	D						0.5	
	S						0.5	
L89	V	1		1			0.5	
	I/T	1		1			0.5	
M/R		1						
	M		0.5	6			0.5	

Table 3: Spearman correlation coefficients between scores

	Mac	Mar	P	ST	VP	REGA	TUW	TW
Mac								
Mar	0.08							
P	0.56	0.38						
ST	0.71	0.12	0.70					
VP	0.46	0.25	0.67	0.61				
REGA	0.71	0.34	0.84	0.83	0.61			
TUW	0.37	0.51	0.62	0.53	0.45	0.65		
TW	0.39	0.39	0.76	0.60	0.63	0.65	0.67	

### Comparison of prediction accuracies and correlations with virologic response

Tables 4 and 5 show the comparison summaries for all patients (Table 4) and patients without at least a log of antiviral activity from their OBR (Table 5). Figures 1 and 2 show the Receiver Operator Curves for the 5 weighted scores. The results are consistent and can be summarized as follows:

- Virtual Phenotype, the BI tipranavir weighted score and Parkin's weighted score perform the best.
- The unweighted tipranavir score and weighted scores by Stanford and the Rega Institute do not perform as well but do better than the simple scores developed by MacArthur and Marcelin/Calvez.
- The predictions and correlations for all scores, but especially TW and VP, are stronger for the patients with less OBR support.

Table 4: Spearman correlations and prediction accuracies for all TPV-treated patients

Score	Week 8		Week 24		Week 48	
	R	AUROC	R	AUROC	R	AUROC
Mac	0.08	70.8%	0.10	70.8%	0.12	70.5%
Mar	0.12	70.7%	0.15	70.6%	0.13	69.5%
P	0.21	73.9%	0.23	73.9%	0.22	73.1%
ST	0.14	71.5%	0.15	72.8%	0.18	71.8%
VP	0.29	74.9%	0.27	74.5%	0.26	75.6%
REGA	0.17	72.6%	0.21	73.5%	0.20	73.1%
TUW	0.17	72.3%	0.21	73.3%	0.21	73.0%
TW	0.26	74.1%	0.27	74.1%	0.26	73.8%

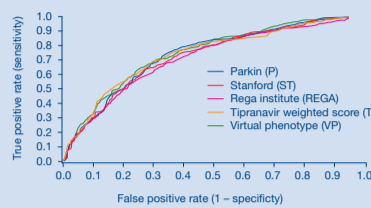


Figure 1: Receiver Operator Curves for the five weighted scores in the evaluation of week 8 virologic response (1 log drop in VL), adjusted for OBR activity

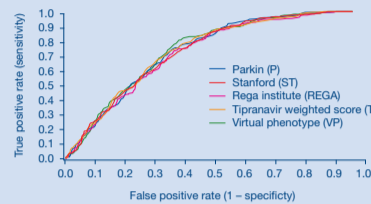


Figure 2: Receiver Operator Curves for the five weighted scores in the evaluation of week 24 virologic response (VL < 50 copies/mL), adjusted for OBR activity

Table 5: Spearman correlations and prediction accuracies for TPV-treated patients without at least 1 log in predicted antiviral activity from their OBR

Score	Week 8		Week 24		Week 48	
	R	AUROC	R	AUROC	R	AUROC
Mac	0.18	64.5%	0.18	71.3%	0.19	71.9%
Mar	0.21	65.4%	0.21	69.7%	0.18	69.5%
P	0.35	71.1%	0.38	77.3%	0.36	76.2%
ST	0.25	66.3%	0.28	74.5%	0.28	73.3%
VP	0.41	72.4%	0.38	78.5%	0.35	78.4%
REGA	0.30	68.5%	0.34	75.8%	0.31	75.9%
TUW	0.29	67.8%	0.31	74.4%	0.28	74.2%
TW	0.38	70.6%	0.40	76.8%	0.37	75.8%

### Response rates by existing cut-offs and comparison with proposed cut-offs

The response rates by clinical interpretation of each score are presented for week 24 response in Figure 3. The results of the search for optimal cut-offs for each score are presented in Table 6. A few key points about these results:

- The Rega Institute predicts 52% of patients to be fully resistant to tipranavir and only 10% to be susceptible. It is clear that their cut-offs need to be revised.
- The lack of score range for MacArthur and Marcelin/Calvez, due to their simplicity, also results in an overprediction of resistance, especially for MacArthur.
- The percentage of responders for those patients predicted to be susceptible is highest for the BI tipranavir weighted score while maintaining a low percentage of responders in the group of patients predicted to be resistant (of 95 patients predicted by TW to be resistant to tipranavir, only 6 responded).
- In general, the existing cut-offs were similar to the new cut-offs proposed in Table 5. Rega's cut-offs are the exception, though, with proposed cut-offs shifting considerably from the existing cut-offs.
- After applying the same methodology to each score to arrive at optimal cut-offs, the BI tipranavir weighted score clinical interpretation outperformed the other scores in terms of accurately predicting week 24 virologic response.

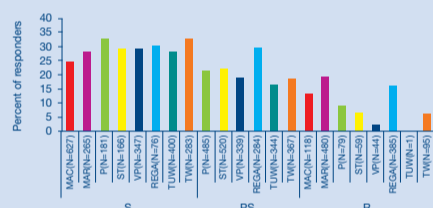


Figure 3: Week 24 response rates by genotypic interpretation algorithm

Table 6: Optimal cut-offs for each score

Score	Existing cut-offs			Proposed cut-offs <sup>1</sup>			Numbers of patients (% responders) in each resistance classification <sup>2</sup>		
	CC01	CC02	AUROC	CC01	CC02	AUROC	S	PS	R
Mac	1	N/A	71.0%	1	3	71.1%	627 (24.4%)	116 (13.8%)	2 (0.0%)
Mar	1	N/A	70.1%	0	3	71.1%	265 (28.3%)	474 (19.8%)	6 (0.0%)
P	2	7	72.7%	3	7	73.6%	289 (32.2%)	377 (18.3%)	79 (8.9%)
ST	30	60	72.1%	41	60	72.7%	386 (28.5%)	300 (18.3%)	59 (6.8%)
VP	1.2	5.4	73.5%	1.6	5	74.3%	473 (28.5%)	225 (14.7%)	47 (2.1%)
REGA	2	4	72.6%	3.75	5.75	73.5%	304 (31.9%)	344 (18.3%)	97 (8.3%)
TUW	3	7	71.9%	1	7	73.5%	134 (41.0%)	610 (18.7%)	1 (0.0%)
TW	3	10	75.0%	3	10	75.0%</			