A PILOT STUDY TO INTEGRATE HIV DRUG RESISTANCE GOLD STANDARD INTERPRETATION ALGORITHMS USING NEURAL NETWORKS

Y. SINGH AND M. MARS

Abstract

Introduction: There are several HIV drug resistant interpretation algorithms which produce different resistance measures even if applied to the same resistance profile. This discrepancy leads to confusion in the mind of the physician when choosing the best ARV therapy.

Aim: The aim of this study is to combine the interpretation of three gold standard interpretation algorithms using multilayer perceptron neural networks in order to produce a single resistance measure.

Methods: The REGA, HIVdb and ANRS algorithms were applied to protease and reverse transcriptase genome sequences. This output together with the relevant IC50 based resistance measures were fed into three different multilayer perception neural networks.

Results: The three neural network models increased the average accuracy of the protease sequence from 61% to 76.3% and 58% to 63% for reverse transcriptase sequences.

Conclusion: Results show that combining the international gold standards using neural networks produces better ARV drug resistance prediction models.

Keywords: Machine learning, Artificial intelligence, Neural networks, HIV drug resistance

1. INTRODUCTION

1.1 Background

The human immunodeficiency virus (HIV) is an enveloped virus that belongs to the family of retroviridae and to the genus lentivirus. This means that the viral infection is characterized by a long latent period, and thus accounts for the HIV's ability to infect and destroy the immune system of a human over a long period of time. There are two known strains of the HIV, i.e. HIV-1 and HIV-2 (Wilson,2002). Functionally, the HIV-2 is not as pathogenic as the HIV-1, meaning that the rate of replication and infection of the HIV-2 is substantially slower than that of the HIV-1. The HIV-1 accounts for the vast majority of HIV infections in the world, and it is estimated that 95% of all infection are due to the HIV-1 (Quinn,1998).

There are currently almost 5.6 million infected with HIV in South Africa, which is approximately 11% of the South African population (AIDS Committee of Actuarial Society of South Africa, 2008). It is also estimated that there are almost 500 000 patients who exhibit AIDS defining conditions in South Africa (Health Systems Trust, 2011). The prevalence of HIV in some countries as well as the prevalence in the nine provinces of South Africa is shown in Table One [Health Systems Trust, 2011; Central Intelligence Agency, 2011). The high prevalence of HIV in Africa, and the contrast between developing and developed countries is clearly seen.

Table 1: Prevalence of HIV for each country as reported by the Central Intelligence Agency of the United States (Central Intelligence Agency, 2011), and for each province in South Africa as per Health Systems Trust (Health Systems Trust, 2011).

Country	Prevalence %	Country	Prevalence %
Swaziland	25.90	France	0.40
Botswana	24.80	Spain	0.40
Lesotho	23.60	Italy	0.30
Zimbabwe	14.30	India	0.30
Zambia	13.50	Netherlands	0.20
Namibia	13.10	United Kingdom	0.20
Mozambique	11.50	Australia	0.10
Malawi	11.00	Germany	0.10
Uganda	6.50	Greece	0.10
Kenya	6.30	New Zealand	0.10
Tanzania	5.60	Afghanistan	0.01
Nigeria	3.60	Svalbard	0.00
South Africa - Kwazulu-Natal	14.9	South Africa - Eastern Cape	10.6
South Africa - Mpumalanga	12.5	South Africa - Limpopo	7.0
South Africa - North West	12.4	South Africa - Northern Cape	6.7
South Africa - Free State	12.0	South Africa - Western Cape	5.6
South Africa - Gauteng	11.2		

There is currently no cure or publically available vaccine for HIV infection. However HIV may be managed by highly active antiretroviral therapy. Highly active antiretroviral therapy comprises of a regimen of three antiretroviral (ARV) drugs from at least two of the possible five drug classes, namely, protease inhibitors, reverse transcriptase inhibitors, integrase inhibitors, fusion inhibitors, and entry inhibitors. Each ARV class works by inhibiting different stages in the replication cycle of HIV from the attachment to CD_4 cells to the cleavage and subsequent formation of viron.

Several factors contribute to the failure of highly active antiretroviral therapy. These include poor treatment, stage of the disease, drug potency, patient adherence, achievable drug levels, drug resistance (Richman et al., 2000), and toxic effects of the drugs. Of these factors, drug resistance is arguably the most critical. Drug resistance is defined as the inability of antiretroviral drugs to reduce the HIV viral load below detectable levels i.e. significantly inhibit HIV replication. It was reported in 2009, that 37% of patients that required ARV treatment actually received ARV drugs in South Africa (Adam et al., 2009). It is inevitable that drug resistance will become a concern in the treatment of HIV/AIDS infected patients due to the high replication rates of the virus, selective pressure caused by the ARV drugs and initial infection by resistant strains of HIV.

Phenotypic assays are a direct in-vitro method of measuring ARV drug resistance. A phenotypic test comprises of cultivating the HIV that has infected a patient in a sterile and controlled laboratory environment. The concentration of a specific ARV to reduce the replication rate of the HIV by 50 % in the culture is compared to the concentration of ARV required to reduce the replication rate of the wild-type HIV by 50% in a similar environment. The ratio of these two concentrations is called the IC_{50} score. The IC_{50} score is compared to predetermined literature based susceptible cutoffs to determine if a particular patient is resistant to ARVs as shown in Equation One.

$$\text{Resistance to ARV} = \begin{cases} \text{Susceptible}, & IC_{50} < \text{Susceptible cutoff} \\ \text{Intermediate}, & \text{Susceptible cutoff} \le IC_{50} \le \text{Resistant cutoff} \\ \text{Resistant}, & IC_{50} > \text{Resistant cutoff} \end{cases} \quad (1)$$

Phenotypic tests are relatively expensive, time consuming, susceptible to error and each test detects resistance to a single drug and thus many tests are required to determine multiple drug resistances (Bartlett et al., 2004). Electronic computerized algorithms (Jaideep et al., 2003) 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 which ARV drugs the patients are resistant to. These computer based tests are faster and cheaper than phenotypic tests.

Computer based gold standard interpretation 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. During treatment of complex ARV resistance these different international gold standards are a source of discrepancy and create confusion in the mind of the physician in terms of choosing the best ARV therapy.

Thus the aim of this study is to combine the interpretation of three gold standard interpretation algorithms using multilayer perceptron neural networks in order to produce a single resistance measure and to determine if the combined algorithm better predicts HIV drug resistance.

The next section describes the three international gold standards and how multilayer perceptron neural networks work.

1.2. Interpretation algorithms

One type of interpretation algorithm 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 gold standards in ARV drug resistance interpretation algorithms. REGA was developed by the laboratory for clinical and evolutionary virology, Rega Institute for Medical Research, Katholieke Universiteit Leuven. The HIV-db program was developed by the Division for Infectious Diseases, Stanford University Medical Center, Stanford University. The French National Agency for AIDS Research AC11 Resistance group developed the ANRS algorithm.

These interpretation algorithms were developed using different datasets, subtypes, and are analyzed on drug-naive and -experienced patients. These differences have led to the creation of different interpretation algorithms. Initial 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 suggest some discordance between interpretation algorithms [Jaideep et al, 2003; de Luca, 2003; de Luca, 2004; Wang et al., 2011; Verge et al, 2006; Snoek et al., 2006; Vercauteren, 2006; Poonpiriya, 2008; Yebra, 2010).

1.3. Multilayer Peceptron Neural Network

All neural networks contain an electronic representation of the human neuron. These networks interconnect these neurons in various manners in order to perform learning. Similar to the dentrites attached to the axon of the human neuron, the electronic neuron has channels which transport input (α) into the core of the neuron. The electronic neuron has weights (ϖ) and an external stimuli called the bias (β). The electronic neurons may or may not fire depending on the value of the product (θ) and summation of the inputs, weights and bias, and on the type of activation function (Ψ) that is used. Mathematically a neuron can be described as shown in Equations Two and Three, and graphically as shown in Figure One.

$$\vartheta = \sum_{i=0}^{\text{num of inputs}} \alpha_i \varpi + \beta \qquad (2)$$

$$\psi(\vartheta) = \begin{Bmatrix} \text{Fire} \\ \text{Dont Fire} \end{Bmatrix} \qquad (3)$$

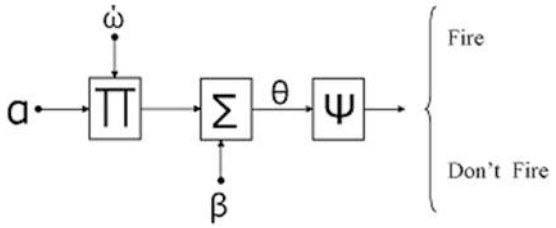

Figure 1: Shows the interaction between inputs, bias and weights.

The simple perceptron can only perform linear separation. In most pattern classification problems non-linear separation is required and this is achieved by layering perceptrons into a multilayer perceptron (MLP). These layers are divided into input, hidden and output layers as indicated by Figure Two. Each neural network neuron receives input from the neural network neurons in the previous layer and is then multiplied by weights. The weighted inputs are summed, and passed through an activation function which scales the output to either being 0 or 1. This scaled value is then broadcast to each of the neurons in the next layer. To use the network to solve a problem, the input space values are applied to the inputs of the first layer, and the signals are allowed to propagate through the network, eventually predicting the output values. These perceptrons are independent of each other, and are thus trained as such.

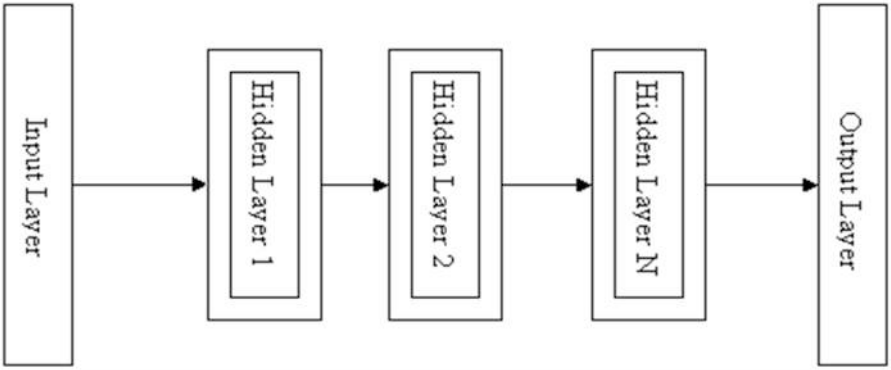

Figure 2: Shows the structure of an MLP with input, hidden and output layers.

2. METHODOLOGY

2.1 Data Processing

Genotype-Phenotype datasets that consisted of 2,928 protease genes and 1,981 reverse transcriptase gene sequences were obtained from the Stanford HIV drug resistance database (http://hivdb.stanford.edu/). This database contains publically available de-identified HIV drug resistance data. The amino acid for each protease and reverse transcriptase sequence in the dataset was then processed into its respective three base nucleotide sequence. The nucleotide list was then fed into to the online HIValg V6.0.11 program hosted by Stanford University. This web application takes as input the nucleotide list, converts this to an amino acid list, determines mutations that occur in the genome, and calculates a resistance measure, namely, resistant, susceptible and intermediate, for a sequence by applying the REGA, HIVdb and ANRS algorithms individually to it. These resistance measures then form the input to the machine learning algorithm.

The predicted output resistance measure 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 and these acted as the output of the machine learning algorithm during the learning step. This process is shown in Figure Three.

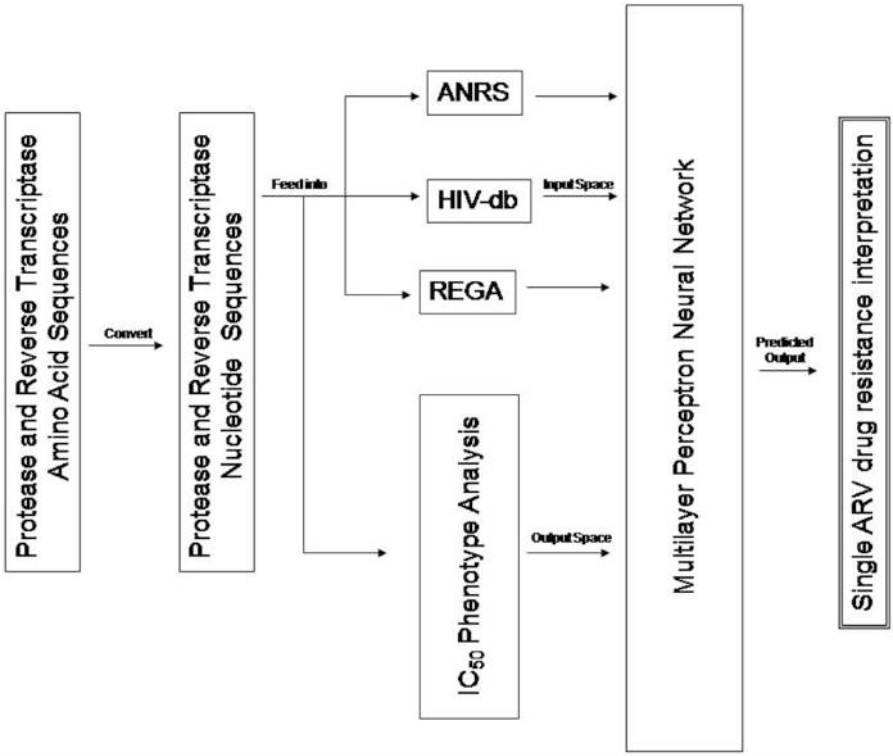

Figure 3: Shows the steps involved in the methodology employed to obtain a single ARV drug resistance interpretation.

2.3. Neural Network

Three multilayer perception neural networks called MLPNN1, MLPNN2, and MLPNN3 respectively were created using NeuroSolutions (http://www.neurosolutions.com/). MLPNN1 consisted of one hidden layer with 50 processing elements. MLPNN2 consisted of two hidden layers with 10 and 5 neurons respectively. MLPNN3 comprised of three hidden layers with 10, 15 and 5 neurons per layer respectively. All three neural networks used a momentum learning rule, a step size of 0.1, and a momentum of 0.7. The multilayer perceptron neural networks are shown in Figures Five-Seven.

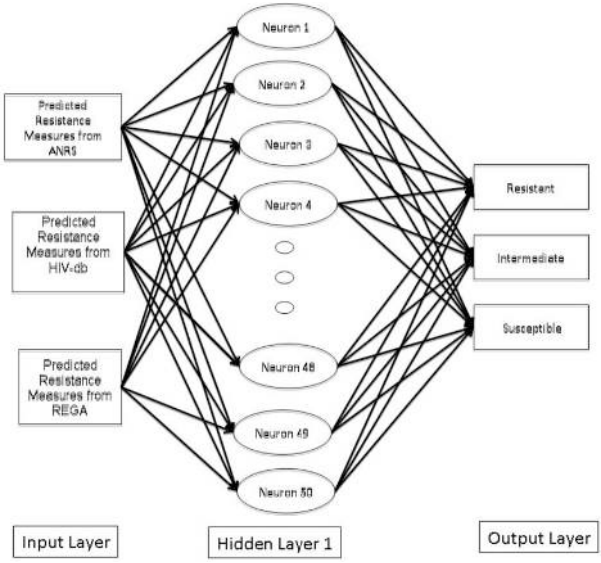

Figure 5: Shows the structure of the MLPNN1, with one hidden layer and 50 neurons.

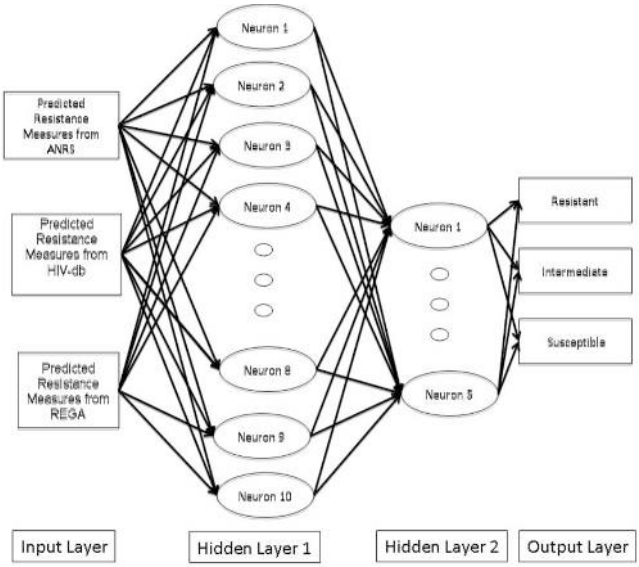

Figure 6: Shows the structure of the MLPNN2, with two hidden layers consisting of 10 and 5 neurons.

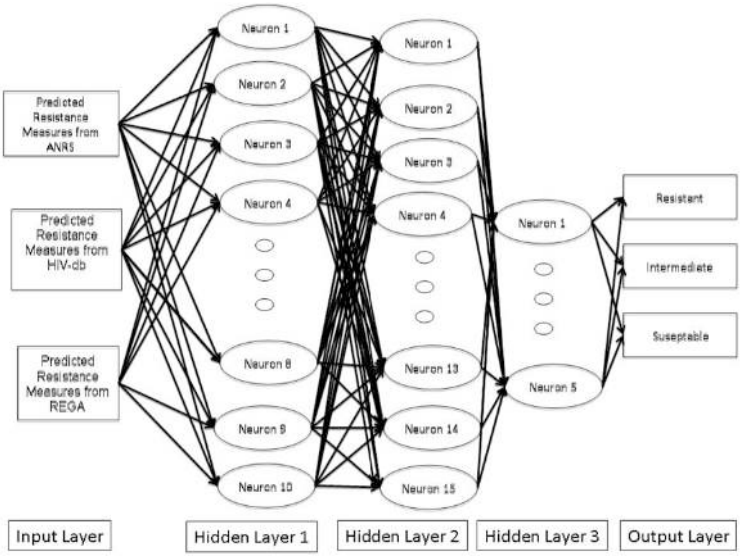

Figure 7: Shows the structure of the MLPNN3, with three hidden layers consisting of 10, 15 and 5 neurons.

2.4 Statistical Analysis

Standard deviation, z-score and p-score were calculated in order to perform a proportion Z-test. The standard deviation was calculated as shown in Equation Four and the Z-score as calculated from Equation Five. P-scores were determined by solving Equation Six. These were calculated to determine if the accuracies obtained were statistically different from chance. Similarly the proportional Z-test was used to determine if the difference between the algorithms were statistically significant.

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

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

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

3. RESULTS

In order to determine which of the three neural networks best classifies the resistance profiles, the accuracies, average sensitivities, specificities, positive predictive value and negative predictive value for the protease (Table Two) and reverse transcriptase sequences (Table Three) were calculated. Accuracy is defined as the percentage of resistance profiles that have resistance measures as identified by the IC_{50} score.

Sensitivity for the resistance measure was determined by calculating the number of predicted patient's resistance profiles that was correctly predicted as resistant compared to all patient resistance profiles that should have been predicted as resistant. These were calculated for the intermediate and the susceptible measures as well. The overall sensitivity was found by averaging the three sensitivities for the resistance, intermediate and susceptibility measures.

Specificity for the resistance measure was determined by calculating the number of predicted patient's resistance profiles that was correctly predicted as not resistant compared to all patient resistance profiles that should have been predicted as not begin resistant. The overall specificity was found by averaging the three specificities for the resistance, intermediate and susceptibility measures.

Positive predictive value is defined as the average proportion of patients with resistance profiles correctly diagnosed as resistant, susceptible and intermediate. Negative predictive value was similarly calculated.

Table 2: shows accuracies, sensitivities, specificities, positive predictive value and negative predictive value for the protease sequences using the ANRS, HIV-db and REGA algorithms.

Measure	ANRS	Hiv-Db	REGA	MLPNN1	MLPNN2	MLPNN3
Accuracy/%	62	62	61	75	77	77
Sensitivity/%				92	91	92
Specificity/%				78	78	81
PPV/%				73	75	78
NPV/%				93	93	93

Table 3: shows accuracies, sensitivities, specificities, positive predictive value and negative predictive value for the reverse transcriptase sequences using the ANRS, HIV-db and REGA algorithms and the multilayer perceptron models.

Measure	ANRS	Hiv-db	REGA	MLPNN1	MLPNN2	MLPNN3
Accuracy/%	57	58	59	63	63	63
Sensitivity/%				66	65	67
Specificity/%				75	76	76
PPV/%				81	82	82
NPV/%				58	56	63

Table 4: shows the Z-scores when comparing accuracies, sensitivities, specificities, positive predictive value and negative predictive value for the protease and reverse transcriptase sequences against chance using the ANRS, HIV-db and REGA algorithms. (* indicates p <0.001)

Measure	Protease Sequences			Reverse Transcriptase Sequences		
	MLPNN1	MLPNN2	MLPNN3	MLPNN1	MLPNN2	MLPNN3
Accuracy	27.1*	29.2*	29.2*	11.6*	11.6*	11.6*
Sensitivity	45.5*	44.4*	45.5*	14.2*	13.4*	15.1*
Specificity	30.3*	30.3*	33.5*	22.3*	23.1*	23.1*
PPV	27.1*	24.9*	30.3*	27.6*	28.5*	28.5*
NPV	46.5*	46.5*	46.5*	7.1*	5.3*	11.6*

In order to determine which of the three neural network algorithms performed the best in terms of most effectively modeling the domain, each of the algorithms were compared against each other using the proportional Z-test as per Equations 4 to 6. These results are reflected in Table 5. Furthermore to determine if the difference between the best neural network and the three international gold standards are significant, the proportional z test was performed. These results are shown in Table 6.

Table 5: shows the Z-scores when comparing accuracies, sensitivities, specificities, positive predictive value and negative predictive value for the protease and reverse transcriptase sequences against chance. (* indicates p <0.001)

	Protease Sequence			Reverse transcriptase Sequence		
	MLPNN1		MLPNN2	MLPNN1		MLPNN2
	MLPNN2	MLPNN3	MLPNN3	MLPNN2	MLPNN3	MLPNN3
Accuracy	2.1*	2.1*	0 (p=1)	0 (p=1)	0 (p=1)	0 (p=1)
Sensitivity	-1.7 (p=0.045)	0 (p=1)	-1.6 (p=0.54)	-0.92 (p=0.17)	0.93 (p=0.17)	-1.9 (p=0.028)
Specificity	0 (p=1)	3.2*	-3.4*	1.1 (p=0.13)	1.0 (p=0.15)	0 (p=1)
PPV	-2.3*	3.0*	5.3*	1.1 (p=0.13)	1.1 (p=0.13)	0 (p=1)
NPV	0 (p=1)	0 (p=1)	0 (p=1)	-1.8 (p=0.035)	4.4*	-6.4*

Table 6: shows the Z-scores when comparing accuracies of ANRS, HIV-db and REGA against MLPNN3 the protease and reverse transcriptase sequences. (* indicates p <0.001)

Gold Standard	Protease	Reverse Transcriptase
ANRS	13.6*	5.3*
HIV-db	13.6*	4.4*
REGA	14.4*	3.6*

4. ANALYSIS

For the protease sequence, the three neural network models produced an average accuracy of 76.3 ± 1.1% (95% CI), sensitivity of 91.6 ± 0.5% (95% CI), specificity of 79 ± 1.7% (95% CI), PPV of 75.3 ± 2.5 %(95% CI), and NPV of 93 ± 0% (95% CI). For the reverse transcriptase sequence, the three neural network models produced an average accuracy of 63 ± 0% (95% CI), sensitivity of 66 ± 1% (95% CI), specificity of 75.6 ± 0.6% (95% CI), PPV of 81.6 ± 0.6% (95% CI), and NPV of 59 ± 3.6% (95% CI). Table Four shows that the MLPNN1, MLPNN2, and MLPNN3 predicted ARV drug resistance measures more effectively than chance as shown by the p<0.001 for all of the accuracy, sensitivity, specificity, PPV and NPV's for both the protease and reverse transcriptase sequences in Table Four. This indicates that the three machine learning algorithms have created a mathematical model that emulates the domain knowledge and that the predication is not based on random chance.

Table Five shows that there is statistical difference between the MLPNN1 and MLPNN2 algorithms in terms of accuracy and positive predictive value with the protease sequences, and from Table 4 it may be deduced that MLPNN2 better predicts ARV drug resistance measures. Similarly, there MLPNN3 has a statistically significantly higher accuracy, sensitivity and PPV than MLPNN1. MLPNN3 also has a higher specificity and PPV than MLPNN2. Thus one may deduce that MLPNN3 better predicts ARV drug resistance measures using protease sequences.

Table 5 also shows that there is statistical difference between the MLPNN1 and MLPNN3 algorithms in terms of negative predictive value with the reverse transcriptase sequences, and from Table 4 it may be deduced that MLPNN3 better predicts ARV drug resistance measures. Similarly, the MLPNN3 has a statistically significantly higher NPV than MLPNN2. Thus one may deduce that MLPNN3 better predicts ARV drug resistance measures using reverse transcriptase sequences.

It is evident from the $p<0.001$ for all comparisons shown in Table 6, that combining the international gold standards using neural networks produces a statistically significantly better ARV drug resistance prediction model.

5. CONCLUSIONS

Results from this study show that combining the international gold standards using neural networks produces a statistically significantly better ARV drug resistance prediction model. Further research should be directed to improving the effectiveness of the combination by using other machine learning techniques and expanding the data set in terms of more predictors and tuples of data.

6. REFERENCES

Adam, M.A. and L.F. Johnson, Estimation of adult antiretroviral treatment coverage in South Africa. . South African Medical Journal 2009. 99(9): p. 6.

AIDS Committee of Actuarial Society of South Africa, ASSA2008 Model: ProvOutput. 2008

Bartlett J.C. and Gallant J.E, Medical management of HIV infection, J.E. Jho Hopkins University School of Medicine, 2004.

Central Intelligence Agency USA, CIA World Factbook 2011, HIV/AIDS adult prevalence rate, publisher: Directorate of Intelligence, ISBN: ISBN 978-1616080471

De Luca, A., A. Cingolani, and S. Di Giambenedetto, Variable prediction of antiretroviral treatment outcome by different systems for interpreting genotypic human immunodedciency virus type 1 drug resistance. Journal of Infectious Disease, 2003. 187(12): p. 1934-1943.

De Luca, A., A. Cozzi-Lepri, and C.F. Perno, Variability in the interpretation of transmitted genotypic HIV-1 drug resistance and prediction of virological outcomes of the initial HAART by distinct systems. Antiretroviral Therapy, 2004. 9(5): p. 743-752.

Health Systems Trust. Aids sick: Indicator data. Health Indicators 2011 October 2011; Available from: http://indicators.hst.org.za/healthstats/85/data.

Jaideep, R. and others, Rapid communication: Hiv-1 protease and reverse transcriptase mutation patterns responsible for discordances between genotypic drug resistance interpretation algorithms. Journal of Acquired Immune Deficiency Syndromes, 2003. 33: p. 8-14.

Poonpiriya, V., S. Sungkanuparph, and P. Leechanachai, A study of seven rule based algorithms for the interpretation of HIV-1 genotypic resistance data in Thailand. Journal of Virological Methods, 2008. 151: p. 78-86.

Richman D. and Staszewki S., A practical guide to HIV drug resistance and its implications for antiretroviral treatment strategies, International Medical Press, 2000.

Snoeck, J., R. Kantor, and R.W. Shafer, Discordances between interpretation algorithms for genotypic resistance to protease and reverse transcriptase inhibitors of human immunodeÞciency virus are subtype dependent. Antimicrobial Agents and Chemotherapy 2006. 50: p. 694-701.

Quinn T.C., Molecular variants of HIV-1 and their impact on vaccine development, International Journal of STD and AIDS 9 (1998), no. 1, 2–2.

Vercauteren, J. and A. Vandamme, Algorithms for the interpretation of HIV-1 genotypic drug resistance information. Antiviral Research, 2006. 71: p. 335-342.

Vergne, L., J. Snoeck, and A. Aghokeng, Genotypic drug resistance interpretation algorithms display high levels of discordance when applied to non-B strains from HIV-1 naive and treated patients. FEMS Immunology and Medical Microbiology, 2006. 46: p. 53-62.

Wang, D., et al., A comparison of three computational modeling methods for the prediction of virological response to combination HIV therapy. Artificial Intelligence in Medicine. 47: p. 63-74.

Wilson D. Naidoo S. Bekker L. Cotton M. Maartens G., Handbook of HIV medicine, Oxford University Press Southern Africa, 2002.

Yebra, G., et al., HIV-1 non-B subtypes: High transmitted NNRTI-resistance in Spain and impaired genotypic resistance interpretation due to variability. Antiviral Research, 2010. 85: p. 409-417.