Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Medicine

journal homepage:www.elsevier.com/locate/apjtm

Document heading doi:

Predicting a single HIV drug resistance measure from three international interpretation gold standards

Singh Yashik*, Mars Maurice

Department of Telehealth, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, 719 Umbilo Road, Umbilo, Durban, South Africa

ARTICLE INFO

Article history:
Received January 2012
Received in revised form 15 March 2012
Accepted 15 May 2012
Available online 20 July 2012

Keywords:
Drug resistance
Antiretroviral therapy
Highly active
HIV
Artificial intelligence
Expert systems

ABSTRACT

Objective: To investigate the possibility of combining the interpretation of three gold standard interpretation algorithms using weighted heuristics in order to produce a single resistance measure. **Methods:** The outputs of HIVdb, Rega, ANRS were combined to obtain a single resistance profile using the equally weighted voting algorithm, accuracy based weighing voting algorithm and the Bayesian based weighted voting algorithm techniques. **Results:** The Bayesian based voting combination increased the accuracy of the resistance profile prediction compared to phenotype, from 58% to 69%. The equal weighted voting algorithm and the accuracy based algorithm both increased the prediction accuracy to 60%. **Conclusions:** From the result obtained it is evident that combining the gold standard interpretation algorithms may increase the predictive ability of the individual interpretation algorithms.

1. Introduction

1.1. Background

Resource poor countries like those in Africa face a triple burden of disease. Human immunodeficiency virus, tuberculosis and malaria are changing the population dynamics of many resource poor countries by creating a larger young population and a significantly smaller middle to old age population. There are currently almost 5.6 million infected with HIV in South Africa, which is approximately 11% of the South African population[1]. It is also estimated that there are almost 500 000 patients who exhibit AIDS defining conditions[2].

Although there is currently no cure or vaccine for HIV infection, it can be managed by highly active antiretroviral therapy. The management of HIV is however complicated due to antiretroviral drug resistance. For the purpose of this paper, drug resistance may be defined as the inability of the antiretroviral (ARV) drugs to adequately suppress viral load. It was reported in 2009, that 37% of patients that required ARV treatment actually received ARV drugs[3]. The factors that facilitate antiretroviral drug resistance are high replication rates of the virus, selective pressure caused by the ARV drugs and initial infection by resistant strains of HIV. Thus it is inevitable that drug resistance will become concern in treatment of HIV/AIDS infected patients. Without effective ARV treatment patients are susceptible to opportunistic infections which may be fatal e.g. pneumonia and tuberculosis. AIDS related conditions, including ARV drug resistance causes 43.6% of deaths in the HIV positive population in South Africa[4].

It thus follows that drug resistance testing may become routine in the treatment of HIV infected patients. Currently, phenotypic laboratory tests are used to determine if a patient is resistant to ARVs. These methods are however expensive, time consuming and prone to error. Electronic computerized algorithms[5-8] may also be used to determine ARV drug resistance, and have many advantages over phenotype testing. Computer based genotype interpretation algorithms usually determine mutations in the patient's pol gene region, and uses this information to determine the ARV

*Corresponding author: Singh Yashik, Department of Telehealth, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, 719 Umbilo Road, Umbilo, Durban, South Africa.
Tel: +27312604543
Fax: +27312604737
E-mail: singhy@ukzn.ac.za

drugs to which patients are resistant to. These computer based tests are faster and cheaper than phenotypic tests.

Computer based algorithms are usually based on an experts' understanding of the domain, available datasets that are used for machine learning and understanding of published literature. This has led to the creation of many different interpretation algorithms, which produce different resistance measures even if applied to the same resistance profile.

The aim of this study is to combine the interpretation of three gold standard interpretation algorithms using weighted heuristics in order to produce a single resistance profile.

1.2. Interpretation algorithms

One type of interpretation algorithms is based on domain knowledge. These interpretation algorithms are logic or decision tree based expert systems that are made up of rules describing interactions between certain mutations and/or combination of mutations with resistance. This means that all computational decisions concerning resistance are based on known mutation-resistance rules found in published scientific literature.

REGA, Agence Nationale de Recherches sur le SIDA (ANRS) and HIV-db are widely used as the goal standards in interpretation algorithms. REGA was developed by the laboratory for clinical and evolutionary virology, Rega institute for medical research, Katholieke Universiteit Leuven. It classifies ARV resistance according to three levels, susceptible, intermediate and resistant. Susceptible indicates that for a given resistance profile, a particular ARV drug will be effective in managing the HIV infection. Intermediate indicates that the ARV drug is partially effective, and if the ARV is not effective the resistance profile is classified as resistant. REGA also incorporates the idea of a genotypic susceptibly score that is assigned to each drug for each of the three levels. Resistance to the therapy is confirmed when the sum of the individual genotypic susceptibility scores are lower than a pre-determined cutoff. The HIV-db program was developed by the Division for Infectious Diseases, Stanford University Medical Center, Stanford University. The expert system itself generates a resistance profile using Boolean based rules and also includes penalties. HIV-db classifies HIV resistance according to five levels: susceptible, potential low-level resistance, low-level resistance, intermediate resistance, and high-level resistance. The French National Agency for AIDS Research AC11 Resistance group developed the ANRS algorithm based on statistically determining the relationship between genome mutations and virological outcomes of patients failing ARV therapy. The ANRS classification of resistance is similar to REGA.

These interpretation algorithms were however developed using different datasets, subtypes, analyzed on drug-naive and -experienced patients etc. These differences have led to the creation of many different interpretation algorithms. Initially, studies suggested that the interpretation algorithms produce different resistance measures even if applied to the same resistance profile. Recent studies with updated interpretation rules still suggested some discordance between interpretation algorithms.

Four interpretation algorithms (ANRS-3-02, TRUGENE VGI-6, Rega 5.5 and HIVdb-8-02) were studied[9] and it was concluded that there was a discrepancy in the interpretations of 33% of all the data tested. The most discordant were NRTI's. De Luca *et al*[10] concluded that discrepancies in the interpretation algorithms may influence the use of resistance testing over virological outcomes. De Luca *et al*[11] later studied the application of 13 interpretation algorithms on drug naive patients and concluded that there were discordances. Wang *et al*[12] also determined that there was a high level of discordance between the interpretation of NRTI resistance, and goes on to suggest that there should be "standardization of unique interpretative rules". Vergne *et al*[13] also confirmed some discrepancies and attributed them to the application of the interpretation algorithms to drug -naive or -experienced patients. Snoek *et al*[14] confirmed that there were low discordances between the algorithms tested and suggested it may be due to subtypes. Vercauteren *et al*[15] felt that the newer versions of interpretation algorithms were converging but there are some discrepancies in their interpretation. Poonpiriya *et al*[16] also found that some discrepancies in the seven interpretation algorithms they studied. Yebra *et al*[17] concluded that there were some discordance in interpretation of subtype B sequences, but more variation in non-B subtype interpretations.

1.3. Weighted voting techniques to combine classifiers

The combination of classifiers has been shown in many studies to produce higher accuracies than the base classifiers alone[18-20]. One method of combining classifiers is to use weighed voting techniques[21,22]. Three examples of such ensemble techniques are simple additive voting, accuracy based voting and Bayesian based weighted voting algorithms[23].

In the simple additive voting technique, each individual classifier has an equal weighting contribution to the single combined output. The single resistance measure is obtained by determining the resistance measure that is most suggested by the gold standard interpretation algorithms. If the three gold standard interpretation algorithms each suggest different resistance measures, the middle measure is reported. Accuracy based voting technique uses the normalized accuracy of each gold standard interpretation algorithm as the weight additively assigned to the resistance profile. The resistance profile with the highest summed weight is reported as the single resistant output.

In order to apply the Bayesian based voting, assume C is a confusion matrix for each expert classifier (e) with M possible classes, such that C is the number of class i patterns that classifier k assigns to class j. Thus the probability that a pattern x actually belongs to class i, given that k experts assign in to class j, can be calculated from Equation 1. Thus

for any pattern x such that the K experts classify it as $e(k) = jk$ for $1< k < K$, the belief, B, that the pattern x belongs to class i may be obtained from Equation 2. Applying Bayes formula, B(i) may be approximated by Equation 3. Thus any pattern x can be assigned to class i if B(i) > B(j) for all $1 < j < M$ and $i \neq j$.

$$P(x \in C_i \mid e_k(x) = j) = \frac{C_{ij}^k}{\sum_{i=1}^{M} C_{ij}^k} \quad (1)$$

$$B(i) = P(x \in C_i \mid e_k(x) = j_1, \cdots, e_k(x) = j_k) \quad (2)$$

$$B(i) = \frac{\prod_{k-1}^{k} P(x \in C_i \mid e_k(x) = j_k}{\sum_{i=1}^{M} \prod_{k-1}^{k} P(x \in C_i \mid e_k(x) = j_k} \quad (3)$$

2. Materials and methods

The methodology employed in this study was broken down into four parts, namely, data pre-processing, data processing, classification algorithm combination and statistical analysis of the combined interpretation algorithms.

2.1. Data pre-processing

Genotype-Phenotype datasets that consisted of 2 928 protease gene and 1 981 reverse transcriptase gene sequences were obtained from the Stanford HIV drug resistance database (http://hivdb.stanford.edu/). This database contains publicaly available de-identified HIV drug resistance data. The dataset consisted of a sequence identifier, subtype of the sequence, phenotyping method, isolate identifier, fold resistance of drug compared to the wild type, amino acids at various positions, and mutation lists.

The amino acid for each protease and reverse transcriptase sequence in the dataset was then processed into its respective three base nucleotide sequence. This was accomplished by feeding individual PR and RT sequences into a Java application that replaced each amino acid with the triplet nucleotide as shown in Table 1.

Table 1.
Triplet coding for each amino acid in the PR and RT sequences.

Amino acid	Triplet codon	Amino acid	Triplet codon
Ala/A	GCU	Leu/L	UUA
Arg/R	CGU	Lys/K	AAA
Asn/N	AAU	Met/M	AUG
Asp/D	GAU	Phe/F	UUU
Val/V	GUU	Tyr/Y	UAU
Ile/I	AUU	His/H	CAU
Thr/T	ACU	Gly/G	GGU
Glu/E	GAA	Gln/Q	CAA
Cys/C	UGU	Pro/P	CCU
Ser/S	UCU	Trp/W	UGG

2.2. Data processing

The nucleotide list was fed into to the online HIValg V6.0.11 program hosted by Stanford university (http://sierra2.stanford.edu/sierra/servlet/JSierra?action=hivalgs). This web application takes as input the nucleotide list, converts this to an amino acid list, determines mutations that occur in the genome, and gives resistance profiles of that sequence by applying the REGA, HIVdb and ANRS algorithms to it.

The predicted output resistance profile for each sequence was linked with its original amino acid sequence and its associated phenotype IC_{50} score. With respect to this study, the IC_{50} scores were treated as a true measurement of drug resistance. Based on the ranges of the IC_{50} score, an actually resistance measurement was determined.

2.3. Combining the classifiers

The outputs of HIVdb, Rega, ANRS were combined to obtain a single resistance profile using the equally weighted voting algorithm, accuracy based weighing voting algorithm and the Bayesian based weighted voting algorithm techniques. Excel was used to create the mathematical model that performed the combination.

2.4. Statistical analysis

Standard deviation, z-score and p-score were calculated in order to perform a proportion test. This was calculated to determine if the accuracies obtained where statistically different from chance. The null hypothesis proposed was that the average accuracy observed resulted purely from chance *i.e.* Proportion(P)=0.5. The standard deviation was calculated as shown in Equation 4 and the Z-score as calculated from Equation 5. P-score was determined by solving Equation 6. Similarly the proportional Z test was used to determine if the difference between the algorithms were statistically significant.

$$\sigma = \sqrt{\frac{P = (1 - P)}{n}} \quad (4)$$

$$z - score = \frac{p - P}{\sigma} \quad (5)$$

$$P(Z \leqslant z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{-u}{2}} du \quad (6)$$

Another statistical method employed to determine if significant differences were produced by the different combination models was described by Salzberg[23]. It was reported that a P-score for comparing classifiers may be computed using the binomial distribution, which gives the probability of s successes in n trials as shown in Equation 7.

$$\frac{n!}{s!\ (n-s)!}(pq)^n \quad (7)$$

Thus, if we expect no difference between the algorithms, then $p=q=0.5$ and the p-score may be calculated as shown in Equation 8.

$$p-score=\sum_{s=x}^{n}\frac{n!}{s!\ (n-s)!}(0.5)^n \quad (8)$$

Further analysis was performed to determine the distribution of the discrepancies were between the resistance measures in the misclassified data elements. Discrepancies were divided into three groups: G1, G2 and G3. G1 represented no discrepancies between the Bayesian based voting algorithm and each of the three gold interpretation algorithms. G2 represent a single resistance measure difference between the Bayesian based voting algorithm and each of the three gold standard interpretation algorithms, *i.e.* intermediate classified as susceptible or resistant or vice versa. G3 represented a two resistance measure difference between the Bayesian based voting algorithm and each of the three gold interpretation algorithms, *i.e.* susceptible classified as resistant and *vice versa*.

3. Results

Accuracy was calculated in order to determine the effectiveness and to compare the gold standard interpretation algorithms and the weighted voting algorithms. Accuracy is defined as the percentage of resistance profiles that have resistance measures as identified by the IC_{50} score. The accuracy obtained by each gold standard interpretation algorithm and weighted voting algorithm is shown in Table 2.

Table 2.
The accuracy of the HIV-db, REGA, ANRS, equal weighting, accuracy weighting and bayesian weighting algorithms in predicting phenotype.

Drug	ANRS	HIVdb	REGA	Equal	Accuracy	Bayesian
Atv	*	*	69.8	69.8	69.8	69.8
Idv	51.6	51.3	51.9	52.8	52.8	69.8
Lpv	68.7	72.5	77.1	75.7	75.7	80
Nfv	57.4	55.5	54.2	55.9	57.9	65.4
Sqv	61.7	53	59.6	61.9	62.0	66.7
Tpv	64.6	61.7	57.8	66.7	62.4	75.1
3tc	67.9	68.8	69.0	68.9	68.8	75.3
Abc	50.1	48.7	50.6	48.0	48.0	55.6
Azt	*	54.9	55.1	54.3	53.6	59.4
D4t	44.5	49.3	48.9	47.9	47.9	58.2
DDI	50.2	48.2	48.0	50.1	50.5	55.2
DLV	*	64.8	67.4	63.8	62.4	72.4
EFV	65.9	64.6	65.9	62.9	62.9	70.0
NVP	73.6	74.6	75.7	77.6	78.0	81.0
TDF	48.1	45.5	46.5	47.9	47.8	79.0

* indicates that the gold standard did not produce any results.

The Salzberg method as described in Equations 8 was performed in order to compare the three weighted voting algorithms with each another to determine if there are any statistical differences between them. Table 3 describes the P-value obtained from the Salzberg when comparing the weighing algorithms. The comparison was also confirmed using the proportional Z-test. Table 4 shows the Z-score of the comparison of equal, accuracy, and Bayesian weighting algorithms.

Table 3.
P-value of the comparison of the weighing algorithms using Salzberg's method[23].

Drug	Equal weighting		Accuracy weighing Bayesian weighting
	Accuracy weighting	Bayesian weighting	
atv	0.9	0.9	1.0
idv	0.9	< 0.01	< 0.01
lpv	<0.01	< 0.01	0.9
nfv	0.07	< 0.01	< 0.01
sqv	0.03	< 0.01	< 0.01
tpv	0.01	< 0.01	<0.01
3tc	0.9	< 0.01	0.9
abc	0.9	< 0.01	< 0.01
azt	0.9	< 0.01	< 0.01
d4t	0.9	< 0.01	< 0.01
ddi	0.9	< 0.01	< 0.01
dlv	0.9	< 0.01	< 0.01
efv	0.9	0.01	0.03
nvp	0.9	0.8	0.9
tdf	0.9	0.9	0.9

Table 4.
Z-score of the comparison of equal, accuracy, Bayesian weighting algorithms.

Drug	Equal Weighting		Accuracy Weighing Bayesian weighting
	Accuracy weighting	Bayesian Weighting	
Atv	0.00	0.00	0.00
Idv	0.00	20.40*	20.40*
Lpv	0.00	4.80*	4.80*
Nfv	−2.60	11.70*	9.30*
Sqv	1.96*	5.90*	5.80*
Tpv	5.40*	9.70*	14.70*
3tc	0.10	7.40*	7.5*
Abc	0.00	10.20*	10.20*
Azt	1.00	6.60*	7.50*
d4t	0.00	7.40*	7.4*
Ddi	−0.60	6.90*	6.30*
dlv	1.80	10.10*	11.80*
efv	0.00	8.50*	8.5*
nvp	−0.50	3.80*	3.30*
tdf	0.10	2.70*	12.80*

* indicates a statistically significant difference.

In order to determine if there is a statistical difference between the weighted voting method and the gold standard interpretation algorithms, the proportional Z-test was performed. Table 5 shows the Z scores of ANRS, REGA and HIVdb compared to Equal, Accuracy and Bayesian weighting algorithms. In order to better understand the nature of the discrepancies of the Bayesian based weighting algorithm the percentages of the G1, G2 and G3 values were calculated as shown in Table 6.

Table 5.

Z scores of ANRS, REGA and HIVdb compared to equal, accuracy and Bayesian weighting algorithms.

Drug	Equal weighting			Accuracy weighting			Bayesian weighting		
	ANRS	HIV-db	REGA	ANRS	HIV-db	REGA	ANRS	HIV-db	REGA
Atv	X	X	0.0	X	X	0.0	X	X	0.0
Idv	1.7	2.1*	1.2	1.7	2.1*	1.2	25.3*	25.8*	24.8*
Lpv	8.4*	3.8*	−1.6	8.4*	3.8*	−1.6	13.6*	8.8*	3.3*
Nfv	−2.0	0.5	2.3*	0.7	3.2*	5.0*	10.6*	13.3*	15.2*
Sqv	0.3	12.2*	3.0	0.4	12.4*	3.1*	6.4*	18.8*	9.2*
Tpv	2.6*	6.4*	11.7*	−2.7	0.9	6.1*	13.1*	17.1*	22.8*
3tc	1.2	0.1	−0.1	1.1	0.0	−0.2	9.0*	7.8*	7.6*
Abc	−3.0	−1.0	−3.7	−3.0	−1.0	−3.7	7.8*	9.9*	7.0*
Azt	X	−0.8	−1.1	X	−1.8	−2.0	X	6.1*	5.8*
d4t	5.1*	−2.0	−1.4	5.1*	−2.0	−1.4	13.2*	5.7*	6.3*
Ddi	−0.1	2.7*	3.0*	0.4	3.3*	3.6*	7.1*	10.1*	10.4*
Dlv	X	−1.2	−4.4	X	−3.0	−6.1	X	9.4*	6.1*
Efv	−3.7	−2.1	−3.7	−3.7	−2.1	−3.7	5.1*	6.7*	5.1*
Nvp	4.7*	3.5*	2.2*	5.1*	3.9*	2.6*	8.6*	7.4*	6.1*
Tdf	−0.3	3.6*	2.1*	−0.4	3.4*	1.9	2.5*	6.4*	4.8*

* indicates a statistically significant difference. X ANRS, HIVdb or REGA did not produce results for the associated drug.

Table 6.

The percentage discrepancy between the resistance measures defined as G1, G2 and G3.

Drug	ANRS			HIVdb			REGA			Phenotype	
	G1	G2	G3	G1	G2	G3	G1	G2	G3	G2	G3
Atv	X	X	X	X	X	X	100.0	0.0	0.0	28.7	1.5
idv	38.9	60.1	0.0	51.0	49.0	0.0	96.9	3.1	0.0	30.4	0.3
lpv	84.0	8.6	0.0	86.9	13.1	0.0	91.1	8.9	0.0	22.5	1.8
nfv	44.8	39.9	0.0	37.1	62.8	0.1	37.9	62.1	0.0	35.1	0.4
sqv	59.8	26.0	0.2	53.7	46.3	0.0	50.4	49.6	0.0	32.1	1.2
tpv	87.0	7.4	0.3	71.4	28.6	0.1	75.5	24.5	0.0	24.1	0.8
3tc	92.7	3.7	0.0	98.4	1.6	0.0	97.9	2.1	0.0	13.3	17.4
abc	77.1	11.0	5.3	60.1	39.9	0.0	69.2	30.7	0.1	23.6	20.9
azt	X	0.0	X	97.2	2.8	0.0	88.7	11.3	0.0	31.1	12.0
d4t	70.7	0.8	28.0	68.2	30.7	1.1	71.0	22.2	6.8	21.9	19.8
ddi	59.9	26.8	0.0	49.8	50.1	0.1	50.7	48.9	0.4	43.6	4.4
dlv	X	0.0	X	90.3	9.7	0.1	98.9	1.0	0.1	13.0	19.4
efv	100.0	0.0	0.0	87.5	12.4	0.1	96.9	1.6	1.5	11.8	22.3
nvp	96.4	0.0	3.6	94.4	5.5	0.1	97.7	1.8	0.5	6.6	17.4
tdf	63.4	7.2	24.8	63.9	14.9	21.2	65.6	12.7	21.7	14.6	6.4

X= no results produced by the algorithm.

4. Discussion

The key finding of this study was that is it possible to combine the gold standard interpretation algorithms in such a way as to improve their individual predictive ability. Table Two shows the accuracy of the ANRS, HIVdb, REGA, equal weighting, accuracy weighting, and Bayesian weighting algorithms. The equal, accuracy and Bayesian algorithms produce the same accuracy for atv. This is due to the fact that ANRS and HIVdb do not predict a resistance profile for atv. ANRS also does not predict a resistance profile for azt and dlv. Thus the weighting algorithms in this case will be based only on the output of HIV-db and REGA.

Overall the average accuracy of the equal weighting algorithm was 60.3% (SD = 10.0), the accuracy weighting algorithm 60% (SD=9.9) and the bayesian weighting algorithm 69% (SD=8.6). Applying Equations 4-6, resulted in *Z*-score of 9.9 ($P<0.001$) for equal and accuracy weighted voting and 18.3 ($P<0.001$) for the Bayesian based weighting algorithm. This indicates that the output of the algorithms were not obtained by chance. The *Z*-score for a comparison between the equal and accuracy weighting algorithms was calculated using Equations 4-6, and a score of 0.26 ($P=0.79$) was obtained. This indicates that there is no overall statistical difference between equal and accuracy weighting algorithms. However, on comparing the equal and Bayesian weighting algorithms, a *Z*-score of 6.1 ($P<0.01$) was obtained. Similar results were obtained when comparing accuracy and Bayesian algorithms. This indicated that the Bayesian algorithm overall outperforms the accuracy and equal

weighting algorithms. Similar results where obtained for the PR and RT drugs separately.

These results based on overall accuracy, is further proven using the Sailzberg[24] method. On comparing equal and accuracy weighting algorithms, Table Three shows that there is no statistical difference in 73% drugs. However, there are statically differences between 73% of all drugs when comparing equal and Bayesian weighting algorithms, and 67% of all drugs when comparing accuracy and Bayesian weighting algorithms. This indicates that there is a significant difference between the Bayesian weighting algorithm when compared to the accuracy as well as the equal based voting algorithms. From the average accuracies it may be deduced that the Bayesian based weighting algorithm performed the best in terms of predicting ARV resistance profiles as determined by IC_{50}, while there is a slight difference between the equal and accuracy weighting algorithms[25,26].

Similar results where obtained when the *Z*-scores are assessed for the comparison of the weighing algorithms using Salzberg's method[23]. On comparing equal and accuracy weighting algorithms, there is no statistical difference in 86.7% drugs. However, there are statically differences between 93.3% of all drugs when comparing equal and Bayesian weighting algorithms, and accuracy and Bayesian weighting algorithms.

The weighting algorithms where compared to the REGA, HIV-db, and ANRS algorithms. The equal weighting algorithm showed difference in 33.7% of all the ARV drugs when compared to the ANRS algorithm. Fifty percent of the ARV drugs showed significant differences when HIV-db and Equal weighting algorithms were compared. Similarly the REGA and equal weighting algorithms had a statistical difference in accuracy in 35.7% of the ARV's. This indicates that combining the gold standard interpretation algorithms does have a positive effect on predicting resistance measures.

Similar to the equal algorithm there is an improvement in the prediction ability of the accuracy algorithm. The accuracy weighting algorithm showed difference in 25% of all the ARV drugs when compared to the ANRS algorithm. Fifty percent of the ARV drugs showed significant differences when HIV-db and accuracy weighting algorithms were compared. Similarly the REGA and accuracy weighting algorithms had a statistical difference in accuracy in 35.7% of the ARV's.

The Bayesian weighting algorithm showed difference in all the ARV drugs when compared to the ANRS algorithm. Similarly all of the ARV drugs showed significant differences when HIV-db and Bayesian weighting algorithms, and REGA and Bayesian weighting algorithms were compared. These results indicate that there are significant differences between the Bayesian weighting and HIV-db, ANRS, and REGA algorithms, and from the average accuracies it may be deduced that the Bayesian algorithm better predicts ARV resistance measures.

The result shows the weighting of ANRS, HIV-db and REGA for each drug using the Bayesian weighing algorithm model. It is evident from the Figure 1 there is no clear single gold standard which determines the outcome of the Bayesian weighting algorithm.

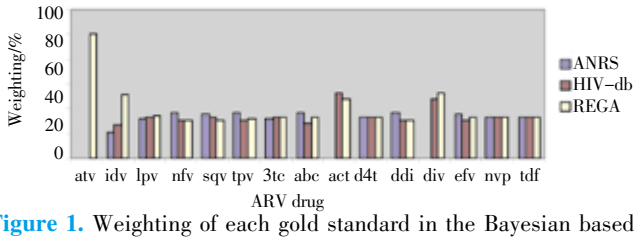

Figure 1. Weighting of each gold standard in the Bayesian based weighting algorithm.

The higher percentage of Group 2 and lower percentage of Group 3 as shown in table Six, is also advantages in terms of the efficiency of the algorithm. This shows that the discrepancies are mostly one resistance measure away and that there is little major discordance.

From the results obtained it is evident that combining the gold standard interpretation algorithms may increase the predictive ability of the individual interpretation algorithms. This prototype study indicates that using the Bayesian weighted voting method to improve the accuracy of the three individual goal standard interpretation algorithms is the most successful.

This prototype study should be extended in order to create a more efficient and effective algorithm. A dataset with a greater number of data elements should be used and other techniques to combine the gold standard interpretation algorithms should be considered. Possible options include machine learns algorithms, such as, support vector machines, multilayer perception neural networks and genetic programming. Thought should be given to designing this tool in such a manner that it may be integrated into an electronic medical record and a mobile interface design. This will serve as preliminary attempts to combine the fields of practicality defined medical informatics, bioinformatics and even telemedicine.

Conflict of interest statement

We declare that we have no conflict of interest.

References

[1] AIDS Committee of Actuarial Society of South Africa. ASSA2008 Model: ProvOutput 2011. [Online]Available from: http://aids.actuarialsociety.org.za/ASSA2008-Model-3480.htm. [Accessed on October 2011].

[2] Health Systems Trust. Aids sick: indicator data 2011. [Online] Available from: http://indicators.hst.org.za/healthstats/85/data. [Accessed on October 2011].

[3] Adam MA, Johnson LF. Estimation of adult antiretroviral treatment coverage in South Africa. *S Afr Med J* 2009; **99**(9): 661-667.

[4] Health Systems Trust. Percentage of deaths due to AIDS health indicators, statistics South Africa: statistical release P0302 mid-year estimates 2011. [Online] Available from: http://indicators.hst.org.za/indicators/StatsSA/.[Accessed on October 2011].

[5] Reddy AS, Garg SK. Hybrid-genetic algorithm based descriptor optimization and QSAR models for predicting the biological activity of *Tipranavir analogs* for HIV protease inhibition. *J Mol Graph Model* 2010; **23**(8): 852-862.

[6] Hernandez N, Kiralj R, Ferreira MMC, Talavera I. Critical comparative analysis, validation and interpretation of SVM and PLS regression models in a QSAR study on HIV-1 protease inhibitors. *Chemometr Intell Lab Syst* 2009; **98**(1): 65-77.

[7] Maillard A, Chapplain JM, Tribut O, Bentué-Ferrer D, Tattevin P, Arvieux C, et al. The use of drug resistance algorithms and genotypic inhibitory quotient in prediction of lopinavir and ritonavir treatment response in human immunodeficiency virus type 1 protease inhibitor-experienced patients. *J Clin Virol* 2007; **38**(2): 131-138.

[8] Stuermer M, Dauer B, Haberl A, Mueller A, Gute P, Klauke S, et al. Drug resistance to Tipranavir (TPV) or Darunavir (DRV) according to new interpretation algorithms in PI-naive HIV-1 infected patients. *Antiviral Res* 2007; **74**(3): A29-A29.

[9] Ravela J, Betts BJ, Brun-Vézinet F, Vandamme AM, Descamps D,Van Laethem K, et al. Rapid communication: HIV-1 protease and reverse transcriptase mutation patterns responsible for discordances between genotypic drug resistance interpretation algorithms. *J Acquir Immune Defic Syndr* 2003; **33**: 7-14.

[10]De Luca A, Cingolani A, Giambenedetto, SD. Variable prediction of antiretroviral treatment outcome by different systems for interpreting genotypic human immunodefciency virus type 1 drug resistance. *J Infect Dis* 2003; **187**(12): 1934-1943.

[11]De Luca A, Cozzi-Lepri A, Perno CF. Variability in the interpretation of transmitted genotypic HIV-1 drug resistance and prediction of virological outcomes of the initial HAART by distinct systems. *Antiviral Ther* 2004; **9**(5): 743-752.

[12]Wang D, Larder B, Revell A, Montaner J, Harrigan R, De Wolf F, et al. A comparison of three computational modeling methods for the prediction of virological response to combination HIV therapy. *Artif Intell Med* 2009; **47**: 63-74.

[13]Vergne L, Snoeck J, Aghokeng A, Maes B, Valea D, Delaporte E, et al. Genotypic drug resistance interpretation algorithms display high levels of discordance when applied to non-B strains from HIV-1 naive and treated patients. *FEMS Immuno Med Micro* 2006; **46**: 53-62.

[14]Snoeck J, Kantor R, Shafer RW. Discordances between interpretation algorithms for genotypic resistance to protease and reverse transcriptase inhibitors of human immunodeficiency virus are subtype dependent. *Antimicrob Agents Chemother* 2006; **50**: 694-701.

[15]Vercauteren J, Vandamme J. Algorithms for the interpretation of HIV-1 genotypic drug resistance information. *Antiviral Res* 2006; **71**: 335-342.

[16]Poonpiriya V, Sungkanuparph S, Leechanachai P. A study of seven rule based algorithms for the interpretation of HIV-1 genotypic resistance data in Thailand. *J Virol Methods* 2008; **151**: 78-86.

[17]Yebra G, de Mulder M, del Romero J, Rodríguez C, Holguín A. HIV-1 non-B subtypes: High transmitted NNRTI-resistance in Spain and impaired genotypic resistance interpretation due to variability. *Antiviral Res* 2010; **85**: 409-417.

[18]Goumas SK, Dimou IN, ZervakisME. Combination of multiple classifiers for post-placement quality inspection of components: A comparative study. *Inform Fus* 2010; **11**(2): 149-162.

[19]Abate AF, De Marsico M, Ricciardi S, Riccio D. Normal maps *vs.* visible images: Comparing classifiers and combining modalities. *J Visual Lang Compt* 2009; **20**(3): 156-168.

[20]Park SC, Pu J, Zheng B. Improving performance of computer-aided detection scheme by combining results from two machine learning classifier. *Acad Radiol* 2009; **16**(3): 266-274.

[21]Pang Y, Shang M, Yuan Y, Pan J. Scale invariant image matching using triplewise constraint and weighted voting. *Neurocomput* 2012; **83**: 64-71.

[22]Yuan Y, Pang Y, Wang K, Shang M. Efficient image matching using weighted voting. *Pat Recog Let* 2011; **33**(4): 471-475.

[23] Singhal S, Jaiswa P. Presentation of tuberculosis in TB-HIV co-infection patients and the treatment outcome with directly observed short course therapy. *Asian Pac J Trop Biomed* 2011; **1** (supp2): 266-267

[24]Salzberg S. On comparing classifiers: Pitfalls to avoid and a recommended approach. *Data Mining Knowl Disc* 1997; **1**: 317-328.

[25]Momoh MA, Muhamed U, Agboke AA, Akpabio EI, Osonwa UE. Immunological effect of aqueous extract of Vernonia amygdalina and a known immune booster called immunace® and their admixtures on HIV/AIDS clients: a comparative study. *Asian Pac J Trop Biomed* 2012; **2**(3): 181-184.

[26] Osahon AI, Ukponmwan CU, Uhunmwangho OM. Prevalence of HIV seropositivity among patients with squamous cell carcinoma of the conjunctiva. *Asian Pac J Trop Biomed* 2011; **1**(2): 150-153.