HIV RESISTANCE RESPONSE DATABASE INITIATIVE

CONCEPT SHEET

BACKGROUND

One of the main reasons for HIV treatment failure is that the virus readily develops resistance to the drugs ranged against it. HIV replication is highly error prone with no proofreading mechanism and genetic changes can make the virus less susceptible to antiretroviral drugs. Under conditions of drug therapy but incomplete viral suppression, selection pressure is exerted on the viral population within the individual patient that can lead to the emergence of drug resistant strains as the dominant viral quasi-species leading to viral rebound and ultimately clinical failure.

Resistance tests are now increasingly used to help identify drug combinations that will overcome resistance in the individual patient. There are two types of test in use: the genotype test that reads the genetic code of HIV and enables mutations in the genetic code to be identified and phenotypic tests in which part of the viral genome is spliced into a laboratory strain of the virus that can then be 'grown' in the presence of drugs in the laboratory. The type of test most commonly used, because of its relative convenience and low cost, is the genotype. Once the test has identified mutations in a patient's virus, however, this information requires interpretation to predict what effect they will have on susceptibility of that virus to each of the available drugs. This poses tremendous challenges because there are more than 200 different mutations that can affect drug resistance in complex ways, and because of the enormous number of possible drug combinations available. Generally, this interpretation is performed by using sets of "rules" or algorithms based on review of the scientific literature. Many different interpretation systems are in everyday use, which vary in content and quality and rely heavily on data relating the genotype to phenotype, rather than to the response of patients in the clinic

The new approach pioneered by the RDI involves relating the genotype, and other clinical information, directly to the virological response of patients to different combinations of drugs in clinical practice, using a substantial database of clinical data and artificial intelligence. Many scientists and clinicians believe that this approach has the potential to greatly improve the interpretation of genotypic information and ultimately the quality of treatment decision-making.

THE CONCEPT

The concept is to develop a relational database that will enable HIV genotypic resistance data to be related directly with *in vivo* virological response to antiretroviral agents using computational modelling. A range of sophisticated modelling methods, including artificial neural networks, random forests and support vector machines are being used to enable the accurate prediction of viral load response to combination therapy from baseline genotype and other information.

The development of the database is an international collaborative initiative. A number of private and public research groups around the world have already made significant contributions to the database & continue to provide updated data.

DATA

All data entered into the RDI database is first stripped of any patient identifiers before being sent to the RDI and each data set assigned a unique identifier.

The data currently required by the RDI for its computational modelling are as follows:

Mandatory data:

HIV RNA viral load measures	Viral load sample dates
HIV genotypes (RT and protease sequence data)	Genotype sample dates
CD4 counts	CD4 count dates
Details of antiretroviral drug therapy	Antiretroviral drug start and stop dates
Antiretroviral treatment history	

Optional data:

Gender	Age
Measures/estimates of adherence	Plasma drug levels
Viral load assay details	Genotype assay details
Phenotype test results	Phenotype sample dates

QA standards and minimum data requirements

There will be strict quality assurance of the database with clinical data having to conform to one of two standards:

1. Clinical trials standards - compliance with Good Clinical Practice (GCP) standards
2. Cohort standards - compliance with the principles of the Declaration of Helsinki

Data meeting either of these standards will be flagged in the database as 1 or 2 respectively.

Sequence Data

Sequence data will also be fully quality assured and must comply with the following minimal length and standard criteria:

Protease: A bi-directional sequence of the protease gene from codons 10 to 99.

RT: A bi-directional sequence of the reverse transcriptase gene from codons 41 to codon 235.

PROGRESS TO DATE

The RDI is registered in the UK as an independent, not-for-profit organization, a core research team of four and an international advisory group of 10. As of October 2010 the database held data from approximately 70,000 patients.

Neural network, random forest and support vector machine models have been developed and have produced statistically reliable predictions of virological response. Correlations between predicted and actual changes in viral load have had r^2 values typically of 0.6-0.7. The mean absolute difference between the predicted and actual viral load values are in the region of 0.5 log copies/ml, similar to the reproducibility of the assays. The most recent models have predicted the probability of the follow-up viral load following the initiation of a new regimen being reduced to <50 copies HIV RNA/ml with an accuracy of approximately 80%.

HIV-TRePS, the RDI's HIV Treatment Response Prediction System was launched in October 2010. This unique experimental system enables healthcare professionals to obtain predictions of how their patients may respond to different combinations of antiretroviral drugs of their choice plus alternative regimens in common clinical use.

Scientific Advisory Group

1. Dr Julio Montaner — BC Center for Excellence in HIV/AIDS, Vancouver, BC, Canada
2. Dr Richard Harrigan — BC Center for Excellence in HIV/AIDS, Vancouver, BC, Canada
3. Dr Jose Gatell — Hospital Clinic of Barcelona, Barcelona, Spain
4. Dr Carlo Torti — University of Brescia, Brescia, Italy
5. Dr Brian Gazzard — Chelsea and Westminster Hospital, London, UK
6. Dr John Baxter — CPCRA, USA
7. Dr Sean Emery — NCHECR, Sydney, Australia
8. Dr Victor DeGruttola — Harvard School of Public Health, Boston, MA, USA
9. Dr Lidia Ruiz — Fundacion IrsiCaixa, Barcelona, Spain
10. Dr Anna Maria Geretti — The Royal Free Hospital, London, UK