

# Modelling Treatment Response Could Reduce Virological Failure in Different Patient Populations

Andrew D Revell<sup>1\*</sup>, Dechao Wang<sup>1</sup>, Gabriella d'Ettorre<sup>2</sup>, Frank DE Wolf<sup>3</sup>, Brian Gazzard<sup>4</sup>, Giancarlo Ceccarelli<sup>5</sup>, Jose Gatell<sup>6</sup>, María Jesús Pérez-elías<sup>7</sup>, Vincenzo Vullo<sup>8</sup>, Julio S Montaner<sup>9</sup>, H Clifford Lane<sup>10</sup> and Brendan A Larder<sup>1</sup>

<sup>1</sup>The HIV Resistance Response Database Initiative (RDI), London, UK

<sup>2</sup>Policlinico Umberto I, Rome, Italy

<sup>3</sup>Netherlands HIV Monitoring Foundation, Amsterdam, The Netherlands

<sup>4</sup>Chelsea and Westminster Hospital, London, UK

<sup>5</sup>Italian Red Cross, Rome, Italy

<sup>6</sup>Hospital Clinic of Barcelona, Barcelona, Spain

<sup>7</sup>Ramón y Cajal Hospital, Madrid, Spain

<sup>8</sup>University of Rome Sapienza, Rome, Italy

<sup>9</sup>BC Centre for Excellence in HIV/AIDS, Vancouver, Canada

<sup>10</sup>National Institute of Allergy and Infectious Diseases, Bethesda, MD, USA

## Abstract

**Background:** HIV drug resistance can cause viral re-bounce in patients on combination antiretroviral therapy, requiring a change in therapy to re-establish virological control. The RDI has developed computational models that predict response to combination therapy based on the viral genotype, viral load, CD4 count and treatment history. Here we compare two sets of models developed with different levels of treatment history information and test their generalisability to new patient populations.

**Methods:** Two sets of five random forest models were trained to predict the probability of virological response (follow-up viral load <50 copies/ml viral RNA) following a change in antiretroviral therapy using the baseline viral load, CD4 count, genotype and treatment history from 7,263 treatment change episodes. One set used six treatment history variables and the other 18 - one for each drug. The accuracy of the models was assessed in terms of the area under the receiver-operator characteristic curve (AUC) during cross validation and with 375 TCEs from clinics that had not contributed data to the training set.

**Results:** The mean AUC achieved by the two sets of models during cross validation was 0.815 and 0.820. Mean overall accuracy was 75% and 76%, sensitivity 64% and 62% and specificity 81% and 84%. The AUC for each committee tested with the independent test set was 0.87 and 0.855. Mean overall accuracy was 89% and 87%, sensitivity 67% and 61% and specificity 90% and 87%. There were no significant differences between the two sets.

The models correctly predicted 330 (92%) of the 357 treatment failures observed in practice and were able to identify alternative regimens that were predicted to be effective for up to 267 (75%) of the failures and regimens with a higher probability of response for all cases.

**Conclusions:** Computational models can predict accurately the virological response to antiretroviral therapy from a range of variables including genotype and treatment history even for patients from unfamiliar settings. This approach has potential utility as a useful aid to treatment decision-making and may reduce treatment failure.

**Keywords:** Antiretroviral therapy; Computer models; HIV drug resistance; Genotype; Treatment history; Predictions; Treatment outcome

## Introduction

The long-term suppression of HIV replication and resulting dramatic improvements in clinical outcome resulting from combination antiretroviral therapy (cART) is a major success story. Nevertheless this requires potentially life-long therapy and the careful selection and sequencing of drugs, particularly to re-establish viral suppression following virological failure, which often occurs with the emergence of HIV drug resistance. When treatment fails in well-resourced settings a genotypic resistance test is routinely performed to identify any resistance-associated mutations [1]. The results are typically interpreted using one of the many rules-based interpretation systems available via the Internet [2]. These indicate whether the patient's virus is likely to be sensitive or resistant to each drug but do not directly provide any indication of the relative antiviral effects of combinations of drugs. With 25 or more drugs available for use in combination and more than a hundred mutations involved in drug resistance, the selection of the optimum new regimen can be challenging.

The RDI was established as a not-for-profit, global collaboration in 2002 to collect sufficient data from clinical practice to make it possible to model accurately the virological response to ART, as a treatment support tool [3]. Artificial neural network models were developed to predict virological response from genotype, viral load and CD4 count [4,5]. Limited treatment history information was added in an attempt to take into account the potential for minority populations of virus with drug resistance resulting from previous rounds of therapy that are present at levels too low for detection by population sequencing.

**\*Corresponding author:** Andrew D. Revell, Ph.D, The HIV Resistance Response Database Initiative (RDI), London, UK, Tel: +44 207 226 7314; Fax: +44 207 226 7314; E-Mail: [andrewrevell@hivrdi.org](mailto:andrewrevell@hivrdi.org)

Received May 22, 2012; Accepted July 05, 2012; Published July 07, 2012

**Citation:** Revell AD, Wang D, d'Ettorre GD, Wolf FDE, Gazzard B, et al (2012) Modelling Treatment Response Could Reduce Virological Failure in Different Patient Populations. J AIDS Clinic Res S5: 008. doi:10.4172/2155-6113.S5-008

**Copyright:** © 2012 Revell AD, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

This was found to improve accuracy [5,6]. Recent models have been demonstrated to predict antiretroviral treatment response with 80% accuracy [7,8]. This compares with a 60-70% predictive accuracy for genotypic sensitivity scores derived from genotyping with rules-based interpretation [8,9].

Random Forest (RF) models are being used to power an experimental web-based HIV treatment response prediction system (HIV-TRePS). Two clinical pilot studies involving experienced HIV physicians demonstrated that this system is a useful aid to clinical practice [10]. One-third of treatment decisions were revised based on the system's predictions and the revised regimens were predicted to produce significantly greater virological responses and involve fewer drugs in the new regimen.

An alternative system for predicting short-term treatment responses, using a combination of three different computational models trained with a European dataset, has also been evaluated and shown to be comparable to estimates of response provided by HIV physicians [11].

The development of the most accurate computational models possible can involve optimising the number of input variables on which the models base their predictions: too few and potentially useful information may be missing that could contribute to the accuracy of the models, too many and accuracy may suffer unless the training data set is sufficiently large [12]. Early models were developed using a limited set of historical variables that studies had shown to be the most influential [5,6].

Previous studies have indicated that models are more accurate in their predictions of response for independent data from patients treated in 'familiar' settings - the clinics that contributed data to the training data set, than from 'unfamiliar' settings [5].

In this study we describe the development and comparison of models developed with limited and comprehensive treatment history variables and their testing with an independent data set from unfamiliar settings.

## Methods

The RDI database currently holds anonymised, longitudinal data from approximately 90,000 patients from more than 30 clinics, cohorts and studies around the world. In order to train models to predict virological response to treatment, data are extracted from the database that relate to a change in antiretroviral therapy. The complete package of data relating to that change is termed a treatment change episode (TCE).

For the current study, TCEs were extracted that had all the following data available to be used as input variables during model development (Figure 1): Baseline plasma viral load on therapy ( $\log_{10}$  copies of HIV RNA/ml; sample taken  $\leq 8$  weeks prior to treatment change); CD4 cell count on therapy (cells/ml; sample taken  $\leq 12$  weeks prior to treatment change); baseline genotype on therapy ( $\leq 12$  weeks prior to treatment change); drugs in previous (baseline) regimen; drugs in antiretroviral treatment history; drugs in the new regimen; time to follow-up (number of days, between four and 48 weeks following introduction of the new regimen); and follow-up viral load.

There were 18 drugs included as binary variables; present=1, not present=0): zidovudine, didanosine, stavudine, abacavir, lamivudine/emtricitabine, tenofovir DF, efavirenz, nevirapine, etravirine, indinavir, nelfinavir, saquinavir, (fos) amprenavir, lopinavir, atazanavir, darunavir,

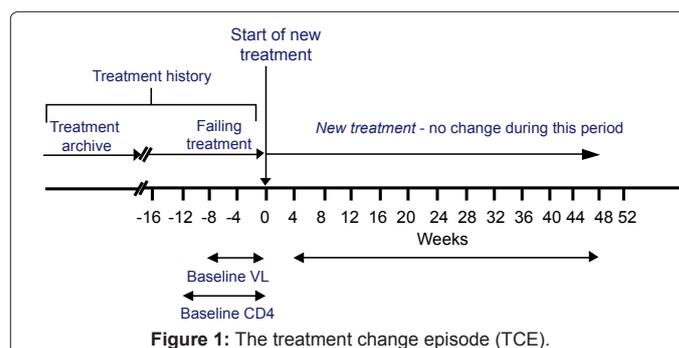
enfuvirtide, raltegravir. Maraviroc and tipranavir were not included in the models as there were insufficient follow-up data for these inhibitors in the RDI dataset. The following 62 mutations from the baseline genotype were selected from previous studies and published lists (refs) to be used as binary variables in the modelling. HIV reverse transcriptase mutations (33): M41L, E44D, A62V, K65R, D67N, 69 insert, T69D/N, K70R, L74V, V75I, F77L, V90I, A98G, L100I, L101I/E/P, K103N, V106A/M, V106I, V108I, Y115F, F116Y, V118I, I38A/G/K, Q151M, V179D/F/T, Y181C/I/V, M184V, Y188C/L/H, G190S/A, L210W, T215F/Y, K219Q/E, P236L; HIV protease mutations (29): L10F/I/R/V, V11I, K20M/R, L24I, D30N, V32I, L33F, M36I, M36L/V, M46I/L, I47V, G48V, I50V, I50L, F53L, I54 (any change), 58E, L63P, A71 (any change), G73 (any change), T74P, L76V, V77I, V82A/F/S, V82T, I84V/A/C, N88D/S, L89V, L90M)

The TCEs identified with complete data were censored using the following rules established in previous studies: no more than 3 TCEs from the same change of therapy (using multiple follow-up viral loads) were used (all with viral load determinations  $\geq 4$  weeks apart); TCEs involving drugs no longer in current use either at baseline or in the new regimen (e.g. ddC, delavirdine, loviride, emivirine, capravirine, atevirdine and adefovir) were excluded; TCEs involving drugs not adequately represented in the database (tipranavir and maraviroc) were excluded; TCEs that include an unboosted protease inhibitor (PI) other than nelfinavir, or ritonavir as the only PI, in the baseline or new regimen positions were excluded; TCEs with viral load values of the form ' $<X$ ' where X is  $>50$  or 1.7 log copies were excluded as the absolute values were not known.

The 7,263 qualifying TCEs were used to train two committees each of 5 RF models to predict the probability of the follow-up viral load being less than 50 copies/ml, using methodology described in detail elsewhere [7,8]. The first committee used just 6 simplified treatment history variables, which have been identified in previous studies as those with most influence on the accuracy of the models (any exposure to zidovudine, lamivudine/emtricitabine, enfuvirtide, raltegravir, any non-nucleoside reverse transcriptase inhibitor or any protease inhibitor). The second committee used individual treatment history variables for each of the 18 drugs covered by the system. The total number of input variables was 89 for the simple treatment history models and 101 for the individual treatment history models.

The output variable was the follow-up viral load coded as a binary variable:  $\leq 1.7$  log or 50 copies/ml=1 (response) and  $>1.7$  log or 50 copies/ml=0 (failure). The models were trained to produce an estimate of the probability of the follow-up viral load being  $<50$  copies/ml.

The two committees of 5 RF models were developed using a 5 x cross validation scheme whereby 20% of the TCEs were selected at random and the remainder used to train numerous models and their



performance gauged by cross validation with the 20% that had been 'left out'. Model development continued until further models failed to yield improved accuracy. This process was followed through five iterations until all the TCEs had appeared in a validation set once. The best performing RF model was selected from all those developed using each partition to be included in the final committee of models.

A dataset of 375 TCEs from clinics not included in the training data was obtained from the Stanford TCE repository ([www.hivdb.stanford.edu/](http://www.hivdb.stanford.edu/)) and set aside as an independent test set [13].

Receiver-operator characteristic (ROC) curves were plotted using the actual virological responses observed in the clinic at follow-up versus the predictions of the models. The performance of the models as predictors of response was evaluated in terms of the area under the ROC curve (AUC) as the primary measure plus the sensitivity and the specificity, using the optimum operating point from the cross validation as the cut-off for classifying the models' outputs as predicting virological response or failure. This was done for each of the five models in the two committees, during cross validation and with each committee as a whole (using the committee average prediction for each TCE) using the test set of 375 TCEs.

The accuracy of prediction of the models was compared to that of the following three rules-based genotype interpretation systems that are in common use as a tool to assist in the selection of effective drug combinations following virological failure: Stanford University's HIVdb system; Agence Nationale de Recherches sur le SIDA (ANRS) and REGA, accessed on 20<sup>th</sup> June 2012 via the Stanford University HIV Drug Resistance Database web site ([hivdb.stanford.edu](http://hivdb.stanford.edu)). For HIVdb the total mutation score was used as a predictor of response and for ANRS and REGA the total genotypic sensitivity scores were used.

Finally the combination of all ten RF models was used to identify potentially effective alternative regimens, using no more drugs than those in the regimen actually used in the clinic for the 357 cases of treatment failures. This was achieved by providing the models with all the baseline data for the cases and then obtaining predictions of the probability of response for alternative regimens commonly used in clinical practice. Those regimens

	Training set	Test set
<b>TCEs</b>	7,263	375
Mean (median) baseline VL (log <sub>10</sub> c/mL)	4.20 (4.11)	4.34 (4.30)
Mean (median) baseline CD4(cells/μL)	269 (230)	277 (235)
<b>Treatment History</b>		
Mean (median) number of previous drugs	5.82 (5)	5.56 (5)
NRTI experience	99%	100%
NNRTI experience	65%	54%
PI experience	75%	79%
Enfuvirtide experience	6%	0%
Raltegravir experience	1%	1%
<b>Mean (median) resistance mutations</b>	8.48 (6)	8.34 (8)
<b>Responders</b> (<1.7log <sub>10</sub> c/mL follow-up VL)	2,550 (35%)	18 (5%)
<b>Failures</b> (≥ 1.7log <sub>10</sub> c/mL follow-up VL)	4,713 (65%)	357 (95%)

TCEs=treatment change episodes. VL=viral load. NRTI=nucleoside reverse transcriptase inhibitor. NNRTI=non-nucleoside reverse transcriptase inhibitor. PI=protease inhibitor

Table 1: Data description.

Model	AUC*	Sensitivity (%)	Specificity (%)
<b>Simple treatment history</b>			
<b>Cross validation</b>			
1	0.829	63.64	84.23
2	0.824	67.51	81.68
3	0.827	65.24	80.43
4	0.784	61.21	79.60
5	0.813	64.08	80.72
<b>Mean</b>	0.815	64.34	81.33
<b>Committee average performance with 375 test TCEs</b>	0.870	66.67	90.48
95% CI	0.87-0.97	41-87	87-93
<b>Individual treatment history</b>			
<b>Cross validation</b>			
1	0.837	64.86	85.06
2	0.821	61.18	84.42
3	0.834	62.52	86.56
4	0.798	60.43	83.49
5	0.811	61.55	81.26
<b>Mean</b>	0.820	62.11	84.16
<b>Committee average performance with 375 test TCEs</b>	0.855	61.11	87.47
95% CI	0.76-0.95	36-83	85-92
<b>Results from common rules-based genotype interpretation systems**</b>			
<b>Stanford HIVdb</b>	0.591	47.06	41.26
95% CI	0.49-0.69	[23, 72]	[36, 47]
<b>ANRS v2011.05</b>	0.573	23.53	61.03
95% CI	0.45-0.69	[7, 50]	[56, 66]
<b>REGA v8.0.2</b>	0.566	35.29	54.44
95% CI	0.44-0.69	[14, 62]	[49, 60]

\*AUC=area under the (receiver-operator characteristic) curve.

\*\* Based on 366 TCEs after those containing maraviroc or raltegravir were removed

Table 2: Performance of the models during cross validation and independent testing.

with a probability of response above the optimum operating point for the models as a prediction system, derived during cross validation, were deemed as predicted to be effective.

## Results

### Characteristics of the datasets

The characteristics of the datasets are summarised in Table 1. The training and test set were comparable in terms of baseline characteristics with mean viral loads of 4.2 and 4.34 log<sub>10</sub> copies viral RNA/ml respectively (median of 4.11 and 4.30) and mean CD4 counts of 269 and 277 (median of 230 and 235). Both populations were heavily pre-treated with a mean of 5.82 and 5.56 previous drugs (median of five in both cases), and both had significant drug resistance with a mean of 8.48 and 8.34 resistance mutations (median of 6 in both cases). The main difference between the populations was the proportion of responders (defined as a follow-up viral load of <50 copies HIV RNA/ml): 35% of the training set and just 5% of the independent test set.

### Results of the modelling

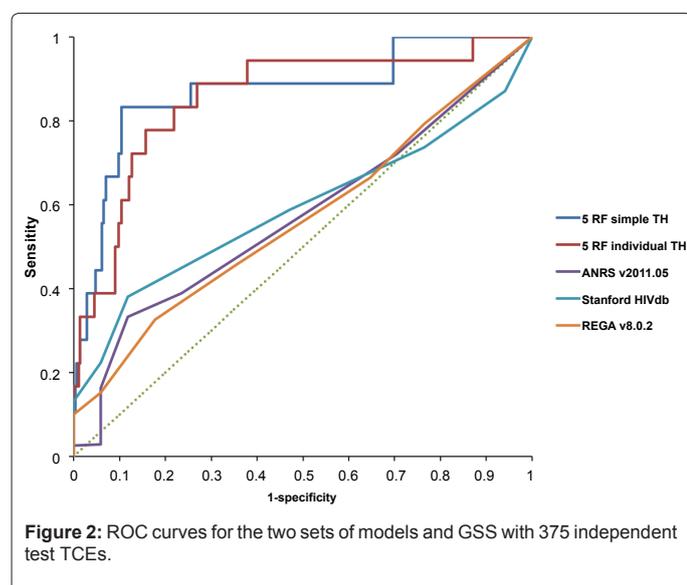
**Cross validation:** The performance characteristics from the ROC curves of the 5 individual models in each of the two committees during

cross-validation and the performance of the two committees with the independent test set are summarized in Table 2. The AUC values achieved by the models using simple treatment history variables during cross-validation ranged from 0.784 to 0.829, with a mean of 0.815. The overall accuracy ranged from 73.16 to 76.97% (mean=75.35%), the sensitivity ranged from 61.21 to 67.51% (mean=64.34%) and the specificity from 79.60 to 84.23% (mean=81.33%).

The AUC values achieved by the models using 18 individual treatment history variables during cross-validation were approximately one percent better, ranging from 0.798 to 0.837, with a mean of 0.820. The overall accuracy ranged from 74.53 to 77.99% (mean=76.43%), the sensitivity ranged from 60.43 to 64.86% (mean=62.11%) and the specificity from 81.26 to 86.56% (mean=84.16%). There were no significant differences between the performances of the two committees using DeLong's test.

**Testing the two committees with the independent set of 375 TCEs:** The ROC curves for the committee average performance of the two committees are presented in Figure 2. The simple treatment history committee achieved an AUC of 0.870. The sensitivity was 66.67% and the specificity 90.48% using the optimum operating point (OOP, the value that when used as a cut-off maximizes sensitivity and specificity) of 0.43. The individual treatment history committee achieved an AUC of 0.855. The sensitivity was 61.11% and the specificity 87.47%, using the OOP of 0.46. Again there were no significant differences between the performance of the two committees using DeLong's test.

The Stanford, ANRS and REGA genotype interpretation systems achieved AUC values of 0.591, 0.573 and 0.566 respectively (Table 2). These values were significantly lower than that of the RDI models.



**Identifying potentially effective alternative regimens:** Since the performance of the two committees of 5 RF models was not significantly different during cross validation and independent testing, both were used to identify potentially effective regimens for the 357 virological failures in the test set of 375 TCEs. The results are presented in Table 3. The models correctly predicted 330 (92%) of the 357 failures observed in practice. They were able to identify alternative regimens that were predicted to be effective for 267 (75%) of the 357 failures, using no more drugs than were used in the clinic. They were able to identify alternative regimens with a higher probability of response than the regimen used in the clinic for all 357 cases of failure.

The data in the test set are from 1997-2010, with 85% at least 10 years old, so some of the treatment decisions made in the clinic will have been made before some of the more recent drugs became available. A frequency analysis of the data indicated that the following drugs, approved for use during the last 10 years, were only infrequently used in the dataset: atazanavir (0 TCEs), etravirine (2 TCEs) enfuvirtide (3), darunavir (5), raltegravir (6). The analysis was repeated without these five drugs. The models were still able to identify alternative regimens that were predicted to be effective for 146 (41%) of the 357 failures, again using no more drugs than were used in the clinic. They were able to identify alternative regimens with a higher probability of response than the regimen used in the clinic for 351 (98%) of the 357 treatment failures.

A further analysis was performed whereby only those drugs that had been approved by the FDA at the time of the treatment decision were included in the analysis for each test TCE. The models were able to identify alternative regimens that were predicted to be effective for 143 (40%) of the 357 failures and regimens with a higher probability of response than the regimen used for 356 (100%) of the failures.

## Discussion

The two sets of models developed in this study performed comparably well as predictors of virological response to antiretroviral therapy. The marginal numerical superiority of the individual treatment history models during cross-validation and of the simple treatment history models during independent testing did not achieve statistical significance. The AUC values of 0.82-0.87 compare very favourably with the values of 0.57-0.59 typically achieved by genotyping with rules-based interpretation as a predictor of outcome indicating that this approach offers additional utility as an aid to treatment selection over the use of genotyping alone.

The performance of the models with an independent test set from 'unfamiliar' clinics was particularly impressive and suggests a high degree of generalizability of the models. A common problem with computational models developed for this sort of task is that during development they find the best 'local solution', i.e. best algorithm for the training data, which may not be generalizable to other datasets. RF models were selected as being less prone to this effect than

Drugs included in the modelling of alternative regimens	Number (%) of actual failures for which the models identified alternative regimens predicted to be effective	Number (%) of actual failures for which the models identified alternative regimens with a higher probability of success
18 drugs (no tipranavir or maraviroc)	267 (75%)	357 (100%)
13 drugs (no tipranavir, maraviroc, lopinavir, atazanavir, darunavir/raltegravir, enfuvirtide): drugs approved in last 10 years)	146 (41%)	351 (98%)
Only those drugs available at the time of the TCE included	143 (40%)	356 (100%)

Table 3: Using the models to identify effective alternatives to regimens that failed.

other modelling methodologies and these results are particularly encouraging in the degree of generalisability. The models significantly out-performed three genotype interpretation systems in common use as aids to treatment selection, which is also encouraging in terms of the potential utility of the system as a clinical aid.

It is interesting to note that the specificity achieved by the models in predicting response was particularly high at 80-90%. This is encouraging in that it minimizes the chances of a false positive - a prediction that a regimen will be effective and then it fails. The test data were largely cases of treatment failure among heavily pre-treated patients. The models correctly predicted 92% of these failures. This compares very favourably with the three genotype interpretation systems, which only correctly predicted 41-61% of the failures. Moreover the models were able to identify alternative, practical regimens that were predicted to produce a virological response for a substantial proportion of the failures observed, even without any of the drugs approved in the last 10 years. This suggests that the models could have considerable utility in salvage and in settings with restricted access to drugs.

The study has some limitations. Firstly, it is a retrospective study and a prospective controlled clinical trial would be required to validate the models in terms of clinical benefit. The training and test data are all from well-resourced settings (North America, Western Europe, Australia and Japan). The models performance and findings may not be as generalizable to other settings. A similar point can be made about the clade of the virus, which will have mostly been B in the cases used in this study, although preliminary studies testing our models with non-B virus indicates that clade may not be a major factor [14,15].

## Conclusions

This study demonstrates that computational models can be developed to predict accurately the virological response to antiretroviral therapy from a range of variables including genotype, treatment history, viral load and CD4 count. This performance is not specific to the settings that provided the training data but can be generalizable to unfamiliar, but similar settings. The models were able to identify potentially effective alternative combinations of drugs for at least 40% of the failures occurring in these unfamiliar clinics, using only drugs that were available at the time, which are now more than 10 years old in the great majority of cases. This approach may therefore have significant utility as a useful aid to treatment decision-making in settings with limited access to more recent drugs and may reduce treatment failure. An experimental online tool, the HIV Treatment Response Prediction System (HIV-TRePS) powered by such models is available for use as a free experimental clinical tool via [www.hivrdi.org](http://www.hivrdi.org).

## Acknowledgements

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E. This research was supported by the National Institute of Allergy and Infectious Diseases. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

## RDI data and study group

The RDI wishes to thank all the following individuals and institutions for providing the data used in training and testing its models:

## Cohorts

Frank De Wolf and Joep Lange (ATHENA, the Netherlands); Julio Montaner and Richard Harrigan (BC Center for Excellence in HIV & AIDS, Canada); Tobias Rinke de Wit and Raph Hamers (PASER-M cohort, The Netherlands); Brian Agan;

Vincent Marconi and Scott Wegner (US Department of Defense); Wataru Sugiura (National Institute of Health, Japan); Maurizio Zazzi (MASTER, Italy).

## Clinics

Jose Gatell and Elisa Lazzari (University Hospital, Barcelona, Spain); Brian Gazzard, Mark Nelson, Anton Pozniak and Sundhiya Mandalia (Chelsea and Westminster Hospital, London, UK); Lidia Ruiz and Bonaventura Clotet (Fundacion IrsiCaixa, Badelona, Spain); Schlomo Staszewski (Hospital of the Johann Wolfgang Goethe-University, Frankfurt, Germany); Carlo Torti (University of Brescia); Cliff Lane and Julie Metcalf (National Institutes of Health Clinic, Rockville, USA); Maria-Jesus Perez-Elias (Instituto Ramón y Cajal de Investigación Sanitaria, Madrid, Spain); Andrew Carr, Richard Norris and Karl Hesse (Immunology B Ambulatory Care Service, St. Vincent's Hospital, Sydney, NSW, Australia); Dr Emanuel Vlahakis (Taylor's Square Private Clinic, Darlinghurst, NSW, Australia); Hugo Tempelman and Roos Barth (Ndlovu Care Group, Elandsdoorn, South Africa), Carl Morrow and Robin Wood (Desmond Tutu HIV Centre, Cape Town, South Africa); Luminita Ene ("Dr. Victor Babes" Hospital for Infectious and Tropical Diseases, Bucharest, Romania).

## Clinical trials

Sean Emery and David Cooper (CREST); Carlo Torti (GenPhereX); John Baxter (GART, MDR); Laura Monno and Carlo Torti (PhenGen); Jose Gatell and Bonventura Clotet (HAVANA); Gaston Picchio and Marie-Pierre deBethune (DUET 1 & 2 and POWER 3); Maria-Jesus Perez-Elias (RealVirfen).

## References

1. Thompson MA, Aberg JA, Cahn P, Montaner JS, Rizzardini G, et al. (2010) Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society-USA panel. *JAMA* 304: 321-333.
2. Liu TF, Shafer RW (2006) Web Resources for HIV type 1 Genotypic-Resistance Test Interpretation. *Clin Infect Dis* 42: 1608-1618.
3. Larder BA, DeGruttola V, Hammer S, Harrigan R, Wegner S, et al. (2002) The international HIV resistance response database initiative: a new global collaborative approach to relating viral genotype treatment to clinical outcome. *Antivir Ther* 7: S111.
4. Wang D, Larder BA, Revell A, Harrigan R, Montaner J (2003) A Neural Network Model Using Clinical Cohort Data Accurately Predicts Virological Response and Identifies Regimens with Increased Probability of Success in Treatment Failures. 12th International Workshop on HIV Drug Resistance, Cabo, Mexico.
5. Larder B, Wang D, Revell A, Montaner J, Harrigan R, et al. (2007) The development of artificial neural networks to predict virological response to combination HIV therapy. *Antivir Ther* 12: 15-24.
6. Larder BA, Wang D, Revell A, Harrigan R, Montaner J, et al. (2004) Previous Drug Exposure Data Significantly Increase the Accuracy of Artificial Neural Networks in Predicting Virological Response to Combination Therapy. 13th International Drug Resistance Workshop, Costa Adeje - Tenerife Sur, Canary Islands, Spain.
7. Wang D, Larder B, Revell A, Montaner J, Harrigan R, et al. (2009) A comparison of three computational modelling methods for the prediction of virological response to combination HIV therapy. *Artif Intell Med* 47: 63-74.
8. Revell AD, Wang D, Boyd MA, Emery S, Pozniak A, et al. (2011) The development of an expert system to predict virological response to HIV therapy as part of an online treatment support tool. *AIDS* 25: 1855-1863.
9. Montaner J, Larder BA, Wang D, Revell A, Wegner S, et al. (2006) Neural Networks Are More Accurate Predictors of Virologic Response to Antiretroviral Therapy Than Rules-Based Genotype Interpretation Systems. CROI, Denver, CO, USA.
10. Larder BA, Revell AD, Mican JM, Agan BK, Harris M, et al. (2011) Clinical Evaluation of the Potential Utility of Computational Modeling as an HIV Treatment Selection Tool by Physicians with Considerable HIV Experience. *AIDS Patient Care and STDS* 25: 29-36.
11. Zazzi M, Kaiser R, Sönnnerborg A, Struck D, Altmann A, et al. (2011) Prediction of response to antiretroviral therapy by human experts and by the EuResist data-driven expert system (the EVE study). *HIV Med* 12: 211-218.
12. DiRienzo G, DeGruttola V (2002) Collaborative HIV resistance-response database: sample size for detection of relationships between HIV-1 genotype and HIV-1 RNA response using a non-parametric approach. *Antivir Ther* 7: S71.

13. Rhee S-Y, Blanco JL, Liu TF, Pere I, Kaiser R, et al. (2012) Standardized representation, visualization and searchable repository of antiretroviral treatment-change episodes. *AIDS Res Ther* 9: 13
14. Larder BA, Revell AD, Wang D, Hamers R, Tempelman H, et al. (2011) Modelling response to antiretroviral therapy without a genotype as a clinical tool for resource-limited settings. *Antivir Ther* 16 (Suppl 1): A42
15. Revell AD, Ene L, Duiculescu D, Wang D, Youle M, et al. (2012) The use of computational models to predict response to HIV therapy for clinical cases in Romania. *GERMS* 2: 6-11