



# virco<sup>®</sup>TYPE HIV-1 powered by *VirtualPhenotype*<sup>™</sup> - LM: a new and improved approach for quantitative Fold Change predictions

Linear Regression Modeling identifies and scores  
resistance-associated mutations from a  
Genotype-Phenotype correlative database

---

## SUMMARY

---

- The virco<sup>®</sup>TYPE HIV-1 assay provides quantitative (predicted Phenotype) and qualitative (Genotype) resistance information.
- *VirtualPhenotype*<sup>™</sup>- LM (VPT-LM) is a new and improved methodology to make phenotypic predictions for virco<sup>®</sup>TYPE HIV-1.
- With VPT-LM, the virco<sup>®</sup>TYPE HIV-1 report gives FC predictions for all drugs and all samples, totally eliminating rules-based calls.
- New Clinical Cut-Offs are added for tipranavir and abacavir. Cut-Offs for other drugs are updated.
- The VPT-LM predictive model uses a Linear Regression Model equation for each ARV drug to predict the drug susceptibility Phenotype (Fold Change in IC<sub>50</sub>, FC) from the Genotype.
- The VPT-LM equations are based on a Genotype/Phenotype correlative database that is representative of currently circulating viral strains. The database is updated with new samples every 2 months.
- Unlike many rules-based algorithms, VPT-LM uses weighted contributions of individual mutations and mutation pairs to make FC predictions.
- VPT-LM is objective and fully data-driven, and makes more accurate FC predictions compared to the first generation VPT, including predictions for complex or uncommon mutation combinations.
- VPT-LM better reflects the dynamic range of phenotypic resistance measured by Antivirogram<sup>®</sup> (AVG).
- Minor changes will be introduced on the virco<sup>®</sup>TYPE HIV-1 report:
  - Column “Database Matches” on page 2 will be removed.
  - Due to very limited use in clinical practice, ddC, DLV, RTV, and unboosted SQV will be removed from the virco<sup>®</sup>TYPE HIV-1 report.
  - The Cut-Off values for virco<sup>®</sup>TYPE HIV-1 will be updated.

---

## 1. Introduction

Antiretroviral (ARV) drug resistance testing is a key component of HIV-1 disease management and is aimed at assisting clinicians and virologists in optimizing the selection of ARV regimens for individual patients.

Genotypic resistance tests are qualitative in nature and identify mutations in the viral genome that affect ARV drug susceptibility, whereas phenotypic tests are quantitative and determine the level of resistance of the HIV-1 virus to ARV drugs, expressed as Fold Change in IC<sub>50</sub> (FC). Genotypic resistance information is typically interpreted with rules-based (computer) algorithms, whereas interpretation of phenotypic data requires Cut-Off values to distinguish resistant strains from susceptible ones.

Virco<sup>®</sup>TYPE HIV-1 combines genotypic information with a predicted Phenotype and Clinical Cut-Off (CCO) values to give optimal guidance to clinical decision-making. The Phenotype is predicted from the Genotype using an extensive database of previously determined Genotypes and Phenotypes from clinical isolates, and a bio-informatics based methodology *VirtualPhenotype*<sup>™</sup> (VPT).

The first generation *VirtualPhenotype*<sup>™</sup> used patterns of resistance-associated mutations to search the database for viruses with similar mutational profiles. The average Fold Change of all database matches for the patient's virus was then calculated and reported. *VirtualPhenotype*<sup>™</sup>- LM is a new version of VPT that uses Linear Regression Modeling and Virco's Genotype/Phenotype correlative database. This new approach identifies resistance-associated mutations, calculates their weighted contribution to resistance, and makes phenotypic predictions. VPT-LM will be introduced in the virco<sup>®</sup>TYPE HIV-1 report as from version 4.0.00.

## 2. Linear Modeling

Linear Regression Analysis or Linear Modeling (LM) is a commonly used statistical technique to study the relationship between observations or variables. The simplest form of a regression model contains a *dependent* variable or "outcome variable" Y and a single *independent* variable X. A typical example is the dependence of the blood pressure (BP) Y on the age X of a person. When age and BP are known for a sufficiently large number of subjects, this information can be used to calculate a Linear Model that could predict BP from age. The accuracy of the prediction depends on the characteristics of the dataset that was used to generate the model [number of data pairs, data diversity (i.e. narrow or broad range of ages), accuracy of BP measurements etc.].

Often the outcome variable Y does not depend on a single independent variable X. A person's BP, for example, is not only influenced by his/her age but also by other variables such as smoking, diet, stress and genetic factors. If data for those additional independent variables is available, a multiple regression model can be calculated that would predict BP from the contribution of each individual variable. The multiple regression model provides a more accurate prediction than the simple model, since it considers not only age but multiple variables that influence BP.

### 3. VirtualPhenotype™- LM

Applied to HIV-1 drug resistance analysis, VirtualPhenotype™- LM uses a set of multiple linear regression models (one for each drug) capable of predicting a sample's Phenotype (the outcome variable Y, FC) from the contribution of individual mutations (independent variables X) in its Genotype. The LMs are derived from a database that contains the Genotypes (all mutations) and Phenotypes (measured FC according to AVG) from 6,143 – 41,958<sup>(3)</sup> clinical isolates. By modeling the relationship between the Genotype and the Phenotype, each mutation and combination of two mutations that has a significant effect on drug susceptibility is identified and assigned a Resistance Weight Factor (RWF) according to its contribution to the FC. VPT-LM analyzes not only individual mutations but also pairs of mutations, thus taking into account the effect that synergistic or antagonistic interactions between mutations may have on phenotypic resistance. The FC of an unknown sample is predicted by adding up the contributions or RWFs of the individual mutations and mutation pairs that were identified in the patient's virus Genotype. (Fig.1)

fig.1

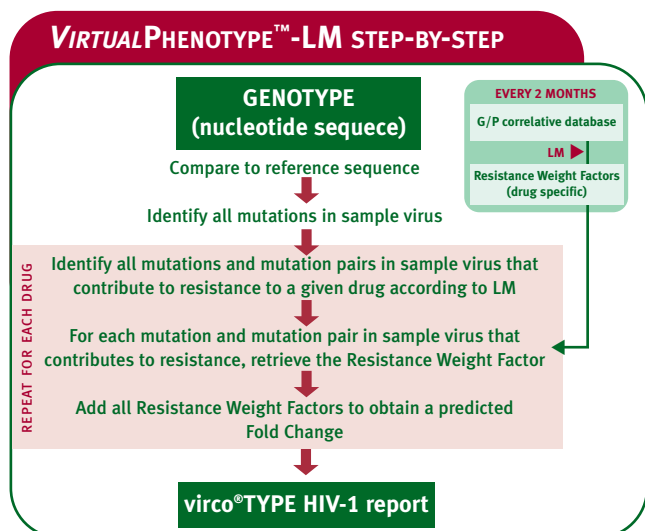
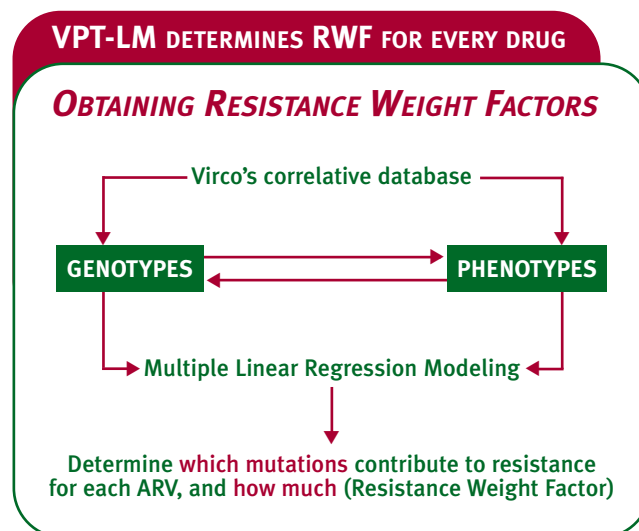


fig.2



#### 3.1 RESISTANCE WEIGHT FACTORS

The establishment of RWFs is an essential step in VPT-LM predictions. RWFs are data-driven indicators of the contribution of a mutation, or pair of mutations, to phenotypic resistance for a given drug. RWFs are expressed as log(FC) (Fig 2).

- **RWFs are drug-specific.** Different mutations contribute to resistance for different drugs, and the contribution of a specific mutation to resistance is different for each drug.
- **RWFs are determined for individual mutations and pairs of mutations.** By including pairs of mutations, VPT-LM accounts for (second order) mutation interactions (synergistic or antagonistic).
- **RWFs can be positive or negative.** Mutations that have a positive RWF contribute to resistance, whereas mutations that have a negative RWF increase drug susceptibility.
- **RWFs are determined by Linear Regression Analysis on a Genotype/Phenotype correlative database.** This database is updated with new samples every two months. After every database update the RWFs are re-calculated.

<sup>(3)</sup> for virco®TYPE HIV-1 v4.0.00. Numbers vary per drug: 6,143 for TPV - 41,958 for NVP

### 3.2 MAKING FC PREDICTIONS

VPT-LM makes predictions of a sample's Phenotype from the Genotype by accounting for the contributions of the individual mutations and mutation pairs identified in that Genotype. The sum of the RWFs of the individual mutations and mutation pairs equals the logarithm of the predicted FC.

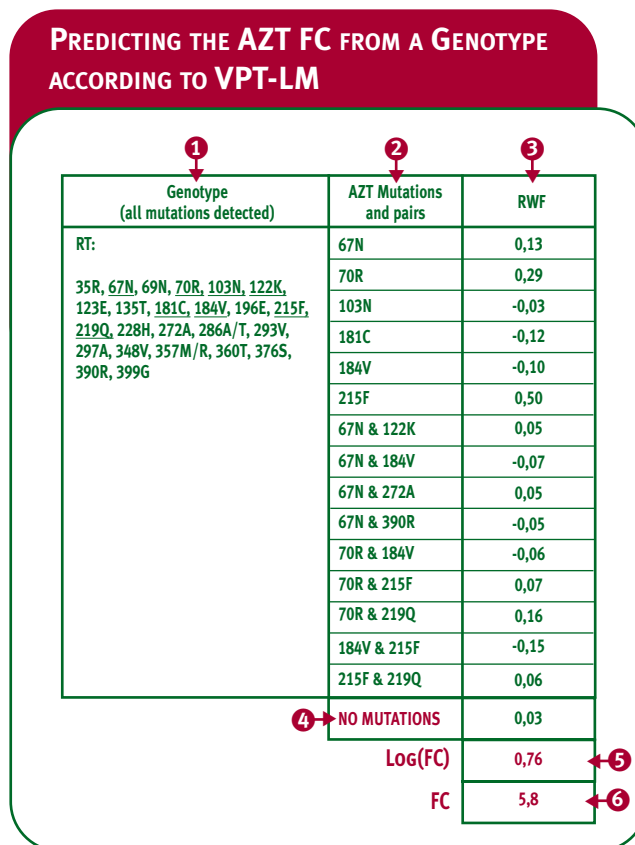
Fig. 3 shows an example of how the FC for AZT is predicted from a viral Genotype.

1. After aligning a sample nucleotide sequence with the sequence of the reference laboratory strain, all mutations in the RT and PR genes are identified.
2. The mutations and pairs of mutations that affect susceptibility to AZT in the LM model are included in the equations used to predict Fold Change. These are underlined in the Genotype, and listed in column 2. For other drugs, other mutations are considered.

3. The corresponding RWFs are retrieved. An RWF is specific for a mutation and a drug. For instance, 67N is also considered for FTC but the RWF of 0.13 log FC is specific for AZT (it is 0.33 log FC for FTC). The negative RFW of 184V for AZT is indicative of its resensitizing effect for that drug; 184V is also considered for 3TC but with a large positive RWF (1.71 log FC), driving the FC prediction to higher resistance levels.
4. A “no mutations” score is added because clinical samples with no resistance-associated mutations often have a FC different from 1. This score is drug-specific. For AZT it is 0.03 log FC, meaning that a virus without AZT resistance-associated mutations will have a predicted FC of  $10^{0.03} = 1.07$ .
5. The sum of all RWFs and the “no mutations” score for this isolate for AZT is 0.76 log FC.
6. The predicted FC is  $10^{0.76} = 5.8$

A limitation of VPT-LM, like any other data-driven system, is that it requires a substantial amount of data before accurate predictions can be made. If novel mutations arise but are underrepresented in the database, they will not be considered in VPT-LM. To overcome this, the databases and VPT-LM are updated every two months. These updates occur more frequently than updates in most rules-based systems. New mutations that effect phenotypic susceptibility may be added to the model every two months.

fig.3



## 4. The virco<sup>®</sup>TYPE HIV-1 report

The introduction of VPT-LM as the method to make FC predictions for the virco<sup>®</sup>TYPE HIV-1 assay v4.0.00 will have some consequences for the virco<sup>®</sup>TYPE HIV-1 report layout.

### 4.1 DATABASE MATCHES

The column 'database matches' on the second page of the report will be removed since VPT-LM makes FC predictions by adding up the contribution of individual mutations, rather than grouping viruses with similar Genotypes (database matches) and averaging the FC values. The accuracy of the FC prediction is given in the 95% confidence interval.

### 4.2 CUT-OFFS

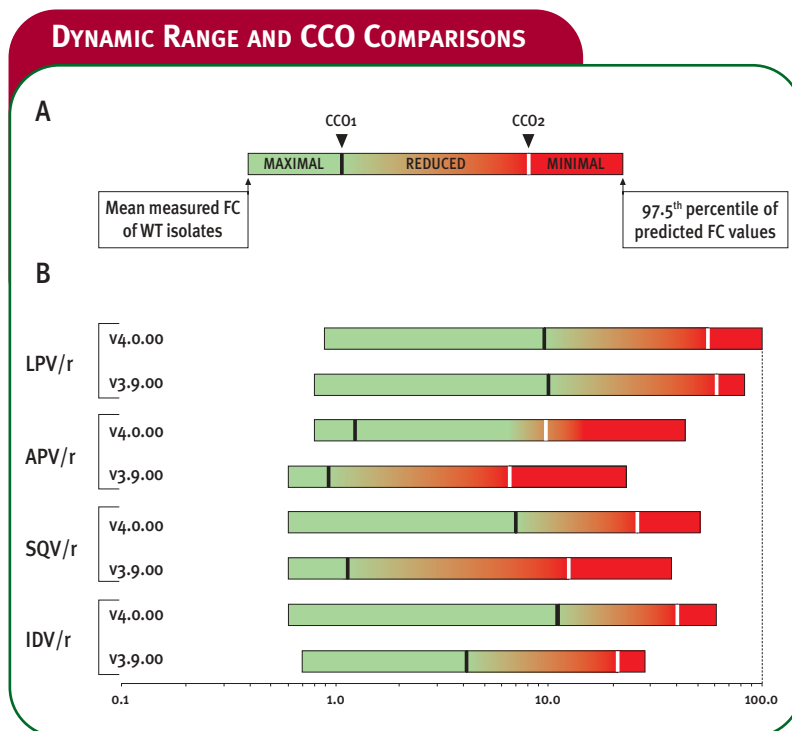
The Biological and Clinical Cut-Offs on the virco<sup>®</sup>TYPE HIV-1 report will be updated to reflect the new FC prediction methodology. Table 1 compares the Cut-Offs for virco<sup>®</sup>TYPE HIV-1 v4.0.00, with the Cut-Offs included on the previous report version 3.9.00.

	ASSAY VERSION					
	VPT (virco <sup>®</sup> TYPE HIV-1 v3.9.00)			VPT-LM (virco <sup>®</sup> TYPE HIV-1 v4.0.00)		
	BCO	CCO1	CCO2	BCO	CCO1	CCO2
AZT		1.9	14.4		1.2	9.6
3TC		1.1	3.7		1.0	3.4
ddI		1.3	3.0		0.9	2.6
d4T		1.1	2.2		0.9	2.0
ABC	2.1				0.8	1.9
FTC	3.7			3.5		
TDF		1.0	2.0		0.9	2.1
NVP	5.2			5.5		
EFV	3.4			3.4		
IDV		0.8	2.2		0.9	4.5
IDV/r		4.1	21.2		10.6	40.1
NFV		1.0	1.5		1.3	7.3
SQV/r		1.1	12.0		7.1	26.5
APV		0.7	1.4		0.9	2.0
APV/r		0.9	6.5		1.2	9.6
FAPV	1.8			2.2		
LPV/r		10.0	61.6		9.7	56.1
ATV	2.0			2.4		
TPV	1.6				1.2	5.4

Table 1: Biological and Clinical Cut-Off values according to first generation VPT (virco<sup>®</sup>TYPE HIV-1 v3.9.00) and VPT-LM (virco<sup>®</sup>TYPE HIV-1 v4.0.00)

fig.4

- With VPT-LM, the Biological Cut-Off (BCO) for ABC and TPV will be replaced with CCO values.
- The remaining BCO (for FTC, FAPV, ATV and the NNRTI) change only modestly.
- Although the definition and the approach to calculating the CCO does not change, the values themselves do change. One reason for this difference is that the dynamic range of FC predictions increases with VPT-LM and better reflects the measured Fold Changes reported by the Antivirogram<sup>®</sup>. As shown in Fig 4, the FC dynamic range increases for all boosted PIs explaining the increase in CCO values. In addition, the new CCOs for SQV/r are based on the currently approved Invirase<sup>®</sup> dose and formulation. CCOs on the previous version, v3.9.00, were derived from a combination of data collected from regimens that included either Fortovase<sup>®</sup> or Invirase<sup>®</sup>.



Dynamic Ranges and CCO for boosted PI according to first generation VPT (virco<sup>®</sup>TYPE HIV-1 v3.9.00) and VPT-LM (virco<sup>®</sup>TYPE HIV-1 v4.0.00)

Despite these changes to the dynamic ranges and the Cut-Off values, the net effect on the Resistance Analysis Calls (Maximal Response, Reduced Response, Minimal Response) is modest. A set of 56,218 samples submitted to Virco for routine clinical analysis in 2004 and 2005 was analyzed by both the VPT and the VPT-LM methodology. The respective Cut-Offs for each assay were used to make Resistance Analysis Calls for all drugs. Overall, among 1,267,970 comparisons, major differences<sup>(2)</sup> were observed in only 1.38 % of all calls.

Table 2 lists the percent major differences per drug.

Drug	% Major Differences
AZT	0.02
3TC	0.02
ddI	0.11
d4T	0.10
ABC	N/A
FTC	2.05
TDF	0.29
<i>NVP</i>	0.96
<i>EFV</i>	2.28
IDV	0.06
IDV/r	1.87
NFV	0.49
SQV/r	0.85
APV	0.14
APV/r	0.02
FAPV	2.45
LPV/r	0.88
ATV	2.96
TPV	n/a

<sup>(2)</sup> 'major difference' is defined as a change from Maximal Response to Minimal Response, or vice versa, for drugs with CCO, and from Susceptible to Resistant or vice versa for drugs with BCO

Table 2: Major Differences in Resistance Analysis Calls between VPT (virco<sup>®</sup>TYPE HIV-1 v3.9.00) and VPT-LM (virco<sup>®</sup>TYPE HIV-1 v4.0.00). For drugs marked in *italics* the comparisons are based on BCO. ABC and TPV were not analyzed because VPT uses BCO and VPT-LM uses CCO.

---

## 5. Conclusions

*VirtualPhenotype*<sup>™</sup>- LM is a new technology to make FC predictions from a viral Genotype, and will be incorporated in *virco*<sup>®</sup>TYPE HIV-1 from assay version 4.0.00 onwards.

Compared to the first generation VPT or to rules-based Genotype interpretation algorithms, VPT-LM has the following benefits:

- It defines the weighted contribution of each mutation to the FC.
- It makes more accurate predictions of the FC than the first generation VPT.
- The assessment of regimen activity based on *virco*<sup>®</sup>TYPE HIV-1 v4.0.00 is more strongly associated with treatment response than assessments based on previous assay versions.
- It can make FC predictions for every virus, even when the Genotype/Phenotype correlative database does not contain sufficient ‘matches’ of viruses with a very similar genotypic profile. It therefore eliminates the need for rules-based calls in the *virco*<sup>®</sup>TYPE HIV-1 report.
- It takes into account mutation interactions.
- It is fully data-driven.

