Poster WePe12.6C04
3rd IAS Conference on HIV Pathogenesis and Treatment
24-27 July 2005, Rio de Janeiro, Brazil

Global neural network models are superior to single clinic models as general quantitative predictors of virological treatment response

A. Revell[1], B.A. Larder[1], D. Wang[1], S. Wegner[2], R. Harrigan[3], J. Montaner[3] and C. Lane[4].

1: The HIV Resistance Response Database Initiative (RDI), London, UK; 2: US Military HIV Research Program, Rockville, MD, USA;
3: The BC Centre for Excellence in HIV/AIDS Vancouver, BC, Canada; 4: National Institute of Allergy and Infectious Diseases (NIAID), Bethesda, MD, USA.

Introduction

- The RDI is developing Artificial Neural Networks (ANN) that predict virological response to combination antiretroviral therapy from genotype
- Previous RDI studies suggested that 'local' models trained with data from a single clinic may be more accurate for patients from that clinic than models trained with additional data from 2-3 other clinics
- Here we compare the accuracy of local models with 'global' models, trained with data from over 200 clinics, for patients from the local clinic and patients from an entirely independent clinic.

Methods

Three-layer ANN were developed using back propagation: input layer (baseline variables), a hidden layer (interconnections) and an output layer (virological response) as follows

Treatment change episodes (TCEs) were extracted from the RDI database. These consisted of the following input variables for the ANN models:

- Baseline viral load (log_{10} copies/ml)
- Baseline mutations (55 resistance mutations)
- Drugs in new regimen
- Time to follow-up viral load

The output variable that the ANN were trained to predict was the virological response: i.e. change in viral load from baseline (ΔVL).

40 and 38 TCEs from clinics A (NIAID) and B (US Military HIV Research Program), were randomly selected to act as local and independent tests data sets.

Ten local ANN models were trained with the remaining data from clinic A (337 TCEs) and 10 global models trained with 3,168 TCEs from >200 clinics, including the 337 TCEs from clinic A. Training was performed as follows:

1. The TCEs were partitioned 10 times into 90% for training and 10% for validation such that each TCE appeared in a validation set once
2. For each of these 10 training sets, 1800 ANN models were developed using the training data and different parameters (learning rate, error thresholds, number of nodes in hidden layer, maximum iteration number).
3. These models were then provided the input variables from the validation set and produced predictions of the output variable - ΔVL. The most accurate model was selected from the 1800 for each partition.
4. By repeating steps 2 and 3 10 times a 'committee' of 10 ANN models was derived.

The performance of the local and global committees was tested by providing the models with the input variables of the test data. The predictions of the 10 models in a committee were averaged for each test TCE to produce the committee average prediction. These predictions were compared to the actual ΔVL for the test TCEs to assess the accuracy of the ANN committees' predictions.

Finally the local and global models were used to identify potentially effective alternative regimens for 19 cases of actual treatment failure following treatment change on the basis of a genotype with rules-based interpretation in the NIAID clinic. Failure was defined as:

- Baseline VL < 3 logs.
- Increase in VL following treatment change (to a regimen containing ≥3 drugs).

The local and global committee predictions were obtained for the top 100 most frequently used 3-5 drug combinations in the RDI database. The best predicted response was recorded in each case.

Results

Test set A (local NIAID clinic)

1. Correlations (Pearson Product Moment) between the predicted and actual ΔVL gave r^2 values of 0.78 ($p<0.00001$) and 0.70 ($p<0.00001$) for the local and global models respectively (Figure 1).
2. The mean absolute difference between the predicted and actual VL was 0.49 vs 0.55 logs for the local and global models respectively.
3. The mean percentage correct VL trajectory predictions was 90% for both ANN committees.

There were no significant differences in performance between the two committees

Test set B (independent clinic)

1. Correlations (Pearson Product Moment) between the predicted and actual ΔVL gave r^2 values of 0.07 (ns) and 0.23 ($p<0.01$) for the local and global models respectively (Figure 2). This difference was statistically signficant ($p<0.01$).
2. The mean absolute differences between the predicted and actual VL were 0.97 vs 0.66 logs for the local and global models respectively ($p<0.05$).
3. The mean percentage correct VL trajectory predictions was 76% vs 89% for the local and global models respectively ($p<0.05$).

Figure 1: Scatterplots of predicted vs actual ΔVL for local and global ANN committees with a local clinic test set

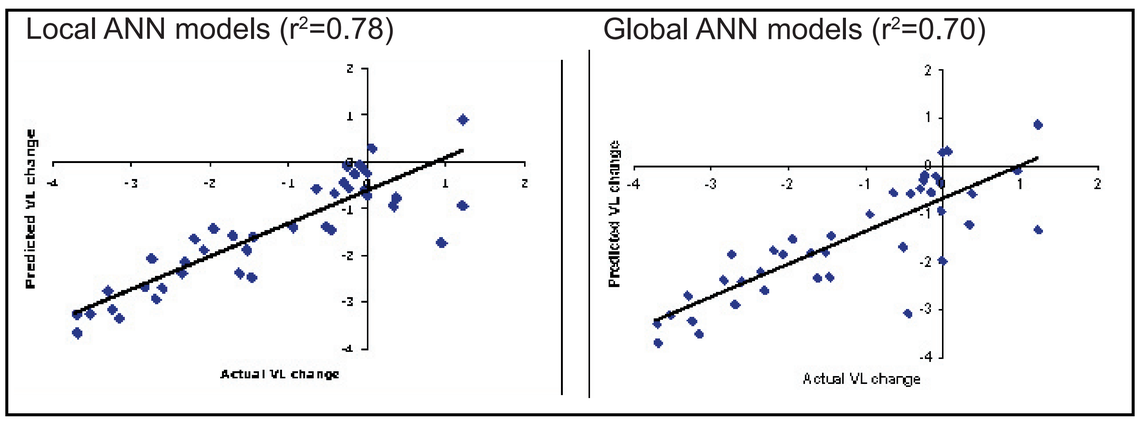
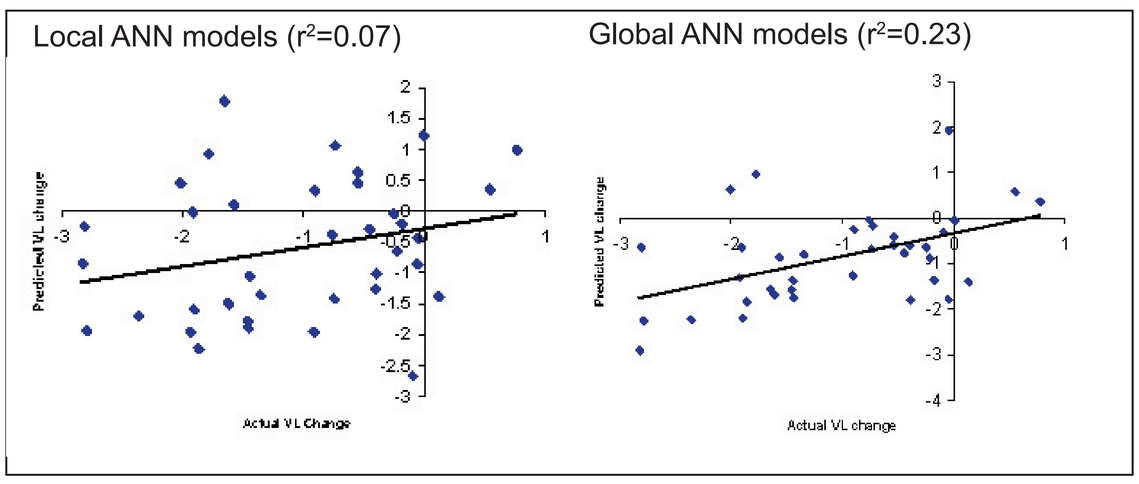


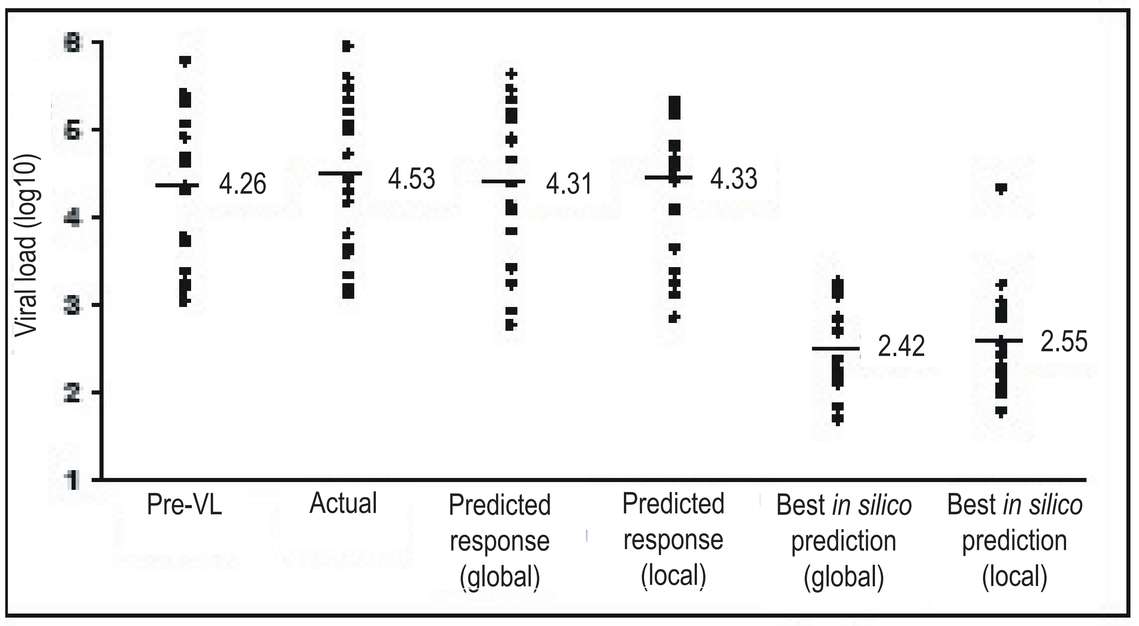
Figure 2: Scatterplots of predicted vs actual ΔVL for local and global ANN committees with an independent clinic test set



The majority of the global models' predictions were within 1 log of the actual ΔVL. The three outlier predictions that were ≥ 1 log different were all cases where the models predicted a drop in VL and the actual change was an increase, possibly due to non-adherence to potentially effective regimens.

Both sets of ANN models successfully identified alternative regimens that were predicted to be effective for the 19 treatment failures (Figure 3). The mean reduction in VL predicted for the best alternative was 1.71 logs for the local and 1.83 logs for the global ANN committees respectively (ns).

Figure 3: Best predicted virological responses with alternative treatments for 19 treatment failures



Conclusions

Global ANN models can perform as accurately as local models, trained with data from a single clinic, in predicting virological response to combination antiretroviral therapy for patients from that clinic

Global models appear superior to local models for predicting virological response to therapy for patients from other clinics, suggesting that they may be the most powerful way to exploit ANN as a generally applicable treatment decision-making tool.

ANN models are able to identify alternative treatment regimens that are predicted to result in virological response in cases of actual virological failure

Acknowledgements

The RDI thanks:

- NIAID and the US Military HIV Research Program and the BC Centre for Excellence in HIV/AIDS for their support and collaboration with this research
- All those centres around the world that have contributed data
- The patients

This project has been funded with Federal Funds from the National Cancer Institute, National Institutes of Health, under contract No. NO1-CO-12400 and from the US Military HIV Research Program (under the Army Cooperative Agreement No. W81XWH-014-2-0005).