Andrew Revell

Putting the experience of treating thousands of patients at every doctor's fingertips

The HIV Resistance Response Database Initiative (RDI), London, UK

Corresponding author: Andrew Revell, Executive Director
RDI HIV Resistance, 14 Union Square, London, N1 7DH UK
e-mail: andrewrevell@hivrdi.org

Twenty-five antiretroviral drugs available from six classes: the pace of the medical response to a virus that was only discovered 27 years ago has been truly remarkable. While combination antiretroviral therapy can now be expected to suppress the virus and prevent or halt disease progression for many years, treatment continues to fail in some individuals, often as a result of lapses in adherence and the development of drug resistance. When this occurs, the therapy has to be changed. Re-suppressing the virus and maintaining it at low levels for the lifetime of a patient requires careful choices to overcome drug resistance, stay one step ahead of viral evolution and minimise toxicity. Faced by the sheer complexity of resistance, the number of potential drug combinations available and the competing clinical and economic considerations affecting the treatment decision, successful individualised sequencing of antiretroviral therapy is highly challenging. These challenges are multiplied in resource-limited settings where there are fewer diagnostics, drugs and experienced HIV specialists to go round. This article describes a new approach: harnessing the clinical experiences of thousands of patients to power individualised and optimised treatment decision-making.

For many years the standard procedure in well-resourced health care settings has been to monitor the HIV patient's viral load regularly, with detection of persistent viraemia triggering a change of antiretroviral drug therapy.[1,2] A genotypic resistance test is usually ordered to identify any drug resistance mutations. The interpretation of this genotype can be complex and is usually performed using rules-based interpretation software that relates individual mutations to viral susceptibility to individual drugs.[3] However, different systems provide different interpretations, which do not always concur well.[5] It is also difficult to use the predicted susceptibility to individual drugs to predict which combination is most likely to be effective: raw genotype sensitivity scores being significant but relatively weak predictors of virological response.[6]

Bioinformatics have been used most commonly to predict the phenotype from a genotype and then relate a cut-off in predicted phenotype to a categorical response.[7] Again, it is difficult to relate this categorical prediction for an individual drug to the relative responses that may be achieved with different candidate drug combinations.

A new approach

Computational models that provide a *quantitative* prediction of virological response to combination therapy (rather than to individual drugs), *directly* from the genotype and other information would clearly be advantageous. However, this can be challenging given that a very large dataset is required to accommodate multiple possible drug-genotype permutations with their respective drug response data.[8] The HIV Resistance Response Database Initiative (RDI) was established in 2002 to take on this challenge.[9]

> *"Being able to model quantitatively the chance of a combination of drugs working always appealed: it is more useful than just getting a 'yes', 'no' or 'maybe' for each individual drug from the rules-based genotype report. The stumbling block was always the huge amount of data needed to train the models to cover the whole range of drugs, genotypes and other information and be as generalisable to different populations as possible. That's why in 2002 we set up the RDI as a wholly independent initiative: to be the single group internationally to collect the data and develop this approach - avoiding counterproductive competition for data."*
>
> *Dr Brendan Larder, Scientific Chair, RDI*

The RDI is not-for-profit research group and a global network of researchers and physicians. It has the mission to improve the clinical management of HIV infection through the application of bioinformatics to HIV drug resistance and treat-

ment outcome data. The group was founded by virologist Dr Brendan Larder, the first person to discover HIV drug resistance, and colleagues with decades of experience in the HIV diagnostic and pharmaceutical industries. Over the eight years since its inception, the RDI has worked with many of the leading clinicians and scientists in the world to develop the world's largest database of HIV drug resistance and treatment outcome data, containing information from approximately 70,000 patients in more than 15 countries spanning three continents. Over the same period, the RDI has conducted more than 30 studies to evaluate and refine its different modelling methods.

The data challenge

Following the establishment of the RDI executive and advisory groups the first task was the collection of sufficient, suitable data. This proved to be and continues to be a significant challenge. Clinics and cohorts that systematically collect all the data that might be required for the modelling were identified and approached. Important issues including patient confidentiality, data ownership, restrictions to data use, publication policy etc were raised and agreed with each centre. Each dataset was delivered in a different format and different codes for some drugs were used. For example, one drug was recorded in 20 different ways by different groups. Sophisticated data interrogation, re-coding and standardisation programs had to be written to quality control and reformat data for use.

Since the intention was to develop models that could help the selection of optimal therapy following failure, once data had been imported 'packets' of all the relevant data surrounding a change of antiretroviral therapy were extracted. These were termed Treatment Change Episodes (TCEs) by the RDI and this has now become a standard term in the field.

Refining the methods – the early years

Over the first few years the RDI developed and refined its methodology, training hundreds of different types of computational models, including artificial neural networks, random forests and support vector machines, to predict virological response to a new treatment from a range of variables including the baseline genotype and viral load.[10,11] Different criteria for the TCE were developed and tested. A number of additional variables were investigated empirically, found to enhance the accuracy of predictions and added to the RDI's modelling strategy. These included CD4 counts and antiretroviral treatment history. One of the limitations of genotyping methods in common use is that resistant virus that emerged during previous rounds of therapy may persist but at levels that are too low for population-based sequencing to detect. Treatment history information may be acting as a surrogate marker for this low-level resistant virus and was found to contribute significantly to the accuracy of predictions while previous genotype information did not.[12]

When tested with independent retrospective data, the models have consistently proved to be accurate predictors of virological response. In addition, the models are able to identify combinations of antiretroviral drugs that they predict to be effective for a substantial proportion of cases of virological failure following genotype-guided changes in therapy.[13] While these results suggested that this approach could indeed have a useful role as an aid to antiretroviral treatment decision-making in the clinic, the group felt it was imperative to subject the system to some clinical evaluation before its release.

Clinical evaluation

In 2006, a prototype web-based interface was developed to allow investigators access to the models and, in 2007, two clinical pilot studies were initiated to assess the system as a clinical tool: a prospective study in Canada and Italy and a retrospective study in the USA.[14] Participating physicians entered baseline data for cases of treatment failure via the RDI website and then registered their treatment intention, based on all the laboratory and clinical information available to them. The baseline information was automatically input to the RDI models, which then made predictions to more than 200 potential alternative combinations of antiretroviral drugs. The physician then received an automated report listing the five alternative regimens that the models predicted would be most effective, plus their own treatment selection, ranked in order of predicted virological response. Having reviewed the report, the physicians then entered their final treatment decision.

Twenty-three physicians entered 114 cases. Overall 33% of treatment decisions were changed following review of the report. The final treatment decisions and the best of the RDI alternatives were predicted to produce greater virological responses and involve fewer drugs than the original selections (Table 1).

Most physicians found the system easy to use and understand. All but one indicated they would use the system if it were available, particularly for highly treatment-experienced cases with challenging resistance profiles. The first clinical evaluation of this approach, by physicians with substantial HIV-experience suggested that it has the potential to deliver clinical and economic benefits.

The launch of a ground-breaking new system

On October 4th 2010 the RDI launched the experimental system, called the HIV Treatment Response Prediction System (HIV-TRePS). It is available free of charge over the Internet and is designed to help physicians select the best treatment for their patients, on an individual basis. Physicians access the system over the Internet and enter their patient's data, and the system predicts how the patient will respond to hundreds of alternative combinations of HIV drugs. Within seconds, the physician receives a report listing the drug combinations that the models predict are most likely to work. An example of a report is included in Figure 1.

Table 1: Comparison of predicted changes in viral load from baseline

	Physician's selection	RDI system		
A: All cases (n=114)	Initial	Final	A*	B*
Mean predicted change in viral load from baseline	-1.88	-1.90	-2.03	-2.03
Median	-1.85	-1.85	-1.98	-1.98
Proportion with >2 log reduction	39%	41%	50%	50%
Statistical significance (vs physician's initial selection)**		p=0.06	p<0.0001	p<0.0001
B: Cases where the treatment decision was changed (n=38)				
Mean	-1.92	-1.99	-2.12	-2.13
Median	-1.91	-1.99	-2.06	-2.07
Proportion with >2 log reduction	39%	50%	58%	58%
Statistical significance (vs physician's initial selection)**		p<0.05	p<0.0001	p<0.0001

* A: Alternative regimens with no more drugs than the physician's initial selection
B: Alternative regimens with no more than six drugs
** One tailed t-tests

Figure 1: Sample report from HIV-TRePS system

HIV TRePS HIV Treatment Response Prediction System

rdi hiv resistance

Case summary for RDI case 382, patient ID 0001

Baseline Data Completed: 10/12/2010

Patient Details

DATE OF BIRTH	SEX	PREGNANT
20/02/1967	M	N

Viral Load

SAMPLE DATE	VIRAL LOAD VALUE
06/12/2010	10,000

CD4

SAMPLE DATE	CD4 VALUE
06/12/2010	200

Genotype

GENOTYPE SAMPLE DATE	PROTEASE MUTATIONS	RT NRTI MUTATIONS	RT NNRTI MUTATIONS
06/12/2010	10F/I/R/V,24I,36I,46I/L 54V/L/M,73S/A,82A/F/S	41L,44D,67N,184V 210W,215Y	103N

Treatment History and Drug Exclusions

PREVIOUS DRUG EXPOSURE	UNAVAILABLE DRUGS	EXCLUDED DRUGS
ABC,F/3TC,AZT,EFV,NVP,IDV/rtv SQV/rtv	None	D4T,MVC,RAL

Predictions*

HIV-TRePS (v1.0) was instructed to model responses at 24 weeks to antiretroviral regimens comprising no more than 4 drugs.

The regimens are listed in order of predicted virological response with the regimen most likely to reduce the viral load to <50 copies ranked first.
The estimated probability of response (viral load <50 copies) is listed next to each regimen.
Those regimens that the system predicts will produce a response (using the cut-off value for the probability of response that produces optimal accuracy) are highlighted in green and those predicted to fail in red.

Predictions of responses at 24 weeks for the top five regimens (containing up to 4 drugs)

RANK	DRUGS USED TO TREAT PATIENT	NO. DRUGS	PROBABILITY VL <50	RANGE OF MODELS PREDICTION	RESPONSE CATEGORY
1	DRV/rtv,ENF,F/3TC,TDF	4	69%	51 - 77%	Response
2	DRV/rtv,ETV,F/3TC,TDF	4	65%	59 - 68%	Response
3	ABC,DRV/rtv,ENF,TDF	4	65%	57 - 77%	Response
4	DRV/rtv,F/3TC,TDF	3	65%	59 - 77%	Response
5	ENF,ETV,F/3TC,TDF	4	64%	45 - 81%	Response

*** TRePS Predictions Disclaimer**

Use of the HIV-TRePS is subject to the following conditions, to which you have agreed:

1. HIV-TRePS is an on line software programme ("Programme") which enables physicians to input specific user-defined information in order to obtain predictions concerning the use of combinations of antiretroviral drugs ("Predictions").
2. The RDI does not accept any responsibility for the accuracy of the data entered by the user or the consequences of any inaccuracies in those data.
3. The Predictions provided by the Programme are produced by an experimental system that is unproven and intended for research use only.
4. Responses to HIV treatment are complex and affected by a number of factors not taken into account by the Programme.
5. The selection of drugs for the treatment of HIV infection is the responsibility of the physician in consultation with the patient and reliance should not be placed on the Predictions for such purposes.
6. The Predictions are not intended to replace professional medical care and attention by a qualified medical practitioner and consequently the RDI does not accept any responsibility for the selection of drugs, the patient's response to treatment or differences between the Predictions and patients' responses.

"This is a very exciting development – the system literally puts the experience of treating thousands of different patients at the doctor's fingertips. This has the potential to improve outcomes for people living with HIV and AIDS around the world, particularly where resources and expertise are scarce."

Dr. Julio Montaner, Past President of the International AIDS Society and Director of the BC Centre for Excellence in HIV & AIDS, Vancouver, Canada

The computational models currently powering the predictions of HIV-TRePS are called random forests and base their predictions on more than 80 variables including baseline CD4 count, viral load and HIV genotype; antiretroviral treatment history; drugs in the new regimen; time to follow-up and the follow-up viral load values. The models estimate the probability of each combination of drugs reducing the amount of virus to below the limit of detection in the blood (50 copies HIV RNA/ml) based on what the system has 'learnt' during its training with thousands of real clinical cases.

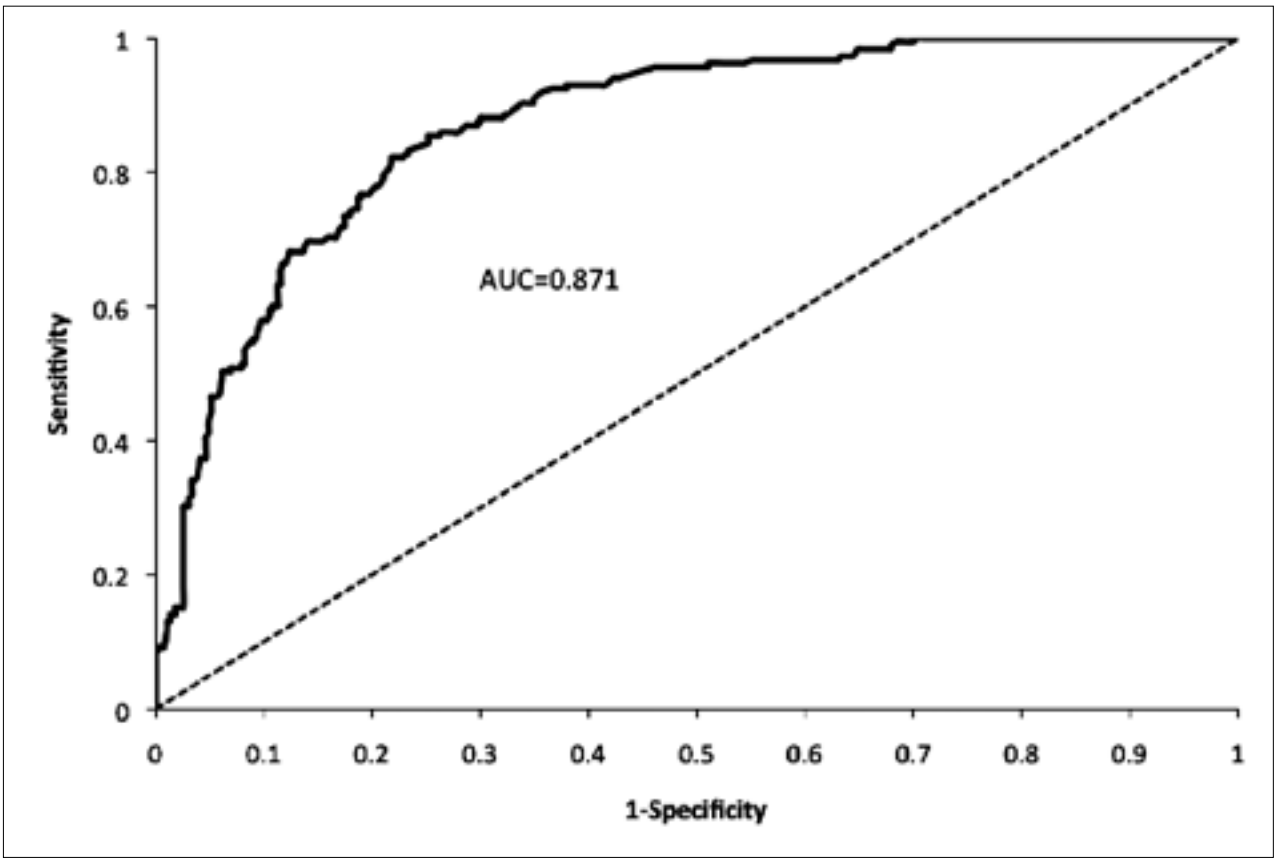
The current models were assessed during cross validation and with an independent test set of data and compared to genotypic sensitivity scores (GSS) from three rules-based genotype interpretation systems in common use to help guide treatment decisions.[15] The models achieved consistently accurate predictions – the overall accuracy was approximately 80%, which was significantly more accurate than the GSS. The receiver operator curve for the RDI models during cross validation is presented in Figure 2.

At the time of writing, HIV-TRePS has over 300 users in 54 countries around the world. But this is not the end, just the end of the beginning. The models that power HIV-TRePS have to be updated and replaced regularly by new ones that keep pace with clinical practice, particularly the development of new drugs.

"We are really excited about the launch of this system, which is a milestone for us, our research partners around the world and also for the use of bioinformatics in medicine. We believe this approach can make a significant difference in a variety of settings and diseases."

Dr Brendan Larder, Scientific Chair of the RDI.

Figure 2. ROC curve for the best-performing RF model during cross validation



The design and application of effective antiretroviral treatment over the long-term poses a particular challenge in resource-limited settings where drugs, diagnostics and sometimes expertise are in short supply. For example genotyping is prohibitively expensive in many such settings. The RDI has already developed prototype models that can predict the response to therapy without a genotype that are only a few percentage points less accurate than the models using genotypic information.[16] A version of HIV-TRePS designed for use in resource-limited settings will be launched shortly.

Finally, this approach could also have potential benefit in other diseases, most obviously where drug resistance can be a problem such as Hepatitis. As data collection in clinical practice becomes more routine and standardised so putting the experience of treating thousands of patients at the fingertips of individual physicians will become more and more of a reality.

HIV-TRePS can be accessed at:
http://www.hivrdi.org/

The RDI's International Advisory Group

- Dr Julio Montaner (BC Centre For Excellence in HIV/AIDS, Vancouver, Canada)
- Dr Carlo Torti (University of Brescia, Italy)
- Dr John Baxter (Cooper University Hospital, Camden, NJ, USA)
- Dr Sean Emery (National Centre in HIV Epidemiology and Clinical Research, Sydney, Australia)
- Dr Jose Gatell (Hospital Clinic of Barcelona, Spain)
- Dr Brian Gazzard (Chelsea and Westminster Hospital, London, UK)
- Dr Anna-Maria Geretti (Royal Free Hospital, London, UK)
- Dr Richard Harrigan (BC Centre For Excellence in HIV/AIDS, Vancouver, Canada)

Acknowledgements

Data sources
The RDI wishes to thank all the following individuals and institutions for providing the data used in training and testing these models.

Cohorts:
Frank De Wolf and Joep Lange (ATHENA, Netherlands)
Julio Montaner and Richard Harrigan (BC Center for Excellence in HIV & AIDS, Canada)
Brian Agan. Vincent Marconi and Scott Wegner (Infectious Disease Clinical Research Program, Uniformed Services University of the Health Sciences, USA)
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, Italy)
Cliff Lane and Julie Metcalf (National Institutes of Health Clinic, Rockville, USA)
Maria-Jesus Perez-Elias (Ramon y Cajal Hospital, Madrid, Spain)
Andrew Carr, Mark Boyd, Richard Norris and Karl Hesse (Immunology B Ambulatory Care Service, St. Vincent's Hospital, Sydney, Australia)
Emanuel Vlahakis (Taylor's Square Private Clinic, Darlinghurst, Australia)
Luminita Ene (Dr. Victor Babes Hospital for Infectious and Tropical Diseases, Bucharest, Romania)
Vincenzo Vullo and Gabriella D'Ettorre (Department of Public Health and Infectious Diseases Sapienza University, Rome, Italy)
Roos Barth (Internal Medicine and Infectious Diseases, UMC Utrecht, The Netherlands)
Annemarie Wensing (Dept of Virology, Medical Microbiology, UMC, Utrecht, The Netherlands)
Hugo Tempelman, (Ndlovu Care Group, Elandsdoorn, South Africa)

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)

References

[1] Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Department of Health and Human Services. January 10, 2011; 1–166

[2] The European AIDS Clinical Society. Guidelines for the Clinical Management and Treatment of HIV infected Adults in Europe Version 5.2. At: http://www.europeanaidsclinicalsociety.org/guidelines.asp (last accessed 3rd February 2011)

[3] Sturmer M, Doerr HW, Preiser W. Variety of interpretation systems for human immunodeficiency virus type 1 genotyping: confirmatory information or additional confusion? Curr Drug Targets Infect Disord 2003;3:373-382

[5] Sturmer M, Doerr HW, Staszewski S, Preiser W. Comparison of nine resistance interpretation systems for HIV-1 genotyping. Antivir Ther 2003; 8:239-244.

[6] Frentz D, Boucher CAB, Assel M, et al. Comparison of HIV-1 Genotypic Resistance Test Interpretation Systems in Predicting Virological Outcomes Over Time. PLoS One 2010;5(7): e11505 (doi:10.1371/journal.pone.0011505).

[7] Beerenwinkel N, Sing T, Lengauer T, et al. Computational models for the design of effective therapies against drug resistant HIV strains. Bioinformatics 2005 21(21):3943-3950.

[8] DiRienzo G, DeGruttola V. 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 2002;7:S71.

[9] Larder BA, DeGruttola V, Hammer S, et al. The international HIV resistance response database initiative: a new global collaborative approach to relating viral genotype treatment to clinical outcome. Antivir Ther 2002;7:S84

[10] Larder BA, Wang D, Revell A, et al. The development of artificial neural networks to predict virological response to combination HIV therapy. Antivir Ther 2007;12:15-24.

[11] Wang D, Larder BA, Revell AD et al. A comparison of three computational modelling methods for the prediction of virological response to combination HIV therapy. Artificial Intelligence in Medicine (2009) 47, 63—74

[12] Larder BA, Wang D, Revell A et al. Treatment history but not previous genotype improves the accuracy of predicting virologic response to HIV therapy**.** Abstract H-1051; 45th ICAAC; 16-19 December 2005; Washington DC, USA.

[13] Wang D, Larder BA, Revell A, Harrigan, R and Montaner, J. A neural network models using clinical cohort data accurately predicts virological response and identifies regimens with increased probability of success treatment failures, Abstract 102, VII International HIV Drug Resistance Workshop, 10-14 June 2003, Los Cabos, Mexico

[14] Larder, BA, Revell, AD, Mican, J et al. 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 2011, 25(1):29-36

[15] Larder BA, Wang D, Revell AD et al. The Development of Computational Models That Accurately Predict Virological Response to HIV Therapy to Power an Online Treatment Selection Tool. Antiviral Therapy 2010; 15 (Suppl 2):A89.