Global Neural Network Models Are Superior to Single Clinic Models as General Quantitative Predictors of Virologic Treatment Response

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

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

Poster 53

14th International Drug Resistance Workshop
Quebec, Canada; June 7-11, 2005

Introduction

Previous RDI studies suggested that 'local' Artificial Neural Network (ANN) models, trained with data from a single clinic, may more accurately predict virological response to combination antiretroviral therapy from genotype for patients from that clinic than models trained with additional data from two to three other centres. Here we compare local models with 'global' models, trained with data from over 200 clinics, in terms of their accuracy in predicting virological response to therapy among patients from the local clinic and patients from an entirely independent clinic.

Methods

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

- Baseline viral load
- Baseline mutations
- Drugs in new regimen
- Time to follow-up viral load

The output variable was the virologic response (ΔVL).

40 and 38 TCEs from clinics A (NIAID) and B (DoD), were randomly selected from the RDI database to act as test data sets. Ten local ANN models were trained with the remaining data from clinic A (337 TCEs) and ten global models with 3,168 TCEs from >200 clinics, including the 337 from clinic A. Training was performed as follows:

1. The TCEs were partitioned ten times into 90% for training and 10% for validation such that each TCE appeared in a validation set once.
2. For each training set 1800 ANN models were developed using different ANN parameters (learning rate, error threshold, number of hidden units, maximum iteration number). These models were provided the input variables from the validation set and produced predictions of virologic response. The most accurate model was selected from each partition.
3. By repeating step 2 ten times a 'committee' of ten ANN models was derived.

The performance of the local and global ANN committees was tested by providing the models the input variables of the test data. Each ANN model in the committee produced a prediction of virologic response for each test TCE. The predictions of the ten models in a committee were averaged for each TCE to produce the committee average prediction (CAP). These CAP predictions were compared to the actual follow-up viral loads for the test TCEs to assess the accuracy of the ANN committee.

Finally, the local and global models were used to identify potentially effective alternative regimens for 19 cases of actual treatment failure from the NIAID clinic. Failure was defined as a baseline VL of <3 logs, with increasing VL after treatment change to a regimen containing 3 or more drugs. The local and global CAPS were obtained for the top 100 most frequently used 3-5 drug combinations in the RDI database. The best predicted response was recorded.

Results

Test set A (local, NIAID clinic)
Correlations (Pearson Product-Moment) between the predicted and actual ΔVL gave r^2 values of 0.78 *vs* 0.70 for local and global models respectively (Figure 1). The mean absolute difference between predicted and actual VL was 0.49 vs 0.55 logs (ns). The mean percentage correct VL trajectory prediction was 90% for both ANN committees.

Test set B (independent , DoD clinic)
Correlations between predicted and actual ΔVL gave r^2 values of 0.07 *vs* 0.23 for the local and global models respectively (Figure 2). The mean absolute differences between predicted and actual VL were 0.97 *vs* 0.66 logs ($p<0.05$) and the mean percentage correct VL trajectory predictions were 76% and 89% for the local and global models respectively .

The majority of global models' predictions were within 1 log of the actual VL. The three outlier predictions that were ≥2 logs different from actual VL were all cases where the models predicted a drop in VL and the actual change was an increase, possibly due to non-adherence to effective regimens.

Figure 1: Correlations between predicted and actual ΔVL for local and global ANN models tested with local clinic data

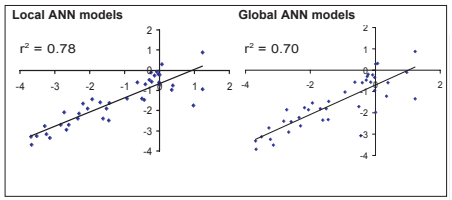
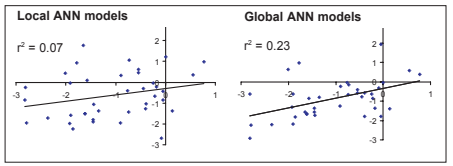
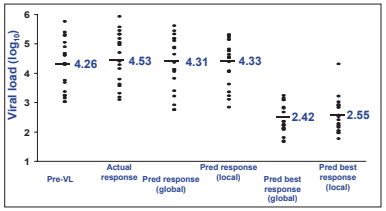


Figure 2: Correlations between predicted and actual ΔVL for local and global ANN models tested with independent clinical data



Both sets of ANN models successfully identified alternative regimens for the 19 treatment failures that were predicted to be effective (Figure 3). The mean reduction in VL predicted for the best alternative was 1.71 logs and 1.83 logs for the local and 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 patients from other clinics, suggesting that they may be the most powerful way to exploit ANN as a generally available treatment decision-making tool.

Acknowledgments

The RDI thanks NIAID, the US Military HIV Research Program and the BC Centre for Excellence for their support and collaboration with this research, and all those centres around the world that have contributed data.

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).