Prediction of Effect of Pegylated Interferon Alpha-2b plus Ribavirin Combination Therapy in Patients with Chronic Hepatitis C Infection

Tetsuro Takayama, Hirotoshi Ebinuma, Shinichiro Tada, Yoshiyuki Yamagishi, Kanji Wakabayashi, Keisuke Ojiro, Takanori Kanai, Hidetsugu Saito, Toshifumi Hibi*, for the Keio Association for the Study of Liver Diseases

Division of Gastroenterology and Hepatology, Department of Internal Medicine, Keio University School of Medicine, Shinjuku, Tokyo, Japan

Abstract

Treatment with pegylated interferon alpha-2b (PEGIFN) plus ribavirin (RBV) is standard therapy for patients with chronic hepatitis C. Although the effectiveness, patients with high titres of group 1b hepatitis C virus (HCV) respond poorly compared to other genotypes. At present, we cannot predict the effect in an individual. Previous studies have used traditional statistical analysis by assuming a linear relationship between clinical features, but most phenomena in the clinical situation are not linearly related. The aim of this study is to predict the effect of PEG IFN plus RBV therapy on an individual patient level using an artificial neural network system (ANN). 156 patients with HCV group 1b from multiple centres were treated with PEGIFN (1.5 μg/kg) plus RBV (400–1000 mg) for 48 weeks. Data on the patients’ demographics, laboratory tests, PEGIFN, and RBV doses, early viral responses (EVR), and sustained viral responses were collected. Clinical data were randomly divided into training data set and validation data set and analyzed using multiple logistic regression analysis (MLRs) and ANN to predict individual outcomes. The sensitivities of predictive expression were 0.45 for the MLRs models and 0.82 for the ANNs and specificities were 0.55 for the MLR and 0.88 for the ANN. Non-linear relation analysis showed that EVR, serum creatinine, initial dose of Ribavirin, gender and age were important predictive factors, suggesting non-linearly related to outcome. In conclusion, ANN was more accurate than MLRs in predicting the outcome of PEGIFN plus RBV therapy in patients with group 1b HCV.

Introduction

Chronic hepatitis C (CHC) is of global concern because CHC patients frequently develop liver cirrhosis and hepatocellular carcinoma (HCC). Eradication of the hepatitis C virus (HCV) is an effective means of preventing CHC. Pegylated interferon alpha-2b (PEGIFN) plus ribavirin (RBV) combination therapy against the HCV is currently standard therapy for patients with CHC. Although this combination is effective against certain types of HCV, it is effective only in 50–60% of patients infected with the IFN-resistant strain of HCV [1]. HCV genotype 1 is common in the United States [2], Europe, and Japan. In Japan, 70% of CHC patients are infected with HCV genotype 1b [2–6]. The treatment outcome of patients infected with HCV genotype 1b is poor compared to other genotypes and the virus is eradicated from only 50% of these patients [7–11].

Although prolonged treatment with an elevated dose of RBV increases the efficacy of PEGIFN plus RBV treatment [12], the response rate is still relatively low. Furthermore, indices for determining whether to continue or stop treatment are lacking. Seventy-five % of patients treated with IFN experience systemic side-effects [1], the treatment of which adds to the cost and duration of IFN treatment. Therefore, it is important to identify factors predictive of treatment efficacy. Early viral response (EVR), a 2-log decrease in the serum HCV RNA level 12 weeks after commencing therapy, is a useful predictive factor. We also have demonstrated host and viral predictive factors [13–15].

Current guidelines recommend that treatment be discontinued for patients who do not achieve viral clearance from sera until 24 weeks after commencing therapy [1]; however, only 50–70% of patients achieve EVR [1]. Moreover, it is recommended that the decision to discontinue treatment should be made on an individual basis according to the patient’s tolerance of therapy and biochemical or viral responses to treatment [1].

Previous studies, which typically used linear discriminant analysis provided the significant factors, though were unable to predict treatment outcomes at the level of the individual patient. Many clinical analyses have employed classical linear methods even though most data obtained in clinical settings are confounded and variables are not linearly related. A recent study demonstrated that the kinetics of most phenomena in living organisms are non-linear [16]. For these reasons, most data derived from clinical
epidemiological or statistical studies are inappropriate for predicting responses at the level of the individual [16].

Artificial neural networks (ANNs) do not suffer from the problems inherent in traditional prediction methods. An ANN is a learning system based on a computational technique and has been used to simulate the neurophysiological processing ability of the human brain [17]. ANNs recognize complex patterns between inputs and outputs via the learning process. Once the hidden relationship between input and output has been learned, an ANN can correctly predict output from a given input [18,19]. ANNs are considered more suitable than MLRs for solving problems of the non-linear type and for analyzing complex datasets [20–24]. Notably, ANNs can provide conclusive predictions at the individual level [16].

Previous reports have demonstrated that ANNs are superior to classical linear methods in the prediction of responses to interferon-α and RBV [20,21,23–26]. It is unclear whether the results of classical linear studies are representative of clinical conditions because all genotypes and a high number of responders were included in these studies. Moreover, liver biopsy results were often used as input data in classical linear studies. Although liver biopsies are useful, the procedure is associated with a high degree of risk and a large sampling error [27]. Alternative non-invasive and low-cost predictive methods are required.

The aims of this study were to develop a new model for predicting responses to PEGIFN plus RBV combination therapy in CHC patients infected with HCV genotype 1b by using clinical and laboratory data and an ANN and to identify factors that have non-linear relationships with responses.

Results

Response rate and patient backgrounds

One hundred and fifty-six patients (101 men and 55 women; mean age, 57.6 years; range of age, 18–77 years) received PEGIFN plus RBV therapy (table 1). Of the 156 patients, 66 patients (42.3%) achieved SVR.

Input factors and Outcome

We used the clinical data to determine input factors \(X_1 \sim X_{21}\), which were used to predict the outcomes of individual patients using MLR and ANN analysis (table 2). \(X_1 \sim X_{10}\) represented the patient’s sex, age, height, and weight, respectively. \(X_{11} \) and \(X_{15}\) represented previous treatment with interferon and interferon plus RBV, respectively. \(X_{12} \sim X_{17}\) represented the initial doses of PEGIFN and RBV, respectively. \(X_{18} \sim X_{21}\) represented laboratory variables \((X_6, \text{ white blood cell count}; X_{10}, \text{ red blood cell count}; X_{11}, \text{ haemoglobin level}; X_{12}, \text{ platelet count}; X_{13}, \text{ serum AST level}; X_{14}, \text{ serum ALT level}; X_{15}, \text{ serum creatinine level}; X_{16}, \text{ serum total cholesterol level})\). \(X_{17}\) represented the presence of diabetes mellitus, \(X_{18}\) represented the HCV RNA level, \(X_{19}\) represented the total amount of administered PEGIFN and RBV, respectively, and \(X_{20}\) represented EVR, defined as the a 2-log decrease in the serum HCV RNA 12 weeks after therapy began. The outcome was SVR, which was determined 24 weeks after cessation of therapy.

Significant factors for the prediction of SVR

For the prediction of SVR, factor \(X_{21}\) (EVR) was highly significant and had a high \(\chi^2\) value \((p<0.0001; \text{table 3})\). Factor \(X_{20}\) (total amount of RBV administrated) was the next most effective factor \((p<0.05)\). Factors \(X_1\) (sex), \(X_9\) (serum creatinine level), \(X_{15}\) (ALT level), and \(X_{19}\) (presence of diabetes mellitus) were the next most effective factors, but their regression coefficients were not statistically significant. Other factors had little effect on the response to therapy.

Non-linear relation exists between input factors and SVR

Next, we generated the predictive expression by using MLR and ANN to predict outcomes from multiple factors as determined by the aforementioned tests. We randomly divided the whole data into training data set for generation of the predictive expressions and validation data set to evaluate their accuracy (table 4). As shown in table 4, there were no significant difference in all factors between training data set and validation data set. The sensitivities were 0.45 in MLR and 0.82 in ANN (table 5). The specificities were 0.55 in MLR and 0.88 in ANN. The low frequency of both sensitivity and specificity in MLR and improved in ANN suggest that a non-linear relationship exists between inputs and outcomes. We also conducted ROC curve analysis to evaluate the accuracy of each prediction. To validate propriety, we analysed the ROC using validation data without training data. The area under the curves of the ROCs (AUROCs) of MLR was 0.662 and the mean AUROCs for ANNs was 0.884 (figure 1).

To evaluate the superiority in prediction of ANN, we randomly divided the all data into training data and validation data for 4
more times and evaluated the accuracy with the same way. In all trials, the correct answer rate and AUROCs for the ANNs were significantly greater than that for MLRs (p < 0.05) (table 6, 7 and figure 2, 3).

### Table 2. Factors and outcomes used to predict individual patient outcomes.

| Factor                                                                 | Outcome                                                                 |
|------------------------------------------------------------------------|------------------------------------------------------------------------|
| $X_1$: 1 = Male, 2 = Female                                            | SVR 24 weeks after commencement of treatment 0 = no, 1 = yes           |
| $X_2$: Age, 57.6 ± 10.3                                                |                                                                        |
| $X_3$: Height (cm)                                                     |                                                                        |
| $X_4$: Weight (kg)                                                     |                                                                        |
| $X_5$: Previous therapy with interferon 0 = no, 1 = yes                |                                                                        |
| $X_6$: Previous therapy with interferon plus RBV 0 = no, 1 = yes       |                                                                        |
| $X_7$: Initial dose of PegIFN (µg)                                     |                                                                        |
| $X_8$: Initial dose of RBV (mg)                                        |                                                                        |
| $X_9$: Serum creatinine level (g/dL)                                   |                                                                        |
| $X_{10}$: WBC level (/mL)                                              |                                                                        |
| $X_{11}$: RBC level (x10^6/mL)                                         |                                                                        |
| $X_{12}$: Haemoglobin (g/dL)                                           |                                                                        |
| $X_{13}$: Platelets (x10^7/µL)                                         |                                                                        |
| $X_{14}$: AST (IU/L)                                                   |                                                                        |
| $X_{15}$: ALT (IU/L)                                                   |                                                                        |
| $X_{16}$: Diabetes mellitus 0 = no, 1 = not determined, 2 = yes        |                                                                        |
| $X_{17}$: Serum total cholesterol level (mg/dL)                        |                                                                        |
| $X_{18}$: HCV RNA level before treatment (kIU/mL)                      |                                                                        |
| $X_{19}$: Total amount of PegIFN administered (µg/kg/week)             |                                                                        |
| $X_{20}$: Total amount of RBV administered (mg/kg/day)                 |                                                                        |
| $X_{21}$: EVR 0 = no, 1 = yes                                          |                                                                        |

### Table 3. Results of multiple logistic regression analysis.

| Factor                                                                 | Regression coefficient | Standard error | $\chi^2$ value | p value |
|------------------------------------------------------------------------|------------------------|----------------|----------------|---------|
| Constant                                                               | 3.8391                 | 8.2907         | 0.21           | 0.6433  |
| $X_1$                                                                  | 0.9712                 | 0.4989         | 3.79           | 0.0516  |
| $X_2$                                                                  | 0.0250                 | 0.0263         | 0.90           | 0.3429  |
| $X_3$                                                                  | −0.0117                | 0.0474         | 0.06           | 0.8045  |
| $X_4$                                                                  | 0.0334                 | 0.0437         | 0.59           | 0.4439  |
| $X_5$                                                                  | 0.0349                 | 0.2550         | 0.02           | 0.8911  |
| $X_6$                                                                  | 0.0216                 | 0.3366         | 0.00           | 0.9489  |
| $X_7$                                                                  | 0.0118                 | 0.0254         | 0.21           | 0.6438  |
| $X_8$                                                                  | −0.0046                | 0.0031         | 2.24           | 0.1344  |
| $X_9$                                                                  | −3.6052                | 2.0527         | 3.08           | 0.0790  |
| $X_{10}$                                                               | −0.0001                | 0.0002         | 0.24           | 0.6227  |
| $X_{11}$                                                               | 0.0155                 | 0.0100         | 2.39           | 0.1221  |
| $X_{12}$                                                               | −0.3129                | 0.3349         | 0.87           | 0.3503  |
| $X_{13}$                                                               | −0.0389                | 0.0549         | 0.50           | 0.4789  |
| $X_{14}$                                                               | 0.0173                 | 0.0149         | 1.35           | 0.2452  |
| $X_{15}$                                                               | −0.0151                | 0.0086         | 3.10           | 0.0785  |
| $X_{16}$ (0)                                                           | 0.3585                 | 0.4015         | 0.80           | 0.3719  |
| $X_{16}$ (1)                                                           | −1.1350                | 0.6219         | 3.33           | 0.0680  |
| $X_{17}$                                                               | −0.0024                | 0.0088         | 0.08           | 0.7812  |
| $X_{18}$                                                               | 0.0007                 | 0.0002         | 0.72           | 0.3970  |
| $X_{19}$                                                               | 0.5059                 | 0.9554         | 0.28           | 0.5965  |
| $X_{20}$                                                               | −0.2572                | 0.1136         | 5.12           | 0.0236  |
| $X_{21}$                                                               | 1.4625                 | 0.2725         | 28.80          | <0.0001 |

Relative weights of the input factors

We analyzed relative weights of the input factors to identify factors that had a significant effect (including both linear and non-linear relationship) on the result of ANN (figure 4). Relative
weights of the input factors determine how the result changes when the test factor ($X_{\text{test}}$) is excluded. An $X_{\text{test}}$ value greater than 1 indicates that it improves the expression, and a value less than 1 indicates that it does not improve the expression. We analysed the value of all networks and determined the corresponding means and standard deviations. $X_{21}$: EVR was the most important predictive factor in every trial. The means of $X_1$: gender and $X_2$: age, $X_3$: height, $X_5$: previous therapy with interferon, $X_8$: initial dose of Ribavirin, $X_9$: serum creatinine, $X_{15}$: ALT, $X_{16}$: Diabetes mellitus, $X_{18}$: HCV RNA level before treatment were also more than 1, which indicates that they have non-linear relationships with response to therapy.

Impact of post-treatment factors in prediction of treatment

We also tried to predict the effect only using pre-treatment parameters (without using post-treatment parameters: $X_{19}$: Total amount of PEGIFN administered, $X_{20}$: Total amount of RBV administered, and $X_{21}$: EVR), though the sensitivity and specificity were low (MLR; sensitivity 0.45, specificity 0.49, ANN; sensitivity 0.59, specificity 0.71) (table 8).

Discussion

Interactions between clinical, genetic, and environmental factors may affect the efficacy of PEGIFN plus RBV combination therapy.
therapy in CHC patients and should be taken into account by physicians when interpreting indications for therapy. Although there are reports on predictors of the response to treatment against the HCV [28,29], data derived from clinical epidemiology studies and medical statistics do not always result in correct predictions at the level of the individual patient. For instance, both male sex and low total cholesterol level are considered indicative of a good prognosis [28], but the prognosis for a male who also has a high total cholesterol level is unknown. In contrast, ANNs can identify relationships within a patient’s clinical data that may be overlooked when classical linear approaches are used [16]. MLRs are powerful tool to find significant factors and provide the key factors in present our study, though it is not suit to predict the results by using factors non-linearly correlate. Because ANNs are trained using existing data, they are more capable of providing correct answers for individual patients. The ANN also has theoretical advantages over conventional MLRs. Unlike MLRs, ANN can predict both linear and non-linear phenomena and can analyse relationships between different variables at different levels [25].

The incidence of correct answers and the AUC of the MLRs differed greatly from that of the ANN. Moreover, it can say that data used in most previous studies were not validated because input data sets were used to estimate ROCs. Therefore, we used validation data sets to estimate ROCs. If we had used only input data in the ANN, the AUROC would have been equal to 100%, because an ANN can fit input data perfectly. Compared with MLRs, a well-trained ANN can predict both linear and non-linear data.

We note that, although the ANN is a useful model, the network logic of prediction cannot be broken down into simple elements because ANNs process data in a non-linear way [16,18,25,30–32]. We analysed the relative weights of input factors to address this issue. The values of each factor affecting the outcome was analysed (figure 2). EVR was identified as the most important factor. Serum creatinine, initial dose of Ribavirin, gender and age also had high values (figure 2).

Both physicians and patients express concern about the risks associated with treatment because the outcome is difficult to predict at the time decisions are made. The increased demand for individualised treatment necessitates new statistics that can be applied in conjunction with ethical and clinical evidence at the individual level. ANNs also have potential economic benefits in that they reduce unnecessary medical treatment.

A report on the classification of patients was published recently [33]. Although this is a valid strategy, it is difficult to apply under clinical conditions because the ISDR mutant and Th1:Th2 ratio must first be determined. Moreover, there are some conflicting reports on the ISDR mutant [34]. As the aforementioned report did not performed validation, they should not be compared with our results; however, the predictive accuracy of our technique is superior to them.

The predictive expression developed in this study should aid physicians to advise individual patients on whether to continue with PEGIFN plus RBV combination therapy. We also tried to predict the effect only using pre-treatment parameters (table 6). Compare to the table 5, both sensitivity and specificity were dramatically improved by adding post-treatment parameter. Suggesting, post-treatment parameter such as adherence to treatment might affect to the effect of PEGIFN plus RBV combination therapy. As the EVR and total amount of RBV were the most important parameters in our study, the predictive expression could also be used to determine whether to increase the dose of RBV. Because we included the total amount of PEGIFN and RBV in the data sets, the effect of an increased dose can be simulated. Although the magnitude of the dose effect depends on patient’s symptoms and exposure to adverse events, our technique remains a powerful tool for determining the appropriate dose of PEGIFN and RBV.

Although our predictive expression does not predict responses perfectly, our results show that the ANN is a valid method for devising individual treatment regimens in the clinical situation. It is well known that 100% prediction accuracy is impossible to achieve because of random error and multiple biases.

As the outcome of PEGIFN plus RBV treatment may be affected by multiple unknown factors, it is important to update data continuously and to acquire clinical data such as the patient’s demographics, medical history, and laboratory test results. Recent accumulating data revealed the importance of IL28B gene from genome wide study [35–37]. Especially, very recent data clearly showed the significance of SNP rs12979860 in IL28B gene in the prediction of the treatment outcome [38–41]. We could not assess the effect of them in this study since we have not collected those data. Further analyses were needed though it may be improve the accuracy. It is also important to demonstrate that the use of trained ANNs in routine medical practice increases the quality of medical care and reduces costs.

**Methods**

**Patients**

The study was conducted by the Keio Association for the Study of Liver Disease (Supporting Information S1). This study was approved by the Keio University School of Medicine review board and the permission was obtained (ID number 2010-026). One hundred and fifty-six CHC patients (101 men and 55 women; mean age, 57.6 years) infected with genotype 1b HCV and treated with PEGIFN plus RBV combination therapy from December 2004 to May 2007 who had been assessed for sustained viral response (SVR) were enrolled and the data were collected. SVR was defined as an absence of serum HCV RNA 24 weeks after cessation of therapy.

All patients had HCV genotype 1b and HCV RNA levels in excess of 100 kIU/mL as measured by quantitative Cobas Amplicor assays (Roche Diagnostics Co. Ltd, Tokyo, Japan). Exclusion criteria were pregnant women or women of childbearing potential, nursing mothers, male patients whose partner could have become pregnant, anaemia, leucopenia, thrombocytopenia,

| Table 6. The incidence of correct answers (%) provided by multiple logistic regression analysis. |
|-----------------------------------------------|
|Trial 1|Trial 2|Trial 3|Trial 4|
|Correct answers (%)|72.6|64.5|75.8|67.7|

doi:10.1371/journal.pone.0027223.t006

| Table 7. The incidence of correct predictions (%) provided by the ANN. |
|-----------------------------------------------|
|ANN 1|ANN 2|ANN 3|ANN 4|
|Training data set|85.9|83.4|79.6|80.5|
|Validation data set|74.9|75.4|78.2|80.0|

doi:10.1371/journal.pone.0027223.t007
severe dysfunction of organs other than the liver, infection with hepatitis B virus or human immunodeficiency virus, autoimmune hepatitis, primary biliary cirrhosis, and liver dysfunction caused by drugs.

Some of the patients did not undergo a liver biopsy because not all of the centres could perform biopsies. All patients were treated for 48 weeks and were followed up for 48 weeks after treatment.

The purpose of the study and its protocol were explained to all patients and their written, informed consent was obtained.

**PEGIFN plus RBV combination therapy**

PEGIFN-α2b (Schering–Plough K.K., Tokyo, Japan) was administered weekly in doses adjusted for body weight according to the manufacturer’s recommendations in Japan (45 kg or less, 60 μg; 46–60 kg, 80 μg; 61–75 kg, 100 μg; 76–90 kg, 120 μg; and 91 kg or more, 150 μg). RBV (Schering-Plough K.K.) was administered once daily in doses adjusted for body weight according to the manufacturer’s recommendations in Japan (61 kg or less, 600 mg; 61–80 kg, 800 mg; and 81 kg or more, 1000 mg).

The duration of PEGIFN plus RBV therapy was 48 weeks and patients were followed-up for the subsequent 48 weeks. Serum levels of HCV RNA were quantified by amplicor analysis. Blood was analysed at the beginning of treatment and every 4 weeks thereafter.

A questionnaire was used to review demographic data (age, sex, weight, and height), previous treatment, initial dose of PEGIFN, initial dose of RBV, presence of diabetes mellitus, HCV RNA level.
before therapy, amount of PEGIFN and RBV administered, SVR, and serum concentrations of white blood cells (WBCs), red blood cells (RBCs), platelets (Plts), asparate aminotransaminase (AST), alanine transaminase (ALT), creatinine (Cre), and total cholesterol (TC).

As the data were collected from several centres in various prefectures, within-centre bias was excluded.

ANN

To develop the ANN, we used three types of network according to manufacturer’s instruction: multilayer perceptrons (MLPs), radial-basis function networks (RBFs), and linear networks (LINs). Details of the ANN and MLP are provided elsewhere [30]. In brief, a hierarchical ANN consisting of three layers (one input, one hidden, and one output layer) was used to classify the effect as a node in the output layer. MLPs were constructed from three layers (one input, one hidden, and one output layer) to classify effects as a node in the output layer. RBF units respond to the distance of points from the centre. The RBF has a hidden layer of radial units, each of which models a Gaussian response surface. We analyzed the results of 156 patients from multiple centres and formed 100 000 networks.

Training data set and validation data set

We used same training data set for generating the predictive expression by using MLR and ANN, and used validation data set to evaluate the accuracy of the expression generated using training data set.

Data analysis

Accuracy (correct answer rate), sensitivity, and specificity were calculated. Also receiver-operating characteristic (ROC) curves for the MLR and ANN were generated to evaluate their accuracy [25]. Multiple logistic analysis was performed using JMP version 7.0.1 software (SAS Institute Japan, Co., Ltd, Tokyo, Japan) and ANN was analysed using Statistica version 06J software (StatSoft Japan, Co., Ltd, Tokyo Japan).

Relative weights of input factors analysis

The detail of relative weights of input factors analysis (= sensitivity analysis) were described elsewhere [42]. In brief, we analysed relative weights of input factors using a leave-one-input-factor-out (LOFO) in turn with a missing values substitution procedure, which enables predictions to be made in the absence of values for each causal factor, and then assessed effects upon ANN response error. Root mean square error (RMSE) for prediction is defined as:

$$RMSE = \sqrt{\frac{\sum (Y_{actual} - Y_{predicted})^2}{n}}$$

where, $n$ is the number of validation data. $Y_{actual}$ and $Y_{predicted}$ are the outcomes of actual values and predicted ones, respectively. RMSE is an estimate of the typical difference between the predicted and actual values of outcomes. The smaller RMSE the better the prediction accuracy of the models is.

The network original error was accumulated as RMSEoriginal and the network was again used with LOFO data and the error
RMSE_{LOFO} was estimated. Then, the relative weights of input factors was calculated as RMSE_{LOFO}/RMSE_{original}.

Supporting Information

Supporting Information S1 (DOC)

References

1. Strader DB, Wright T, Thomas DL, Seeff LB (2004) Diagnosis, management, and treatment of hepatitis C. Hepatology 39: 1147–1171.
2. Zeln NN, Rakela J, Kravvit EL, Reddy KR, Tominaga T, et al. (1996) Hepatitis C virus genotypes in the United States: epidemiology, pathogenicity, and response to interferon therapy. Collaborative Study Group. Ann Intern Med 125: 634–639.
3. Dushkik G, Main J, Thomas H, Reichard O, Lee C, et al. (1996) Ribavirin treatment for patients with chronic hepatitis C: results of a placebo-controlled study. J Hepatol 25: 591–598.
4. McOmish F, Yap PL, Dow BC, Follett EA, Seed C, et al. (1994) Geographical distribution of hepatitis C virus genotypes in blood donors: an international collaborative survey. J Clin Microbiol 32: 884–892.
5. Noobbaum JB, Pol S, Nalpas B, Landais P, Berthelot P, et al. (1995) Hepatitis C virus type 1b (II) infection in France and Italy. Collaborative Study Group. Ann Intern Med 122: 161–168.
6. Takada N, Takase S, Takada A, Date T. (1995) Differences in the hepatitis C virus genotypes in different countries. J Hepatol 17: 277–283.
7. Berg T, van Wagner M, Naiser S, Sarrazin C, Heintges T, et al. (2006) Extended treatment duration for hepatitis C virus type 1: comparing 48 versus 72 weeks of peginterferon-alfa-2a plus ribavirin. Gastroenterology 130: 1681–1697.
8. Hadiyanius S, Seta H, Jr., Morgan TR, Balan V, Diago M, et al. (2004) Peginterferon-alfa2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. Ann Intern Med 140: 346–357.
9. Zeuzem S (2004) Heterogeneous virologic response rates to interferon-based therapy in patients with chronic hepatitis C: who responds less well? Ann Intern Med 140: 358–367.
10. Jacobson IM, Brown RS, Jr., Freilich B, Afdhal N, Kwo PY, et al. (2007) Peginterferon alfa-2a and ribavirin for the treatment of chronic hepatitis C in previously untreated patients infected with HCV genotypes 2 or 3. J Hepatol 40: 993–999.
11. Jacobson IM, Brown RS, Jr., Freilich B, Althald N, Kwo PY, et al. (2007) Peginterferon alfa-2b and ribavirin for 72 weeks in patients with detectable hepatitis C virus RNA at week 4 of treatment. Gastroenterology 131: 451–460.
12. Nakamura M, Suito H, Ibeda M, Tada S, Kumaagi N, et al. (2000) Possible molecular mechanism of the relationship between NNSB polymorphisms and early clearance of hepatitis C virus during interferon plus ribavirin treatment. J Med Virol 80: 632–639.
13. Saito H, Ehimeh H, Satoj I, Miyagachi S, Tada S, et al. (2000) Immunological and virological predictors of outcome of interferon-alpha therapy of chronic hepatitis C. J Viral Hepat 7: 6–74.
14. Saito H, Saito T, Kawaiashiki K, Nakamoto N, Takahashi A, et al. (2000) The detection of IRF-1 promoter polymorphisms and their possible contribution to T helper 1 response in chronic hepatitis C. J Interferon Cytokine Res 22: 693–700.
15. Grosi E, Buscema M (2007) Introduction to artificial neural networks. Eur J Gastroenterol Hepatol 19: 1046–1054.
16. Takayama K, Fujikawa M, Obata Y, Morishita M (2005) Neural network based optimization of drug formulations. Adv Drug Deliv Rev 57: 1217–1231.
17. Baxt WG (1995) Use of an artificial neural network for the diagnosis of myocardial infarction. Ann Intern Med 115: 843–848.
18. Shirakawa H, Matsumoto A, Joshi S, Komatsu M, Tanaka N, et al. (2008) Pretreatment prediction of virologic response to peginterferon plus ribavirin therapy in chronic hepatitis C patients using viral and host factors. Hepatology 48: 1753–1760.
19. Liu S, Ansari IH, Das SC, Pataki MK (2006) Insertion and deletion analyses identify regions of non-structural protein 5A of Hepatitis C virus that are dispensable for viral genome replication. J Gen Virol 87: 325–327.
20. Grossi E, Ferray J, Thompson A, Simon J, Shianna K, et al. (2009) Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nat Genet 41: 399–401.
21. Tanaka Y, Nishida N, Sugiyama K, Kurokawa M, Matsuoka K, et al. (2009) Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 41: 1105–1109.
22. Lin CY, Chen JY, Lin TN, Jeng WJ, Huang CH, et al. (2011) IL28B SNP rs12979860 is a critical predictor for on-treatment and sustained virologic response in patients with hepatitis C virus genotype-1 infection. PLoS ONE 6: e21854.
23. Ge D, Fellay J, Thompson A, Simon J, Shianna K, et al. (2009) Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nat Genet 46: 399–401.
24. O’Brien TR, Everhart JE, Morgan TR, Lok AS, Chung RT, et al. (2011) An IL28B genotype-based clinical prediction model for treatment of chronic hepatitis C. PLoS ONE 6: e20904.
25. Cyr DD, Lucas JE, Thompson JW, Patel K, Clark PJ, et al. (2011) Characterization of serum proteins associated with IL28B genotype among patients with chronic hepatitis C. PLoS ONE 6: e21854.
26. McCarthy JJ, Li JH, Thompson A, Suchindran S, Lao XQ, et al. (2010) Replicated association between an IL28B gene variant and a sustained response to pegylated interferon and ribavirin. Gastroenterology 138: 2007–2014.
27. Takagaki K, Araki H, Takayama K (2010) Creation of a tablet database containing several active ingredients and prediction of their pharmacological characteristics based on ensemble artificial neural networks. J Pharm Sci 99: 4201–4214.

Acknowledgments

We thank to Mika Inoue individuals for the contribution.

Author Contributions

Conceived and designed the experiments: TT HE ST. Performed the experiments: TT. Analyzed the data: TT. Contributed reagents/materials/analysis tools: TT HE ST YY KW KO TK HS TH. Wrote the paper: TT HE ST.