Identification of Clinical Data Effect on Aplastic Anemia Treatment Using Multilayer Fuzzy Neural Network

Hironori Mutoh\textsuperscript{1}\h041318@mbox.nagoya-u.ac.jp  Seiji Kojima\textsuperscript{2}\kojimas@med.nagoya-u.ac.jp

Hiroyuki Honda\textsuperscript{1}\honda@nubio.nagoya-u.ac.jp

\textsuperscript{1} Department of Biotechnology, School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
\textsuperscript{2} Department of Developmental Pediatrics, Nagoya University, Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan

\textbf{Keywords}: Aplastic Anemia, fuzzy neural network, multilayer

\section{Introduction}
Acquired aplastic anemia (AA) is an uncommon but serious disorder characterized by pancytopenia resulting from nonfunction of the bone marrow. Although bone marrow transplantation (BMT) from a human leukocyte antigen (HLA)-identical sibling donor is the treatment of choice, this approach is limited by the availability of such donors. Immunosuppressive therapy (IST) has been an alternative treatment for patients who do not have suitable donors \cite{1}. Because of unwieldiness and intractableness of biological data, traditional statistical methods haven’t performed well on the analysis. Machine learning algorithms carry the potential for mining significant information from complicated data. Here, we applied one of machine leaning algorithms, Fuzzy neural network (FNN) to AA data.

\section{Materials and Method}
\subsection{Subjects}
Patients with acquired AA were eligible if they met the following criteria: age younger than 18 years, recently diagnosed disease without specific prior treatment, and moderate to very severe AA. Three hundred and twenty eight children with newly diagnosed AA were entered in the study and their response at 3 and 6 month after the treatment was examined. This data was obtained by running nationwide clinical trial for examination of drug mediation effects on AA treatment. Response level was partitioned into 3 levels including complete response (CR), partial response (PR) and non response (NR). Three levels were defined by neutrophil counts, platelet counts and hemoglobin levels. We analyzed only the patients had NR (169 samples) or PR (119 samples) at 3 month because the patient who had CR at 3 month remained CR at 6 month. Nine parameters, age, sex, grade of severity, duration of the disorder, neutrophil counts, platelet counts, reticulocyte counts, medication of Danazol and medication of G-CSF (granulocyte colony-stimulating factor) were used as input variables. The response at 6 month was used as an output variable.

\subsection{FNN model and Parameter selection}
The FNN model is a relatively advanced artificial neural network (ANN) model. The FNN model has the advantages of extremely high prognostic accuracy and the ability to explicitly describe causality between input and output variables as linguistic IF-THEN rules \cite{2}. In this study we also applied multilayer FNN. Multilayer FNN (ML FNN) uses the output value of standard FNN (ST FNN) as an input valuable \cite{3}. All of the input variables are normalized as average is equal to 0.0 and standard deviation is equal to 1.0. If the response at 6 month was better than that of 3 month, the output value was set to -1.0. If not, the output value was set to 1.0. We applied exhaustive combination search method to detect the parameters which affect the response to the treatment. In the exhaustive combination search method, the all possible FNN models with 2 input variables combination were constructed and ranked by the accuracy of the models.
3 Result and Discussion

All possible 2 inputs ST FNN models ($C_2=36$) were constructed using exhaustive combination search method. We separately analyzed patients who had NR (NR group) and PR (PR group) at 3 month after the treatment. The accuracies of the top ranked models in two groups are shown in Fig. 1. For prediction of PR group, the model with higher accuracy was obtained.

The IF-THEN rules extracted from those top ranked models are shown in Fig. 2. According to the rules in PR group, short duration of disorder might give worse response in younger patients than older patients ($P$-value $= 0.0014$). It might be also said that non-mediation of Danazol gives worse response in patients who had small platelet counts ($P$-value $= 0.0040$). Grade of severity is examined by a neutrophil, platelet and reticulocyte counts. Therefore this result implies that the response to the treatment can be related to the severity.

We also applied multilayer FNN model to PR group. Using exhaustive combination search method, 3 inputs ML FNN models including 2 inputs top.

4 References

