Application of a Genotypic Driven Rule-Based Expert Artificial Intelligence Computer System in Treatment Experienced HIV-Infected Patients. Immunologic and Virologic Response.

Paul J. Cimoch MD, Darryl M. See MD**, Michael J. Pazzani Ph.D.**, William M Reiter MD*, Richard H. Lathrop Ph.D.**, Wendy A Fasone RN*, Jeremiah G Tilles MD ***
*Center for Special Immunology
**University of California, Irvine

<u>Background</u>:
A rule-based expert computer program, "Customized Treatment Strategies for HIV" (CTSHIV) has been developed connecting the literature describing specific HIV drug resistant mutations with the patient's genotype. Drug treatment recommendations are made based upon: known drug-resistant mutations, ranking and weighting based upon antiviral activities, overlapping toxicity's, relative levels of drug resistance and proportion of drug-resistant clones in the patients HIV quasispecies. All FDA approved antiretroviral (ARV) drugs are ranked by their estimated ability to avoid both current and nearby drug resistant mutants. The goal of this application is to mediate between the scientific literature and the patients' current infection to help a physician avoid or avert HIV drug resistance.

<u>CTSHIV Program:</u>
The key enabling artificial intelligence (AI) technology is knowledge representation of the relevant scientific literature about HIV drug resistance as a set of sequence pattern rules in the HIV genome. Rule-based expert systems declaratively represent knowledge of a specialized problem and facts about a specific case, and from these draw inferences about the case. Here, the rules encode information on drug resistant mutations of HIV, the facts are the sequences of HIV genome obtained from a specific individual, and the inference to be drawn is a set of drug combinations to be recommended for the patient.

Rule forward chaining from the patient's current HIV sequences yields currently resistant HIV mutants. Rule backward chaining through sequence space yields the nearby putatively resistant mutants. Together, they allow CTSHIV to avoid both sets of mutants.

Current drug resistance is identified by applying the 55 rules in the knowledge base to the HIV sequences from the patient. The rules represent knowledge about HIV drug resistance as a set of if-then rules of the form:

IF < antecedent > THEN < consequent > [weight].

For example, one such rule in CTSHIV is:

IF Methionine is encoded by RT codon 151, THEN do not use AZT, ddI, d4T, or ddC. [weight = 1.0] (*Inversen et al 1996*)

The weight associated with a rule is not a confidence as in many expert systems. Rather, it reflects the estimated level of resistance to a particular drug and is part of the consequent. Weights range from 0.1 (low) to 1.0 (high) based upon expert advice and the level of resistance reported in the literature.

To estimate current resistance, rule weight is multiplied by the fraction of viral sequences that trigger the rule, and combined additively. As a summary metric we use:

$$\mathit{CurrWt}(\mathrm{D}) = \Sigma_{r \in \mathrm{Rules}(\mathrm{D})} \Sigma_{s \in \mathrm{S}} \mathrm{Apply}(r, s)/|\mathrm{S}|$$

Where D is a set of drugs comprising a combination therapy, Rules(D) are the rules that confer resistance to a drug in D, S is the set of the HIV sequences extracted from the patient, and Apply (r, s) yields the rule weight of r if r fires on s and 0 if not. *CurrWt* is comparable only between combinations with the same number of drugs because any superset of a drug combination has equal or greater current weight.

Under this model, the total current level of resistance to a multi-drug combination is the sum of the current resistances to each drug. The effect of this is to identify drug combinations that have little or no current resistance and therefore attack the virus strongly.

Predict Nearby Resistant Mutants

Nearby resistant mutants are predicted by a backward-chaining search through mutation sequence space, beginning with the patient's current HIV sequences. At each step, a sequence that does not fire a rule is used to generate several new sequences that do. The new sequences are identical except that codon positions mentioned by the rule are modified so that the rule does fire. They represent mutants that are close in Hamming distance but resist the drugs mentioned by the rule. Conceptually, every virtual mutant within a pre-determined Hamming distance cut-off is examined. Currently all mutants up to and including Hamming distance three are considered. Branch-and-bound techniques speed the search by pruning unnecessary examinations. Currently CTSHIV runs in about a minute per patient, which is acceptable for now.

Under this model, a mutant resists a drug combination only as strongly as it resists the least-resisted drug in the combination, and a drug combination suppresses a virus population only as strongly as it suppresses the most-resistant member of the population. The effect of this is to identify nearby mutants that resist every drug in a combination, and drug combinations such that no nearby mutant resists every drug.

Rank Alternatives

CTSHIV ranks alternative drug combinations using the current resistance weight (CurrWt) and the nearby mutant resistances (MutScore). The best ranked combinations of 1, 2, 3, and 4 drugs are generated independently. This is done by sorting the combinations by any monotonic function of CurrWt and MutScore. Currently we use Euclidean distance, the square root of $CurrWt^2(D) + MutScore^2(D)$, to rank drug combination D. Values near or at zero indicate little or no resistance, and increasing positive values indicate increasing resistance. The best-ranked combinations represent a satisfying compromise along both metrics simultaneously.

<u>Suggest Clinical Treatment Protocols</u>
The final result of application processing is to recommend the five highest-ranked combinations of 1, 2, 3, and 4 drugs. The next highest-ranked RT-only combination is shown for comparison. Figure A shows 3-drug combinations recommended for an HIV patient. Figure B shows an example of nearby resistant mutants. It is hoped that the CTSHIV output will increase patient adherence, by clearly showing the deleterious effects of failing to take all medication. Figure C shows the projected consequences of non-adherence to the highest-ranked 3-drug combination of Figure A.

CTSHIV associates citations from the scientific AIDS literature with each rule and maintains a trace of rule execution. Therefore, it can explain its conclusions and point the physician at the relevant medical literature. Figure D gives an example. Cost is an important factor in HIV treatment because many of the drugs are difficult to manufacture and inherently expensive. CTSHIV provides cost estimates of the highest-ranked treatments. Cost codes shown in Figures A are described in the Figure legend.

There are important limitations of the approach described. Sequence-based rules capture only part of the domain knowledge about drug resistance, albeit a clinically useful part. Drug resistance may arise for other domain-specific reasons that cannot be represented easily as rules. Current sequencing techniques may provide only partial or no information about minority strains. The rule set is only as complete as current scientific knowledge allows. Currently it may be possible to infer when resistance is likely to occur, based on genome sequences actually seen in the patient that correspond to resistance-conferring mutations described in the scientific literature. However, it is impossible to guarantee the non-existence of an unsuspected resistant mutant.

<u>Methods</u>:
Fourteen patients failing $\leq$ 5 ARV's were studied. CD4 counts and viral load (VL) (Amplicor PCR) were performed at day –21, day zero and every three months. For each patient, the reverse transcriptase and protease portions of the *pol* gene were amplified by RT-PCR, 5 clones were produced, plasmid DNA extracted, and complete sequencing was performed by an ABI sequencer (Applied Biosystems) with the data directly downloaded into the CTSHIV program. The five most effective 2, 3, and 4 drug regimens coupled with an explanation for their choice was displayed for each patient. The clinician chose from the options based upon order of ranking from the program, history of drug intolerance and patient preference. Patients were seen monthly for tolerance, compliance and routine lab assessments. Every three months CD4 counts and VL were performed. If the VL was greater than 1,000 copies/ml then genotyping was performed, a new CTSHIV display was generated and ARV therapy adjusted based upon these new results at the following monthly visit.

Results:

Of the 14 patients treated according to CTSHIV, 12 patients completed 12 months of follow-up. One patient died at 6 months from progressive pulmonary KS. The only treatment naïve patient developed drug related side effects and was lost to follow-up after month 6. Of the 12 patients completing the study, the mean number of baseline RTI's were 3 and 5 of the 12 had also been treated with a protease inhibitor (SQV). The baseline mean CD4 count of the entire group was 498 cells/mm^3 (range 5-1201) and mean VL was 80,205 copies/ml (range 983-751,385 copies/ml). There was no significant statistical baseline difference between the responder and non-responder groups.

At 12 months, 9 of the 12 patients (75%) had undetectable VL (<400 copies/ml) and 1 of 12 (8%) experienced a >1 log drop in VL for a total response rate of 83%. Of the 2 patients (17%) who were non-responders, one did meet the criteria as a responder until month 12 when he failed due to admitted medication non-adherence and the other experienced multiple adverse experiences related to the medications. A mean rise in CD4 cells from baseline of 294 cells/mm^3 (range –67 – 1143) was seen in the responder group. Several patients changed therapies due to side-effects of the medications yet still responded to alternative CTSHIV recommendations. (See Table 1.)

Conclusions:

The current standard for choosing ARV therapy is clinical judgment based upon VL and CD4 cell count. Drug failure is often associated with HIV-genotypic mutations. Patients with persistently elevated VL despite various ARV therapies are a particular challenge to manage and historically have poor treatment outcomes. The use of the CTSHIV system equipped with mutational prediction artificial intelligence fields successfully assisted in the ARV management of the vast majority of this treatment experienced HIV-infected individuals.

Knowledge of current or nearby mutants putatively resistant to one or more drugs is valuable to a physician treating an HIV patient. In conjunction with HAART, such knowledge may help select a combination of drugs less likely to be resisted. Currently there are 11 drugs approved by the FDA for HIV, plus one available for compassionate use. These 12 result in 407 different combination treatments of four or fewer drugs, as some drugs should not be used together. A physician may find it tedious to scan many genetic sequences, be unfamiliar with the latest HIV drug resistant mutations reported, or have difficulty ranking the hundreds of treatment choices for each patient. CTSHIV mediates between the scientific literature and the patient's current infection and may successfully assist a physician choose ARV.

Table 1

Responders <400 VL @ 12mo	**Prior ARV's**	**Baseline VL copies/ml PCR**	**CTS ARV Rec's**	**Day 90 VL**	**Day 180 VL**	**Day 270 VL**	**Day 360 VL**
AHM	3TC, d4T	7856	RTV, SQV, 3TC	<400	<400	<400	<400
STJ	AZT, ddI, 3TC, d4T SQV	751,385	3TC, d4T, IDV[1]	<400	<400	<400	3TC, d4T, NFV <400
NCP	AZT, 3TC	5271	RTV, SQV	<400	Off meds[1] 1615	d4T, AZT, IDV 1396	<400
BGC	AZT, ddI, 3TC	2397	RTV[1], SQV[1], d4T AZT, d4T, IDV	<400	<400	<400	<400
DDA	3TC, d4T SQV, DLV, AZT	983	DLV, d4T	<400	805 AZT, NVP[2]	1467 3TC, d4T, SQV	<400
REJ	AZT, ddI, 3TC	3160	DLV[2], d4T[2]	4469 NFV, SQV, 3TC	<400	<400	<400
KMS	3TC, d4T, ddI, SQV	3161	NFV, NVP, d4T	<400	<400	<400	<400
AXM	3TC, ddC, AZT	2638	NFV, d4T	<400	<400	<400	<400
RRN	AZT, ddC, 3TC	1328	NVP, d4T, 3TC	Off meds[1] 113439	Combivir NFV[1], SQV[1] 89609	Off meds[1] 65462 Crixivan, Combivir 4019	<400

Responders > one (1) log drop compared to baseline	**Prior ARV's**	**BaselineVL copies/ml PCR**	**CTS ARV Rec's**	**Day 90 VL**	**Day 180 VL**	**Day 270 VL**	**Day 360 VL**
AHS	AZT, ddI, 3TC, d4T, DLV	31153	NVP[2], d4T[2], NFV[2]	18954	14593 3TC, ddI, d4T, NFV	<400	1277

Non-Responders	**Prior ARV's**	**Baseline VL copies/ml PCR**	**CTS ARV Rec's**	**Day 90 VL**	**Day 180 VL**	**Day 270 VL**	**Day 360 VL**
RSP	AZT, 3TC, d4T, ddC, SQV	25930	NFV[1], ddC, d4T	<400 d4T[3], ddC[3] RTV[3], SQV[3] Off Meds 80170	AZT[1], NVP[1], IDV[1] 999	Off Meds[1,4] 12212	FTV[2,4], NFV[2,4,] 3TC,[2,4] DLV[2,4], 110156
RDC	AZT, 3TC, d4T, ddC	18629	NFV, SQV	2873	<400	<400	166,627[5]

Withdrawn from Study	**Prior ARV's**	**Baseline VL copies/ml PCR**	**CTS ARV Rec's**	**Day 90 VL**	**Day 180 VL**	**Day 270 VL**	**Day 360 VL**
JKB	AZT, ddC, SQV	94459	AZT[2,4] d4T[2,4] NFV[2,4]	197965 NFV, 3TC, NVP	10017	Hospital Admit- Died pulmonary KS	
AGB	Treat-ment naïve	101824	3TC, IDV[1]	Off meds 196,756 NFP[5], 3TC[5], NPV[5], d4T[5]	Off meds[5] 135,500	>750,000 Lost to Follow Up	

[1]Drug(s) intolerance, side effects
[2]Failed drug(s) therapy
[3]Drug(s)toxicity
[4]Opportunistic infection(s)
[5]Non-compliant with drug therapy

These protocols with 3 drugs are recommended: CurrWt MutScor 0 Mut 1 Mut 2 Mut 3 Mut

A5 SAQUINAVIR NELFINAVIR D4T:	0.06	0.1	0.0	0.0	0.0	0.1
B3 SAQUINAVIR DELAVIRDINE D4T:	0.00	0.2	0.0	0.0	0.0	0.2
C3 SAQUINAVIR NEVIRAPINE D4T:	0.00	0.4	0.0	0.0	0.0	0.4
D4 SAQUINAVIR DELAVIRDINE AZT:	0.00	0.6	0.0	0.0	0.0	0.6
E4 SAQUINAVIR NEVIRAPINE AZT:	0.00	0.6	0.0	0.0	0.0	0.6
RF3 DELAVIRDINE DDI AZT:	0.08	1.2	0.0	0.0	0.2	0.9

FIGURE A: Example 3-drug output from HIV patient ‘‘AA,” showing a favorable resistance profile. For the highest-ranked treatment, current resistance (CurrWt) and nearby mutation score (MutScor) are small, and only a weakly-resistant mutant appears even out to Hamming distance three (3 Mut). The letters A--F identify treatments. Treatment F is the best RT-only treatment (indicated by the prefixed letter R).

Digits after the letters indicate cost codes (0 = $0 to $200, …, 3 = $600 to $800, 4 = $800 to $1000, 5 = $1000 to $1200, etc. per month estimated average wholesale cost). Based on these per-drug costs (per month, estimated average wholesale): AZT = $432; DDI = $125; 3TC = $230; DDC = $186; D4T = $335; NEVIRAPINE = $248; DELAVIRDINE = $222; SAQUINAVIR = $191; INDINAVIR = $451; RITONAVIR = $576; NELFINAVIR = $570

	CurrWt	MutScor	0Mut	1Mut	2Mut	3Mut
A5 D4T NELFINAVIR SAQUINAVIR:	0.06	0.1	0.0	0.0	0.0	0.1

Current: (NELFINAVIR) RT 151:CAG->ATG by R11 (D4T) PRO 90:TTG->ATG by R28 (SAQUINAVIR)

FIGURE B: Example output for HIV patient ``AA" showing one example of the closest mutants inferred to most resist every drug in the top-ranked 3 drug combination of figure 4. Three letters must change simultaneously. Currently nelfinavir is resisted; changing two letters at RT 151 resists D4T and changing one at PRO 90 resists saquinavir.

	CurrWt	MutScor	0Mut	1Mut	2Mut	3Mut
A5 SAQUINAVIR NELFINAVIR D4T:	0.06	0.1	0.0	0.0	0.0	0.1
If stop NELFINAVIR:		0.6	0.0	0.0	0.0	0.6
If stop SAQUINAVIR:		1.1	0.0	0.0	0.1	1.0
If stop D4T:		2.1	0.0	0.1	0.6	1.1

FIGURE C: Example output for HIV patient ``AA" showing the projected result from stopping any single drug in the top-ranked 3 drug combination of Figure 4 (saquinavir, nelfinavir, d4T). Mutants are closer or worse or both. Stopping nelfinavir is bad, stopping saquinavir is worse, and stopping d4T is worst of all.

By RULE36 NELFINAVIR RITONAVIR INDINAVIR reduced by 0.06 since 3 equences had GTT at codon 71 of the PRO gene. Patrick A, Duran M, Cao Y, et al. Genotypic analysis of HIV-1 variants isolated from patients treated with the protease inhibitor nelfinavir, alone or in combination with d4T or AZT and 3TC. ABSTRACT, Fourth Conference on Retroviruses and Opportunistic Infections, Washington, DC, 1997.

By RULE14 3TC reduced by 0.40 since 2 sequences had GTG at codon 184 of the RT gene. Quan Y, Gu Z, Li X, Li Z, Morrow C, Wainberg M. Endogenous reverse transcription assays reveal high-level resistance to the triphosphate of (-)2-dideoxy-3-thiacytidine by mutated 184V human immunodeficiency virus type 1. J Virol 1996;70:5642-5645.

FIGURE D: Partial example output literature