J Med Syst
DOI 10.1007/s10916-006-9046-8

ORIGINAL PAPER

Efficacy of Interferon Treatment for Chronic Hepatitis C Predicted by Feature Subset Selection and Support Vector Machine

Jun Yang · Anto Satriyo Nugroho · Kazunobu Yamauchi · Kentaro Yoshioka · Jiang Zheng · Kai Wang · Ken Kato · Susumu Kuroyanagi · Akira Iwata

Received: 9 August 2006 / Accepted: 10 November 2006

Abstract Chronic hepatitis C is a disease that is difficult to treat. At present, interferon might be the only drug, which can cure this kind of disease, but its efficacy is limited and patients face the risk of side effects and high expense, so doctors considering interferon must make a serious choice. The purpose of this study is to establish a simple model and use the clinical data to predict the interferon efficacy. This model is a combination of Feature Subset Selection and the Classifier using a Support Vector Machine (SVM). The study indicates that when five features have been selected, the identification by the SVM is as follows: the identification rate for the effective group is 85%, and the ineffective group 83%. Analysis of selected features show that HCV-RNA level, hepatobiopsy, HCV genotype, ALP and CHE are the most significant features. The results thus serve for the doctors' reference when they make decisions regarding interferon treatment.

Keywords Chronic Hepatitis C (CHC) · Interferon (IFN) · Support Vector Machine (SVM) · Feature Subset Selection (FSS) · Predict

J. Yang (✉) · K. Yamauchi · J. Zheng · K. Wang · K. Kato
Department of Medical Information and Management Science, Graduate School of Medicine, Nagoya University, 65, Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan
e-mail: junyang@med.nagoya-u.ac.jp

K. Yamauchi
e-mail: kyamauti@med.nagoya-u.ac.jp

J. Zheng
e-mail: jianzhen@med.nagoya-u.ac.jp

K. Wang
e-mail: wangkai@med.nagoya-u.ac.jp

K. Kato
e-mail: katok@med.nagoya-u.ac.jp

A. S. Nugroho
School of Life System Science and Technology, Chukyo University, Japan
e-mail: nugroho@life.chukyo-u.ac.jp

K. Yoshioka
Hepato-Gastroenterology, School of Medicine, Fujita Health University, Japan
e-mail: kyoshiok@fujita-hu.ac.jp

S. Kuroyanagi · A. Iwata
Department of Computer Science and Engineering, Nagoya Institute of Technology, Japan

S. Kuroyanagi
e-mail: bw@nitech.ac.jp

A. Iwata
e-mail: iwata@nitech.ac.jp

Introduction

Patients with Chronic Hepatitis C (CHC) are at high risk of developing liver cirrhosis and/or liver cancer, and prevention of disease progress is very important. It was first reported in 1986 that interferon was used to treat chronic hepatitis C and showed ALT normalization [1]. After that, some studies [2–4] showed the effects of improvement of liver function and liver histology image, and serum HCV-RNA proved negative after interferon treatment. Now interferon therapy is the treatment of choice for chronic hepatitis C. Combination therapy with pegylated interferon alpha (PEG) and ribavirin is currently the standard treatment for patients with CHC. However, there are numerous side effects such as fatigue, a flu-like syndrome and others. Furthermore, a cost-effectiveness analysis in Japan estimates that the health care costs per patient for standard treatment with interferon ranged from $10,500 to $35,000. On the other hand,

Table 1 Main features of patients *Mean ± SD

Patient features	Effective group ($n = 66$)	Ineffective group ($n = 46$)	Total ($n = 112$)
Age (years)*	49.73 ± 13.76	52.93 ± 11.86	51.04 ± 12.75
Sex (Male/Female)	49/17	31/15	80/32
HCV RNA level (kiu/ml)*	380.98 ± 298.12	735.83 ± 281.29	526.72 ± 334.33
Hepatobiopsy	F0(3); F1(27); F2(21); F3(10); F4(5)	F1(11); F2(11); F3(14); F4(10)	F0(3); F1(38); F2(32); F3(24); F4(15)
HCV genotype	1b(33); 2a(16); 2b(10); type2(6); 3a(1)	1b(37); 2a(3); 2b(5); type2(1)	1b(70); 2a(19); 2b(15); type2(7); 3a(1)
GOT (IU/L)*	68.82 ± 43.93	84.98 ± 51.24	75.46 ± 47.52
GPT (IU/L)*	114.09 ± 92.44	128.57 ± 87.44	120.04 ± 90.30
ALP (IU/L)*	266.48 ± 80.09	331.89 ± 150.44	293.35 ± 120.67
CHE (IU/L)*	0.95 ± 0.21	0.86 ± 0.26	0.91 ± 0.23
Treatment interval (weeks)*	30.09 ± 10.20	27.98 ± 9.76	29.22 ± 10.03
γ-GTP (IU/L)*	79.18 ± 99.45	82.22 ± 78.49	80.43 ± 91.06
Ribavirin	35(+)/31(−)	27(+)/19(−)	62(+)/50(−)
LDH (IU/L)*	181 ± 38.15	193.07 ± 37.82	186.18 ± 32.28
LAP (IU/L)*	62.86 ± 16.00	63.72 ± 19.06	63.21 ± 17.25

combination therapy with interferon alpha and ribavirin for CHC shows a sustained response rate of only about 40% after 24–48 weeks of treatment, and the rate increased to 54–56% even though PEG was used instead [5–8]. Obviously, interferon treatment is expensive and has limited efficacy.

Therefore, predicting the efficacy of interferon is crucial before attempting treatment for CHC patients. Some factors such as HCV genotype and HCV-RNA level have been identified to predict the efficacy of interferon therapy [18, 26]. Recently, a genetic model based on patient genotypes was proposed for the prediction of combination treatment of interferon-alpha and ribavirin [27]. However, the efficacy of interferon is affected not only by the genetic factors of virus, but also by dosage of interferon, treatment Interval and the patient's age, gender, pretreatment histology, baseline ALT, gamma-GTP, iron serum level, and so on. Most of these factors were induced in a multivariate analysis to predict roughly the efficacy of interferon, but the outcome was unsatisfactory [28]. In this study, a simple model was established to predict the efficacy of interferon therapy for CHC patients. The first step for the model is the selection of clinical markers (feature subset) that are the key to the classification task and then assessed by Fisher Criterion to determine the feature significance. The second step is to classify and identify the two (effective/ineffective) groups by means of the Support Vector Machine (SVM) using selected clinical markers. SVM is a linear machine that is trained to find the optimal discrimination hyperplane of a given data set, Basically it works with two groups of problems. In this study, we used 30 kinds of clinical markers to analyse and gained better outcome.

Patient and clinical database

The clinical data of all 112 CHC patients (80 male and 32 female, aged 17 to 72 years, with an average of 51 years old) undergoing therapy with interferon were collected in Nagoya University Hospital, Japan during the period from August 1997 to March 2005. Based on the details of the individual patients, 30 kinds of clinical markers were chosen: Sex, Age, HCV-RAN level, HCV genotype, Hepatobiopsy, Total Protein, albumin/globulin (A/G), glutamic oxalacetic transaminase (GOT), glutamic pyruvic transaminase (GPT), lactate dehydrogenase (LDH), alkaline phosphatase (ALP), gamma glutamyl transpeptidase (γ-GTP), leucine aminopeptidase (LAP), cholinesterase (CHE), GOT/GPT, Total Cholesterol, Total Bilirubin, hepaplastin test (HPT) (%), hepaplastin test (HPT) (hours), prothrombin time (PT) (%), prothrombin time (PT) (hours), activated partial thromboplastin time (APTT) (%), activated partial thromboplastin time (APTT) (hours), red blood cell (RBC), white blood cell (WBC), hemoglobin (Hb), Lymphocyte, Platelet, Treatment Interval (weeks) and Ribavirin. The main features of patients are summarized in Table 1.

A few biomarkers such as thymol turbidity test (TTT) and zinc sulfate turbidity test (ZTT) were not adopted in this model because they were not tested in some patients during the monitoring period.

Part of the database was missing (no observation). To deal with this problem, we imputed these missing data by value 0.

Some attributes in the database were nominal data: HCV genotype and hepatobiopsy. A common approach to deal with nominal data is using "1-of-c coding" (c is the number of category) [25]. However, most of the genotype were

belonging to type 1b and type 2 (type 2a and 2b were considered as the same group with 2), 70 samples were belonging to type 1b, 41 samples were belonging to type 2, and only one was of type 3a. Therefore, for the sake of simplicity, we decided to use one dimensional feature to represent the genotype. Type 1b was assigned by -1, type 2 was assigned by $+1$, and the only one sample of type 3a was assigned by 0. Of course this is not always the optimal choice for each situation, but this is considered to be the best and simplest representation for the data used in current experiment. In the case of "hepatobiopsy," although they were nominal data, they showed transition phase. Thus the values of this attribute could be considered to have order among them. This characteristic motivated us to use numerical representation, by assigning the best condition F0 by 0, F1 by 1, F2 by 2, F3 by 3, and the worst condition F4 by value of 4, respectively.

Classification of interferon treatment efficacy

One hundred and twelve patients were divided into two groups according to the results of HCV-RNA test six months after the end of interferon therapy. The patients with undetectable HCV-RNA were defined as the effective group, and those with detectable HCV-RNA were defined as the ineffective group. Among 112 patients, 66 belonged to the effective group; the remaining 46 belonged to the ineffective group.

Methods

The prediction model is constructed of two parts: Feature Subset Selection (FSS) and Support Vector Machine (SVM) as classifier, as depicted in Fig. 1.

Feature Subset Selection (FSS) is the preprocessing part of the model that works to select features useful for the classification. In this study, the selection was conducted based on the individual advantages of each feature. Fisher criterion was used to measure the significance of each feature [9].

Let us denote the D-dimensional input vector as $\vec{x} = (x_1, x_2, \ldots, x_j, \ldots, x_D)^T$. Where the number of examples belonging to the effective group is n_{+1}, and the examples belonging to the ineffective group is n_{-1}, the mean of j^{th} feature of the effective group is $\mu_{j,+1}$, the mean of j^{th} feature of ineffective group is $\mu_{j,-1}$, and their standard deviations are $\sigma_{j,+1}$ and $\sigma_{j,-1}$, respectively. The significance of each feature x_j is measured by the following equation:

$$F(x_j) = \frac{n_{+1}n_{-1}}{n_{+1}+n_{-1}} \frac{(\mu_{j,+1}-\mu_{j,-1})^2}{n_{j,+1}\sigma_{j,+1}^2 + n_{j,-1}\sigma_{j,-1}^2} \tag{1}$$

This criterion can be interpreted as finding one single feature that best discriminates both of the groups in the feature space. The greater this score is, the better is the discrimination power of the feature. Based on this score, each feature was assigned by rank of significance. The feature selection was conducted by selecting a certain number of features from the top.

The second part of the model is the Support Vector Machine (SVM) [10]. SVM has currently received increasing attention due to its promising performance in many situations. In principal, SVM is a linear classifier that is trained to obtain an optimal classification hyperplane on the feature space. The optimal hyperplane is obtained by maximization of the "margin"; a criterion defined by the distance between the training samples and the hyperplane.

Let us denote each example as $\vec{x}_i \in \Re^d$, $i = 1, 2, \ldots, l$. Where l is the number of examples, and each example is labeled by $y_i \in \{-1, +1\}$; -1 represents the ineffective group and $+1$ the effective group. It is assumed that both the effective and ineffective groups are perfectly separated by a hyperplane in D-dimensional feature space. This hyperplane is represented by $\vec{w}.\vec{x}_i + b = 0$. Examples $\vec{x}_i$ that belong to the ineffective group should satisfy $\vec{w}.\vec{x}_i + b \leq -1$, and those that belong to the effective group should satisfy $\vec{w}.\vec{x}_i + b \geq +1$. The optimal margin is obtained by maximizing the distance between the hyperplane and the closest pattern, which is formulated by $1/\|\vec{w}\|$ ($\|\vec{w}\|$ is the norm of vector $\vec{w}$). This can be formulated as a Quadratic Programming (QP) problem, by minimizing Eq. (2) under constraint (3). Minimize:

$$\|\vec{w}\|^2 \tag{2}$$

Subject to:

$$y_i(\vec{x}_i.\vec{w} + b) - 1 \geq 0, \quad \forall i \tag{3}$$

The solution to this problem can be obtained through the Lagrange multiplier.

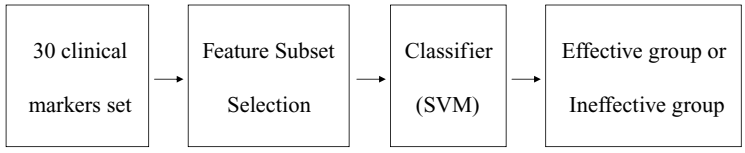

Fig. 1 Predictor model

Table 2 Results of SVM

Dim	SV	Overall ratio of discrimination [%]	Ineffective group ($n = 46$)		Effective group ($n = 66$)	
			Errors	Ratio of discrimination [%]	Errors	Ratio of discrimination [%]
5	70	84	8	83	10	85
10	76	76	18	61	9	86
15	88	81	12	74	9	86
20	85	82	11	76	9	86
25	89	81	13	72	8	88
30	106	78	14	70	11	83

Note. SV (Support Vectors): it shows how many is the support vectors after the training phase completed.

In this study, due to the complexity of the data, we used a non-linear SVM. To work with non-linear problem, the example $\vec{x}$ is mapped onto a higher dimensional feature space by a mapping function $\Phi(\vec{x})$. By this transformation, both groups will become linearly separable in the new feature space. The training phase in SVM is conducted based on the optimization problem as mentioned in the linear SVM. The computation in this optimization must calculate the dot product of two examples in the new feature space, which is denoted by $\Phi(\vec{x}_i).\Phi(\vec{x}_j)$. This computation could be obtained indirectly, without knowing the tranformation function Φ. This strategy is called the *Kernel Trick*. Instead of computing the dot product in the new feature space, it is possible to use the following kernel function, as given by Eq. (4):

$$K(\vec{x}_i, \vec{x}_j) = \Phi(\vec{x}_i).\Phi(\vec{x}_j) \tag{4}$$

The experiments in this study used the Gaussian kernel. The decision function of test sample $\vec{x}$ by non-linear SVM is obtained as follows:

$$f(\Phi(\vec{x})) = \sum_{i=1,\vec{x}_i \in SV}^{l} \alpha_i y_i K(\vec{x}, \vec{x}_i) + b \tag{5}$$

α_i is the Lagrange multiplier corresponding to example $\vec{x}_i$ that takes zero or positive values. SV (Support Vectors) is the subset of training set $\vec{x}_i$ with corresponding $\alpha_i \neq 0$.

In this paper, the classification accuracy of SVM was tested by the leave-one-out cross validation (LOO-CV) method, which can be applied when the samples are small. The procedure consists of picking up one example for testing while the rest of the data are used to train the classifiers, and then testing the removed example. After testing, the classification result is recorded. The process is repeated until all examples have been tested. This method estimates the generalization error of the classifier with a set of tuning parameter β, The tuning parameter set β consisted of : the number of features, and two SVM parameters (soft margin parameter C and Gaussian Kernel function σ). We attempted to evaluate several combination of these three parameters to obtain the best combination for the model. When it is applied to an independent test sample. The final score is obtained by taking the average of the classification rate of each part:

$$\text{LOO} - \text{CV}(\beta) = \frac{1}{l} \sum_{i=1}^{l} \text{correct\_classif\_rate(i, } \beta) \tag{6}$$

In each part of LOO-CV, Fisher Criterion—based Feature Subset Selection (FSS) was applied to the training set, and the significance of the features is ranked based on the score defined by Eq. (1). A number of top-ranked features were selected and used for training by SVM.

We have also attempted the same experiments using *k*-Nearest Neighbour (k-NN) [9] Classifier for comparison with SVM. To classify one example using k-Nearest Neighbour Classifier, first we measure the distance (e.g. Euclidean Distance) between the example from the test set and the whole data of the training set. A number of k examples with the smallest distance are chosen. These data is the nearest neighbours to the test example. The classification of test sample is made by examining the class on the selected k nearest neighbours and taking a vote.

Results

In SVM experiment, When 5, 10, 15, 20, 25 and 30 features were selected, the best identification rate of each dimensionality was presented in Table 2. When five features are selected, the result was optimal: overall identification rate of 84%, identification rate of 85% for the effective group, and identification rate of 83% for the ineffective group.

In *k*-Nearest Neighbor Classifier (k-NN) experiment. When 5, 10, 15, 20, 25 and 30 features are selected, the optimal identification rate by k-NN of each dimensionality is presented in Table 3. The optimal overall identification rate was 81%.

In the study, the rank of the features was further analyzed in each part of the leave-one-out test. The total score of significance is simply defined by the sum of the rank of each feature in each part. The lower score shows the greater the contribution of the feature to the classification task.

Table 3 Results of K-NN

Dim	Overall ratio of Discrimination [%]	Ineffective group ($n = 46$)		Effective group ($n = 66$)	
		Errors	Ratio of Discrimination [%]	Errors	Ratio of Discrimination [%]
5	81	12	74	9	86
10	78	12	74	13	80
15	69	20	56	15	77
20	71	18	61	15	77
25	74	18	61	11	83
30	71	18	61	14	79

HCV-RNA level, hepatobiopsy, HCV genotype, ALP and CHE are chosen in the top rank of the features. The list of significance of the features is presented in Table 4.

Discussion

The hepatitis C virus (HCV) is one of the main reasons for the human hepatitis virus [21]. Hepatitis C readily tends to become chronic and develop into liver cirrhosis and liver cancer. At present, there are about 170,000,000 HCV-infected people, and this number increases by 1/100,000 to 3/100,000 each year [11]. Once they are ill, about 80% will become chronic [12, 13]. Interferon is the most effective drug to treat CHC, but it is very expensive and has various side effects. Therefore, the early prediction of interferon treatment efficacy is very important.

During interferon treatment, the factors that affect its treatment efficacy include HCV-RNA level, HCV genotype, liver damage level, IFN dosage, and treatment interval, among which the HCV-RNA level and HCV genotype are the most important. Many studies reported that a low HCV-RNA level before treatment can be an index for predicting effective of interferon treatment, Interferon treatment efficacy is better for low HCV-RNA patients [14, 24]. Referring to the influence of IFN treatment efficacy by HCV genotype, some research has demonstrated that 1b type efficacy is not as good as those of 1a, 2a, 2b and 3 types [15–18]. The response rate of HCV-1b patients to α—IFN is only 20%–40% whereas that of HCV-2a reaches 70%–80% [19]. As to the liver damage level, the IFN treatment efficacy for patients with liver cirrhosis or obvious fibrosis clearly decreases. Interferon treatment efficacy is better for mild chronic hepatitis, i.e. no fibrosis or liver cirrhosis [20]. In treatment, some studies show that increasing the dosage and treatment course can help eradicate the virus and promote the IFN treatment efficacy [22, 23], The joint use of IFN and ribavirin has cooperative action and better treatment efficacy.

Table 4 shows the list of significance of the features, of which the HCV-RNA level, hepatobiopsy and HCV genotype are the top-ranked features, in agreement with clinical opinion. Treatment interval and ribavirin were expected to rank higher, but it was not reflected in the available dataset. In this study, several types of interferon were used. The interferon dosage cannot be determined each time in comparison with the treatment efficacy, because the patients were not treated with the same type of interferon. Therefore, the total interferon dosage each time is not included in the analysis.

Table 3 and Table 2 show that the difference of overall identification rate between five features and 30 features in *k*-Nearest Neighbor Classifier is 10%, while that of SVM is 6%. The difference of overall identification rate in SVM is less than *k*-Nearest Neighbour. This result shows that SVM is more robust to the existance of irrelevant features, rather than *k*-Nearest Neighbour Classifier. This is because SVM works by mapping the data into new higher dimensional feature space, that makes it easier in finding the discrimination hyperplane on the new space. The results of SVM and *k*-NN

Table 4 List of significance of the features

Rank	Features	Rank	Features	Rank	Features
1	HCV-RNA level	11	A/G	21	Total Cholesterol
2	Hepatobiopsy	12	Treatment interval (weeks)	22	APTT (hours)
3	HCV genotype	13	GPT	23	PT (hours)
4	ALP	14	HPT (hours)	24	Hb
5	CHE	15	PT (%)	25	LAP
6	GOT	16	Sex	26	Total Protein
7	Platelet	17	Ribavirin	27	Lymphocyte
8	GOT/GPT	18	APTT (%)	28	γ-GTP
9	LDH	19	Total Bilirubin	29	HPT (%)
10	Age	20	WBC	30	RBC

Note. $1 \sim 30$ denotes significance rank of the features. 1 represents the most significance.

clearly illustrated that the SVM performance was better than that of *k*-NN.

We observed identification rate of SVM if one of the features is removed. Accordingly, we did the same experiments by using the best 4 features. The result obtained by leave one out cross validation scheme showed that the performance of SVM was overall identification rate of 82 % (identification rate of 85% for the effective group, and identification rate of 78% for the ineffective group, SV: 87 samples). This result showed that taking off one marker worsened the performance of SVM. From this result we concluded that five markers are significant to discriminate the interferon efficacy.

At present, only a few studies have predicted the efficacy of the interferon treatment for CHC patients. In the present study, a different method, which is a combination of the Feature Subset Selection (FSS) and the Support Vector Machine (SVM) was used. System performance was estimated by leave-one-out cross validation. And a higher identification rate of 85% for the effective group, and 83% of for the ineffective group by using the top 5 ranked features was obtained.

In this study, because of the small sample, the reliability of the features selected was limited, Future issues to be addressed include establishing larger clinical databases, reducing the number of support vectors, and finding a better way to assign significance rank to the features for evaluation, in order to obtain the best outcome.

Conclusion

This study clearly showed that a simple model consisting of 5 clinical data with FSS-SVM could identify CHC patients with interferon treatment efficacy (effective group and ineffective group) with a higher degree of accuracy. Thus, the application of this model can be a useful reference for doctors when making decisions regarding interferon treatment.

Acknowledgements J.Y and A.S.N contributed equally to this study. The research of A.S.N is partially supported by a Grant-in-Aid for Private University High-Tech Research Center from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1. Hoofnagle, J. H., Mullen, K. D., Jones, D. B., and Rustgi, V., Treatment of chronic non-A, non-B hepatitis with recombinant human alpha interferon. A preliminary report. *N. Engl. J. Med.* 315:1575–1578, 1986.
2. Davis, G. L., Balart, L. A., Schiff, E. R., and Lindsay, K., Treatment of chronic hepatitis C with recombinant interferon alfa. A multicenter randomized, controlled trial. *N. Engl. J. Med.* 321:1501–1506, 1989.
3. Di Bisceglie, A. M., Martin, P., Kassianides, C., and Lisker-Melman, M., Recombinant interferon alfa therapy for chronic hepatitis C. A randomized, double-blind, placebo-controlled trial. *N. Engl. J. Med.* 321:1506–1510, 1989.
4. Hagiwara, H., Hayashi, N., Mita, E., and Ueda, K., Detection of hepatitis C virus RNA in serum of patients with chronic hepatitis C treated with interferon—alpha. *Hepatology.* 15:37–41, 1992.
5. Lai, M. Y., Kao, J. H., Yang, P. M., Wang, J. T., Chen, P. J., Chan, K. W, Chu, J. S., and Chen, D. S., Long-term efficacy of ribavirin plus interferon alfa in the treatment of chronic hepatitis C. *Gastroenterology*. 111:1307–1312, 1996.
6. Reichard, O., Norkrans, G., Fryden, A., Braconier, J. H., Sonnerborg, A., and Weiland O. Randomised, double-blind, placebo-controlled trial of interferon α-2b with and without ribavirin for chronic hepatitis C. *Lancet*. 351:83–87, 1998.
7. Manns, M. P., McHutchison, J. G., Gordon, S. C., Rustgi, V. K., Shiffman, M, Reindollar, R., Goodman, Z. D., Koury, K., Ling, M. H., and Albrecht, J. K. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet*. 358:958–965, 2001.
8. Fried, M. W., Schiffman, M. L., Reddy, K. R., Smith, C., Marinos, G., Goncales, F. L., Hãussinger, D., Diago, M., Carosi, G., Dhumeaux, D., Craxi, A., Lin, A., Hoffman, J., and Yu, J. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N. Engl. J. Med.* 347:975–982, 2002.
9. Duda, R. O., Hart, P. E & Stork, D. G. Pattern Classification (2nd ed.). Wiley Interscience. New York, NY, 2000.
10. Cristianini, N., and Shawe-Taylor, J. An Introduction to Support Vector Machines (and other kernel-based learning methods). Cambridge University Press. 2000.
11. EASL Intemational Consensus Conference on hepatitis C. Paris, 26-28, February 1999 Consensus statement. *Journal of Hepatology*. 30(5): 956–961, 1999.
12. Kiyosawa, K., Sodeyama, T., Tanaka, E., and Gibo, Y. Interrelationship of blood transfusion, non-A, non-B hepatitis and hepatocellular carcinoma: analysis by detection of antibody to hepatitis C virus. *Hepatology.* 12:671–675, 1990.
13. Tong, M. J., el-Farra, N. S., Reikes, A. R., and Co, R. L. Clinical outcomes after transfusion-associated hepatitis C. *N. Engl. J. Med.* 332:1463–1466, 1995.
14. Kanai, K., Kako, M., Aikawa, T., and Kumada, T. Clearance of serum hepatitis C virus RNA after interferon therapy in relation to virus genotype. *Liver.* 15:185–188, 1995.
15. Kiyosawa, K. The value of hepatitis C virus genotyping to epidemiological and clinical studies. *J. Gastroenterol. Hepatol.* 12:623–624, 1997.
16. Chayama, K., Tsubota, A., Kobayashi, M., and Okamoto, K. Pretreatment virus load and multiple amino acid substitutions in the interferon sensitivity-determining region predict the outcome of interferon treatment in patients with chronic genotype 1b hepatitis C virus infection. *Hepatology.* 25:745–749, 1997.
17. Kanai, K., Kako, M., and Okamoto, H., HCV genotypes in chronic hepatitis C and response to interferon. *Lancet.* 339:1543, 1992.
18. Yoshioka, K., Kakumu, S., Wakita, T., and Ishikawa, T. Detection of hepatitis C virus by polymerase chain reaction and response to interferon-α therapy: relationship to genotypes of hepatitis C virus. *Hepatology.* 16:293–299, 1992.
19. Nousbaum, J. B., Pol, S., Naloas, B., and Landais, P. Hepatitis C virus type 1b(II) infection in France and Italy. *Ann. Intern. Med.* 122:161–168, 1995.
20. Davis, G. L., and Lau, J. Y. Factors predictive of a beneficial response to therapy of hepatitis C. *Hepatology.* 26(suppl):122S–127S, 1997.

21. Kuo, G., Choo, Q. L., Alter, H. J., and Gitnick, G. L., An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science*. 244:362–364, 1989
22. Shiratori, S., Kato, N., Yokosuka, O., and Imazeki, F., Predictors of the efficacy of interferon therapy in chronic hepatitis C virus infection. Tokyo-Chiba Hepatitis Research Group. *Gastroenterology*. 113:558–566, 1997.
23. Poynard, T., Bedossa, P., Chevallier, M., and Mathurin, P., A comparison of three interferon alfa-2b regimens for the long-term treatment of chronic non-A, non-B hepatitis. *N. Engl. J. Med*. 332:1457–1462, 1995.
24. Lau, J. Y., Davis, G. L., Kniffen, J., and Qian, K. P., Significance of serum hepatitis C virus RNA levels in chronic hepatitis C. *Lancet*. 341:1501–1504, 1993.
25. Bishop, C., M., Neural Networks for Pattern Recognition, Oxford University Press, 1995.
26. Takahashi, M., Saito, H., Higashimoto, M., Atsukawa, K., and Ishii, H. Benefit of hepatitis C virus core antigen assay in prediction of therapeutic response to interferon and ribavirin combination therapy. *J. Clin Microbiol*. 43(1):186–191, 2005.
27. Hwang, Y., Chen, E. Y., Gu, Z. J., Chuang, W. L., Yu, M. L., Lai, M. Y., Chao, Y. C., Lee, C. M., Wang, J. H., Dai, C. Y., Shian-Jy Bey, M., Liao, Y. T., Chen, P. J., and Chen, D. S. Genetic predisposition of responsiveness to therapy for chronic hepatitis C. *Pharmacogenomics*. 7(5):697–709, 2006.
28. Kim, S, R., Hayashi, Y., Yoon, S., Taniguchi, M., Yang, M. K., Kim, K. I., Kim, M. M., Saeki, K., Nukata, I., and Imoto, S. Prediction of efficacy of interferon treatment of chronic hepatitis C by multivariate analysis and a new classification. *Pathol Int*. 48(3):215–220, 1998.