

Pretreatment prediction of anemia progression by pegylated interferon alpha-2b plus ribavirin combination therapy in chronic hepatitis C infection: decision-tree analysis

Naoki Hiramatsu · Masayuki Kurosaki · Naoya Sakamoto · Manabu Iwasaki · Minoru Sakamoto · Yoshiyuki Suzuki · Fuminaka Sugauchi · Akihiro Tamori · Sei Kakinuma · Kentaro Matsuura · Namiki Izumi

Received: 16 February 2011 / Accepted: 2 April 2011 / Published online: 17 June 2011
© Springer 2011

Abstract

Background This study aimed to develop a model to predict the development of severe anemia during pegylated interferon alpha-2b plus ribavirin combination therapy.

Methods Data were collected from 1081 genotype 1b chronic hepatitis C patients who were treated at 6 hospitals in Japan. These patients were randomly assigned to a model-building group ($n = 691$) or an internal validation group ($n = 390$). Factors predictive of severe anemia (hemoglobin, Hb < 8.5 g/dl) were explored using data-mining analysis.

Results Hb values at baseline, creatinine clearance (Ccr), and an Hb concentration decline by 2 g/dl at week 2 were

used to build a decision-tree model, in which the patients were divided into 5 subgroups based on variable rates of severe anemia ranging from 0.4 to 11.8%. The reproducibility of the model was confirmed by the internal validation group ($r^2 = 0.96$). The probability of severe anemia was high in patients whose Hb value was <14 g/dl before treatment (6.5%), especially (a) in those whose Ccr was <80 ml/min (11.8%) and (b) those whose Ccr was ≥ 80 ml/min but whose Hb concentration decline at week 2 was ≥ 2 g/dl (11.5%). The probability of severe anemia was low in the other patients (0.4–2.5%).

Conclusions The decision-tree model that included Hb values at baseline, Ccr, and an Hb concentration decline by

N. Hiramatsu and M. Kurosaki contributed equally to this work.

N. Hiramatsu
Department of Gastroenterology and Hepatology,
Osaka University Graduate School of Medicine,
Osaka, Japan

M. Kurosaki · N. Izumi (✉)
Division of Gastroenterology and Hepatology,
Musashino Red Cross Hospital,
1-26-1 Kyonan-cho, Musashino,
Tokyo 180-8610, Japan
e-mail: nizumi@musashino.jrc.or.jp

N. Sakamoto · S. Kakinuma
Department of Gastroenterology and Hepatology,
Tokyo Medical and Dental University, Tokyo, Japan

M. Iwasaki
Department of Computer and Information Science,
Seikei University, Tokyo, Japan

M. Sakamoto
First Department of Internal Medicine,
University of Yamanashi, Yamanashi, Japan

Y. Suzuki
Department of Hepatology, Toranomon Hospital,
Tokyo, Japan

F. Sugauchi
Department of Gastroenterology, Nagoya Koseiin
Medical Welfare Center, Nagoya, Japan

A. Tamori
Department of Hepatology,
Osaka City University Medical School, Osaka, Japan

K. Matsuura
Department of Gastroenterology and Metabolism,
Nagoya City University Graduate School of Medical Sciences,
Nagoya, Japan

2 g/dl at week 2 was useful for predicting the probability of severe anemia, and has the potential to support clinical decisions regarding early dose reduction of ribavirin.

Keywords Data mining · Decision tree · Severe anemia · Chronic hepatitis C · Pegylated interferon · Ribavirin

Abbreviations

Hb	Hemoglobin
PEG-IFN	Pegylated interferon
Ccr	Creatinine clearance
SVR	Sustained virological response
AFP	Alpha-fetoprotein

Introduction

The current standard therapy for chronic hepatitis C is 48 weeks of pegylated interferon (PEG-IFN) plus ribavirin [1]. Sustained virological response (SVR), defined as negative hepatitis C virus (HCV) RNA for 24 weeks after cessation of therapy, can be achieved by the current treatment regimen, but this outcome can be attained in only less than 50% of patients infected with genotype 1 HCV [2, 3]. Hemolytic anemia is a common side effect of ribavirin and is the major reason for dose reduction. Age, gender, baseline platelet level, baseline hemoglobin (Hb) level [4, 5], haptoglobin phenotype [6], drug dose [7], plasma concentration of ribavirin [8], apparent clearance of ribavirin (CL/F) [9], and an early decline in Hb concentration [10, 11] have been reported to contribute to ribavirin-induced anemia. Predicting the possibility of severe anemia before therapy or at the early phase of therapy can help modify ribavirin dosage, decrease the discontinuance rate for ribavirin, and raise the SVR rate.

Data mining is a method of predictive analysis that explores data to discover hidden patterns and relationships in highly complex datasets and enables the development of predictive models. Decision-tree analysis is a core component of data mining and predictive modeling [12], and it is utilized by decision makers in various business fields. Recent publications concerning decision-tree analysis in the medical field indicate its usefulness for defining prognostic factors in various diseases such as prostate cancer [13], diabetes [14], melanoma [15, 16], colorectal carcinoma [17, 18], and liver failure [19]. The results of decision-tree analysis are presented in the form of a flow chart, which is easy to use in clinical practice [20]. This analysis was also used to predict early virological response (undetectable HCV RNA within 12 weeks of therapy) and SVR to PEG-IFN plus ribavirin combination therapy in chronic hepatitis C [21–24]. In the present study, we used decision-tree analysis to explore before- and during-treatment

predictors of severe anemia during PEG-IFN alpha-2b/ribavirin combination therapy and used a prediction algorithm to try to identify chronic hepatitis C patients who are likely to develop severe anemia.

Materials and methods

Patients

This multicenter retrospective cohort study was supported by the Japanese Ministry of Health, Labour and Welfare. Data were collected from 1081 chronic hepatitis C patients who were treated with PEG-IFN alpha-2b plus ribavirin at Osaka University, Musashino Red Cross Hospital, Toranomon Hospital, Tokyo Medical and Dental University, Nagoya City University, Yamanashi University, and their related hospitals. The inclusion criteria applied in the present study were as follows: (1) infection by genotype 1b, (2) HCV RNA ≥ 100 KIU/ml by quantitative PCR (Cobas Amplicor HCV Monitor v 2.0, Roche Diagnostic Systems, CA, USA), (3) lack of co-infection with hepatitis B virus or human immunodeficiency virus, (4) lack of other causes of liver diseases such as autoimmune hepatitis and primary biliary cirrhosis, and (5) completion of at least 12 weeks of therapy. Patients received PEG-IFN alpha-2b (1.5 g/kg) subcutaneously every week and were administered a weight-adjusted dose of ribavirin (600 mg for <60 kg, 800 mg for 60–80 kg, and 1000 mg for >80 kg). The dosage of ribavirin was reduced from 1000 to 600 mg, 800 to 600 mg, or 600 to 400 mg when the Hb concentration decreased to less than 10 g/dl, and was discontinued when the Hb concentration decreased to less than 8.5 g/dl, based on the recommendations in the package inserts. No patient received erythropoietin or blood transfusion for the treatment of anemia. Anemia with Hb < 8.5 g/dl was defined as severe anemia in this study.

For the analysis, patients were randomly assigned to either the model-building ($n = 691$) group or the internal validation ($n = 390$) group. Consent was obtained from each patient. The study protocol conformed to the ethical guidelines of the Declaration of Helsinki and was approved by the institutional review committee. The baseline characteristics and representative laboratory test results are listed in Table 1. There were no significant differences between the clinical backgrounds of the two groups.

Laboratory tests

Blood samples were obtained before therapy and at least once every month during therapy, and were used for hematological tests, blood chemistry analyses, and determination of HCV RNA. Pretreatment levels of HCV RNA

Table 1 Comparison of clinical parameters of model-building and internal validation groups

	All patients (<i>N</i> = 1081)	Model building (<i>N</i> = 691)	Internal validation (<i>N</i> = 390)
Age (years)	55.6 ± 10.5	55.6 ± 10.8	55.6 ± 10.4
Gender (male/female)	612/469	393/298	219/171
Body mass index (kg/m ²)	23.2 ± 3.3	23.4 ± 3.8	23.1 ± 3.0
Creatinine (mg/dl)	0.73 ± 0.16	0.74 ± 0.17	0.73 ± 0.16
AST (IU/l)	62.0 ± 44.8	63.2 ± 48.6	61.4 ± 42.5
ALT (IU/l)	74.6 ± 56.1	75.4 ± 60.5	74.2 ± 53.5
GGT (IU/l)	58.6 ± 57.0	59.5 ± 58.5	58.0 ± 56.2
Albumin	4.0 ± 0.3	4.0 ± 0.3	4.0 ± 0.4
Total cholesterol	171.8 ± 31.7	171.5 ± 32.3	172.2 ± 30.8
HDL cholesterol	50.9 ± 14.5	51.1 ± 14.3	50.5 ± 15.0
LDL cholesterol	95.5 ± 27.7	96.1 ± 27.9	94.1 ± 27.2
Triglyceride	108.8 ± 55.4	107.8 ± 57.3	110.9 ± 51.7
Glucose	111.2 ± 39.0	111.7 ± 39.8	110.3 ± 37.6
Alpha-fetoprotein	14.5 ± 43.9	13.3 ± 37.7	16.8 ± 54.1
White blood cell count (/μl)	4946 ± 1427	4851 ± 1355	4999 ± 1464
Hemoglobin (g/dl)	14.2 ± 1.4	14.2 ± 1.4	14.2 ± 1.4
Platelets (10 ⁹ /mm ³)	166.1 ± 51.4	165.6 ± 51.7	166.4 ± 51.2
Ccr (ml/min)	95.1 ± 26.5	94.8 ± 25.9	95.4 ± 26.9
HCV RNA (KIU/ml)	1978 ± 1442	1937 ± 1382	2001 ± 1476
Fibrosis stage (F0–2/F3–4/ND)	695/148/238	454/85/152	241/63/86
Activity (A0–1/2–3/ND)	457/383/241	295/241/154	162/141/87
PEG-IFN alpha-2b dosage (μg/kg/body weight)	1.48 ± 0.13	1.49 ± 0.13	1.48 ± 0.13
Ribavirin dosage (600/800/1000 mg)	581/457/43	370/298/23	211/159/20
Decline of Hb at week 1	−0.2 ± 0.8	−0.2 ± 0.8	−0.2 ± 0.8
Decline of Hb at week 2	−1.2 ± 1.2	−1.2 ± 1.3	−1.2 ± 1.1
Decline of Hb at week 4	−2.3 ± 1.5	−2.2 ± 1.4	−2.4 ± 1.5
Decline of Hb at week 8	−2.8 ± 1.4	−2.8 ± 1.4	−2.7 ± 1.4

Data are expressed as median ± standard deviation unless otherwise indicated

AST aspartate aminotransferase, ALT alanine aminotransferase, GGT gamma-glutamyltransferase, Ccr creatinine clearance, Hb hemoglobin

were quantified by Cobas Amplicor (Roche Diagnostic Systems, CA, USA).

Database of variables and decision-tree analysis

A database of pretreatment variables was created containing 3 variables from hematological tests (Hb, white blood cells, and platelets), 11 variables from blood biochemical tests (creatinine, albumin, aspartate aminotransferase, alanine aminotransferase, gamma-glutamyl transpeptidase, total cholesterol, HDL cholesterol, LDL cholesterol, triglyceride, fasting blood glucose, and alpha-fetoprotein), creatinine clearance (Ccr), serum level of HCV RNA, liver histology (activity, fibrosis), 3 variables from patient characteristics (age, gender, and body mass index), 2 variables from therapeutic factors (PEG-IFN alpha-2b dosage, ribavirin dosage), and the level of decline of Hb concentration (at the end of 1, 2,

4, and 8 weeks from the start of treatment). Ccr levels were calculated using the Cockcroft–Gault formula [25]. Variables with data deficiency of greater than 15% were not included in the decision-tree analysis. Data deficiency was 21% in liver histology (activity, fibrosis) and 16% in the level of decline of Hb concentration at the end of 1 week. Accordingly, these variables were excluded from the database.

On the basis of this database, we implemented the recursive partitioning analysis algorithm referred to as the decision-tree analysis algorithm [26] to define subgroups of patients with respect to the possibility of severe anemia. The data-mining software used was IBM SPSS Modeler 13 (IBM SPSS Inc, Chicago, IL, USA), as reported previously [21–24]. In brief, the software searched the patient population for the most significant variables and cutoffs to be used for dividing the total population into 2 subgroups, having different probabilities of severe anemia. Thereafter,

the analysis was repeated on all subgroups in the same manner until either no additional significant variable was detected or the sample size was less than 20.

For other statistical analyses, including multivariable analysis, IBM SPSS Statistics software v.15.0 (IBM SPSS Inc, Chicago, IL, USA) was used. Differences in proportions were tested by the chi-squared test. Differences in continuous variables were compared by Student's *t* test. For univariate and multivariate analyses, logistic regression analysis was used to predict ribavirin-induced severe anemia. A value of $P < 0.05$ (two-tailed) was considered to indicate significance.

Results

Decision-tree analysis

Decision-tree analysis was carried out on the data of the model-building group using 27 variables, as described above. The analysis automatically selected 3 predictive variables to produce a total of 5 patient subgroups to build the decision tree (Fig. 1). Baseline Hb was selected as the first splitting variable, with an optimal cutoff of 14 g/dl. The possibility of severe anemia was 6.5% for patients with Hb levels <14 g/dl compared to 1.0% for patients with Hb

levels ≥ 14 g/dl. Among patients with Hb ≥ 14 g/dl, the level of decline of Hb at the end of 2 weeks from the start of treatment, with an optimal cutoff of 2 g/dl, was selected as the second splitting variable. Patients with lower decline levels had a lower probability of developing severe anemia [<2 g/dl (group A) 0.4% vs. ≥ 2 g/dl (group B) 2.5%]. Among patients whose Hb was less than 14 g/dl, Ccr was selected as the second splitting variable, with an optimal cutoff of 80 ml/min. Patients with higher Ccr levels had a lower probability of developing severe anemia [≥ 80 ml/min, 2.4% vs. <80 ml/min (group C) 11.8%]. Among patients with a Ccr ≥ 80 ml/min, the level of decline of Hb at the end of 2 weeks from the start of the treatment was selected as the third splitting variable, with an optimal cutoff of 2 g/dl. Patients with lower decline levels had a lower probability of developing severe anemia [<2 g/dl (group D) 1.4% vs. ≥ 2 g/dl (group E) 11.5%].

The probabilities of severe anemia for the 5 subgroups derived by this process were highly variable. The subgroup of patients with higher Hb levels (≥ 14 g/dl) (groups A and B) had a low probability of developing severe anemia (0.4–2.5%). Also, the subgroup of patients with lower Hb (<14 g/dl) but with a higher Ccr (≥ 80 ml/min) and lower Hb decline levels at the end of 2 weeks from the start of the treatment (<2 g/dl) (group D) showed a low probability of developing severe anemia (1.4%). On the other hand, the

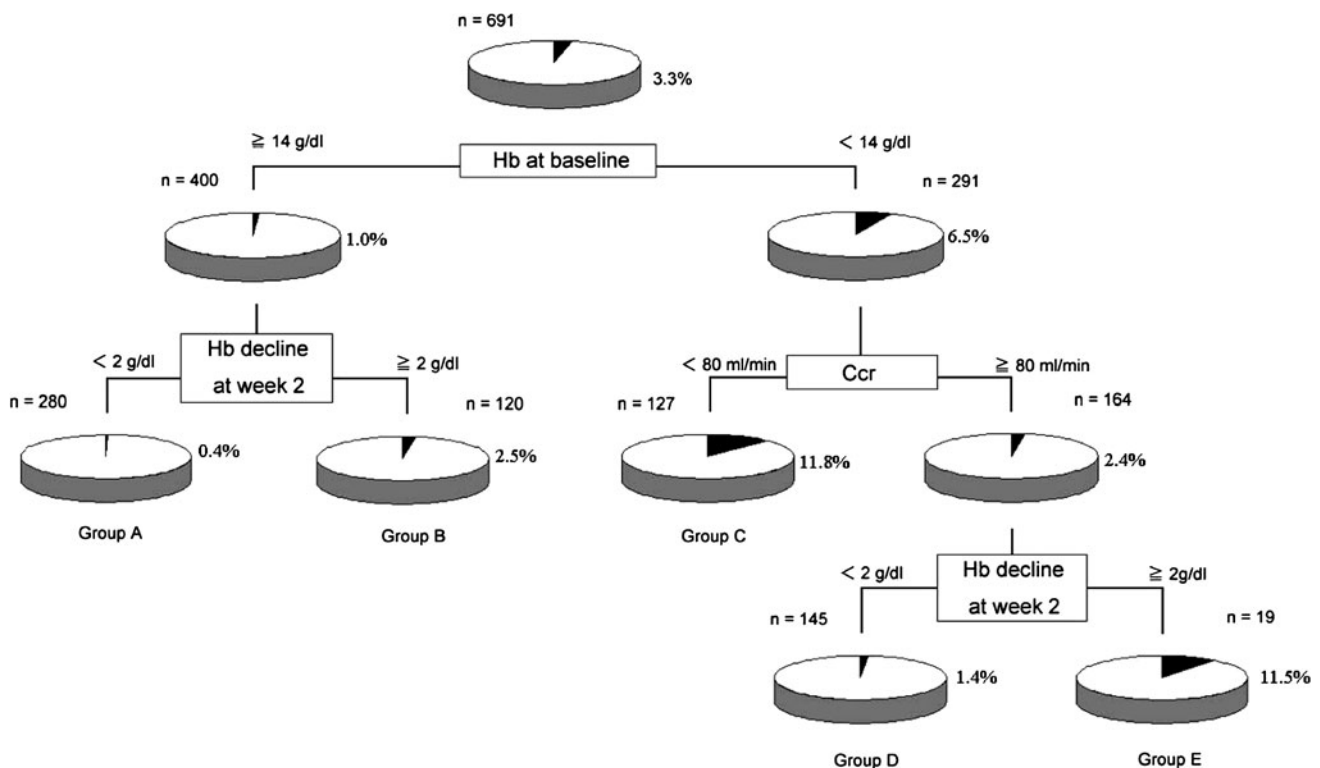


Fig. 1 Decision-tree analysis. Boxes indicate the splitting factors and the cutoff value for the split. Pie charts indicate the rate of severe anemia (Hb < 8.5 g/dl) for each group. Terminal groups classified by the analysis were labeled from A to E. Hb hemoglobin, Ccr creatine clearance

subgroup of patients with lower Hb (<14 g/dl) and lower Ccr (<80 ml/min) (group C) levels showed the highest probability of severe anemia (11.8%). Also, the subgroup of patients with lower Hb levels (<14 g/dl), higher Ccr (\geq 80 ml/min), and higher Hb decline levels at the end of 2 weeks from the start of treatment (\geq 2 g/dl) (group E) showed a high probability of developing severe anemia (11.5%).

Validation of the decision tree

The results of the decision tree were validated with the dataset of the internal validation group, which was independent of the model-building group dataset. Each patient in the validation group was allocated to groups A–E using the flow chart form of the decision tree. The rates of severe anemia (Hb < 8.5 g/dl) were 0.6% for group A, 3.0% for group B, 16.9% for group C, 2.3% for group D, and 11.0% for group E. The rates of severe anemia for each subgroup of patients were closely correlated between the model-building group and the internal validation group ($r^2 = 0.96$) (Fig. 2).

The efficiency and stability of the decision-tree model were validated using the discrimination efficiency curve (Fig. 3). The subgroups were sorted according to the order of incidence rate of severe anemia and validated using the correlation between the cumulative cases (%) and the cumulative incidence of severe anemia (%). The curve of the model-building group was located at the left upper part compared with the standard curve, indicating that the discrimination efficiency was high. Furthermore, the curve of the model-building group was extremely similar to the

curve of the internal validation group, indicating that the stability was high.

Factors associated with severe anemia determined by multivariate logistic regression analysis

We also explored the factors associated with severe anemia using standard statistical analysis. By univariable analysis, age, creatinine, Hb, Ccr, fibrosis stage, and decline of Hb at 2, 4, and 8 weeks from the start of treatment were found to be associated with severe anemia (Table 2) and the odds ratio for these factors were 1.06, 9.61, 0.47, 0.95, 3.14, 0.76, 0.70, and 0.68, respectively, by univariable logistic regression analysis (Table 3). By multivariate analysis, Hb, Ccr, and decline of Hb at 2 weeks from the start of treatment were found to be independently associated with severe anemia (Table 3). Fibrosis was not included in the multivariable analysis because data were not available for 238 patients. Creatinine was not included in the multivariable analysis because creatinine and Ccr were confounding factors due to their close correlation. Decline of Hb at week 2, 4, and 8 were also closely correlated. We selected decline of Hb at week 2 in the multivariable analysis because we think that variables at earlier time points may be more useful in clinical use. As a result, decision-tree and multivariable logistic regression analyses identified the same factors for prediction of severe anemia.

Discussion

Hemolytic anemia, a major common side effect of ribavirin treatment, is one of the most important adverse effects of PEG-IFN and ribavirin combination treatment. Therefore, before- and during-treatment prediction of the likelihood of severe anemia can be very useful for physicians to support clinical decisions concerning the dose reduction of ribavirin. Reducing the dose of ribavirin has been shown to affect the HCV RNA negativity [27], and the discontinuation of ribavirin has been reported to lead to a marked decrease of SVR [9]. Therefore, averting ribavirin discontinuance, even if its dose must be reduced, can lead to an improvement in the SVR rate. It is important to identify patients prone to develop severe anemia leading to ribavirin dose reduction or discontinuance in the early phase of treatment.

Using decision-tree analysis, we constructed a simple model for predicting the incidence of severe anemia during therapy. The analysis highlighted 3 variables relevant to virological response: Hb, Ccr, and the decline of Hb concentration by 2 g/dl at the end of the 2 weeks from the start of treatment. Classification based on these variables identified subgroups of genotype 1b chronic hepatitis C patients with high probabilities of developing severe anemia. The

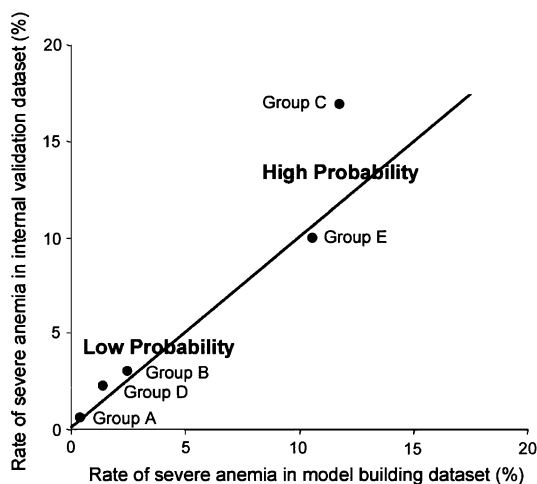


Fig. 2 Validation of decision-tree analysis with the internal validation dataset: subgroup-stratified comparison of the rate of severe anemia. The rate of severe anemia in each subgroup was plotted. The X-axis represents the model-building dataset and the Y-axis represents the internal validation dataset. There was a close correlation between the model-building and internal validation datasets ($r^2 = 0.96$)

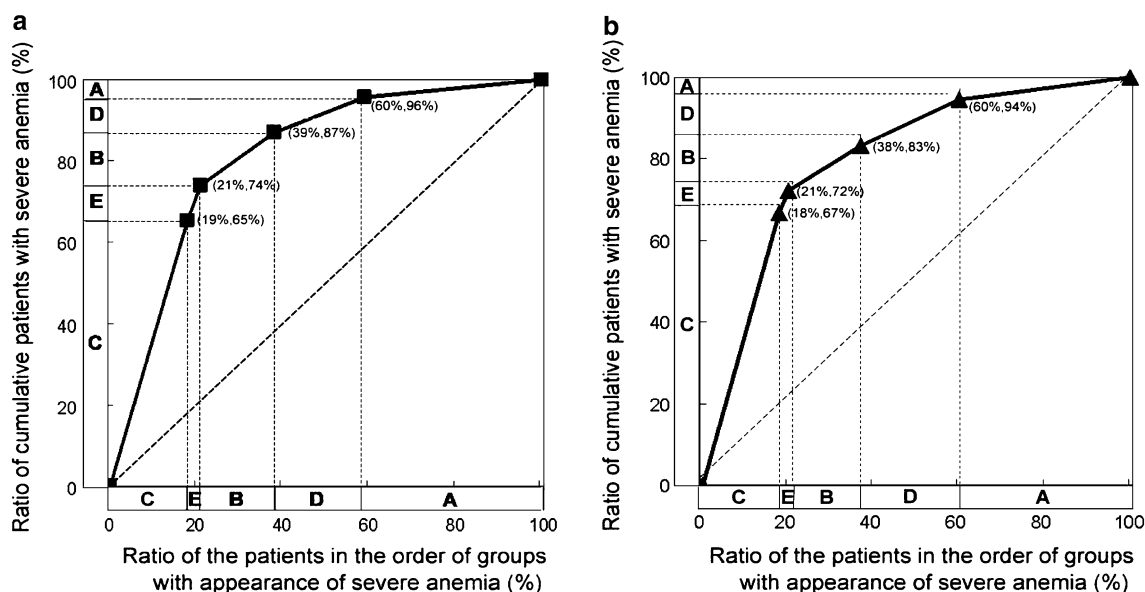


Fig. 3 Validation of the efficiency and stability by the discrimination efficiency curve. **a** Model-building group and **b** internal validation group. The groups were sorted in the order of incidence rate of severe anemia and validated using the correlation between cumulative cases (%) and the cumulative incidence of severe anemia

(%). The X-axis represents the ratio of patients in the order of groups predicting the development of anemia and the Y-axis represents the cumulative patients suffering from severe anemia. The discrimination efficiency and stability of the curve of the model-building group were high

reproducibility of the model was confirmed with the internal validation dataset. An advantage of decision-tree analysis over traditional regression models is that the decision-tree model is user-intuitive and can be readily interpreted by medical professionals without the need for any specific knowledge of statistics. Patients can be allocated to specific subgroups based on a defined rate of severe anemia simply by following the flow chart format. Using this model, an estimate of the incidence of severe anemia can be obtained rapidly, which may facilitate clinical decision making for the reduction of ribavirin dosage. Thus, this model could be readily applicable for clinical practice.

According to the results of the decision tree, patients were categorized into 2 groups. The rates of severe anemia were 0.4–2.5% for the low probability group and 11.5–11.8% for the high probability group. For example, patients in the high probability group may be the most suitable candidates for dose reduction of ribavirin. Decision-tree analysis revealed that the high probability groups are patient groups with lower Hb (<14 g/dl) and lower Ccr (<80 ml/min) levels (group C) and patient groups with lower Hb (<14 g/dl), higher Ccr (\geq 80 ml/min), and higher Hb decline levels at 2 weeks from the start of treatment (\geq 2 g/dl) (group E). In particular, groups C and A were shown to be clinically significant in Fig. 3; group C includes the majority of patients suffering from severe anemia (65% in the model-building group and 67% in the internal validation group) and the very steep tilt angle of

the group C slope means that group C patients have a very high probability of developing severe anemia. On the other hand, group A includes a large number of patients (40% in the model-building group and 40% in the internal validation group), and the very gentle tilt angle of the group A slope implies that group A patients have a very low probability of developing severe anemia.

Predicting the progression of anemia is necessary to decide whether medication can be continued while minimizing the disadvantages of anemia. The apparent clearance of ribavirin (CL/F), which reflects its plasma concentration at 4 weeks after the start of combination therapy, has been used as a predictive factor for developing ribavirin-induced hemolytic anemia before the start of treatment [9, 10]. However, the use of CL/F is not practical for general clinicians, because the calculation of CL/F is complicated. We revealed that a decline of Hb concentration by 2 g/dl at 2 weeks from the start of treatment (“2 by 2” standard) is both sensitive and convenient for identifying patients at high risk for severe anemia [10, 11]. The present study using decision-tree analysis revealed that Hb decline at week 2 was a significant and independent predictor of severe anemia. When considered along with other predictive factors, decision-tree analysis enables more exact identification of the patients prone to severe anemia.

Recently, a genome-wide association technique was used to show that ITPA polymorphism affects ribavirin-induced anemia. Polymorphisms (rs 1127354 and rs 7270101) that cause ITPase deficiency are strongly

Table 2 Comparison of clinical parameters of patients with and without severe anemia

	Anemia (N = 41)	No anemia (N = 1040)	P value
Age (years)	61.0 ± 7.6	55.4 ± 10.6	0.001
Gender (male/female)	18/23	594/446	0.109
Body mass index (kg/m ²)	22.4 ± 2.9	23.2 ± 3.3	0.119
Creatinine (mg/dl)	0.79 ± 0.24	0.7 ± 0.16	0.011
AST (IU/L)	74.2 ± 62.9	61.6 ± 43.9	0.075
ALT (IU/L)	79.6 ± 68.7	74.5 ± 55.6	0.565
GGT (IU/L)	40.7 ± 31.0	59.2 ± 57.6	0.071
Albumin	3.9 ± 0.3	4.0 ± 0.3	0.260
Total cholesterol	177.1 ± 23.1	171.6 ± 32.0	0.258
HDL cholesterol	50.8 ± 8.0	50.9 ± 14.7	0.986
LDL cholesterol	93.6 ± 22.0	95.5 ± 27.9	0.717
Triglyceride	109.1 ± 45.0	108.8 ± 55.8	0.974
Glucose	114.1 ± 32.9	111.1 ± 39.2	0.738
Alpha-fetoprotein	29.0 ± 71.4	13.9 ± 42.5	0.229
White blood cell count (/μl)	4632 ± 1828	4958 ± 1408	0.152
Hemoglobin (g/dl)	12.9 ± 1.3	14.2 ± 1.4	0.0001
Platelets (10 ⁹ /mm ³)	152.1 ± 51.7	166.6 ± 51.3	0.075
Ccr (ml/min)	75.4 ± 23.6	95.9 ± 26.4	0.0001
HCV RNA (KIU/ml)	1807 ± 1456	1985 ± 1442	0.438
Fibrosis stage (F0–2/F3–4/ND)	18/10/13	677/138/225	0.019
Activity (A0–1/2–3/ND)	14/14/13	443/369/228	0.701
PEG-IFN alpha-2b dosage	1.50 ± 0.13	1.48 ± 0.13	0.260
Ribavirin dosage (600/800/1000 mg)	28/13/0	553/444/43	0.105
Decline of hemoglobin at week 1	−0.2 ± 1.1	−0.2 ± 0.8	0.644
Decline of hemoglobin at week 2	−1.6 ± 1.9	−1.2 ± 1.2	0.022
Decline of hemoglobin at week 4	−3.0 ± 1.7	−2.2 ± 1.4	0.005
Decline of hemoglobin at week 8	−3.6 ± 1.6	−2.7 ± 1.4	0.003

Data are expressed as median ± standard deviation unless otherwise indicated
 AST aspartate aminotransferase, ALT alanine aminotransferase, GGT gamma-glutamyltransferase, Ccr creatinine clearance, Hb hemoglobin

Table 3 Univariable and multivariable logistic regression analysis of factors associated with severe anemia

	Univariable analysis			Multivariable analysis		
	Odds	95% CI	P value	Odds	95% CI	P value
Age (years)	1.06	1.03–1.11	<0.0001	1.02	0.96–1.08	0.984
Creatinine (mg/dl)	9.61	1.91–48.4	0.006	–	–	–
Hb (g/dl)	0.47	0.37–0.59	<0.0001	0.40	0.29–0.55	<0.0001
Ccr (ml/min)	0.95	0.94–0.97	<0.0001	0.97	0.95–0.99	0.012
Fibrosis (F3–4)	3.14	1.49–6.60	0.003	–	–	–
Decline of Hb at week 2	0.76	0.61–0.95	0.017	0.54	0.39–0.74	0.0001
Decline of Hb at week 4	0.70	0.57–0.87	0.001	–	–	–
Decline of Hb at week 8	0.68	0.55–0.85	0.001	–	–	–

Hb hemoglobin, Ccr creatinine clearance

associated with protection from ribavirin-induced hemolytic anemia and with a lesser need for ribavirin dose reduction [28–30]. These polymorphisms are very valuable, but the indication for treatment is determined not by them but by viral genotypes or by *IL28B* variations. The present decision tree, which involves a factor attained after initiation of PEG-IFN plus ribavirin therapy, i.e., Hb

decline at week 2, is useful for selecting the best regimen, and can be easily used by general clinicians.

What is unique to the present study is the visualization of the probability of severe anemia by combining factors and its high reproducibility, as revealed by high-quality validation of the internal validation dataset that was completely independent of the model-building dataset. The

factors used in the decision-tree model were clinical parameters that are readily available through the usual work-up of patients. This model can be immediately applied to clinical practice without imposing any cost for additional examinations.

A potential limitation of the present study is that data-mining analysis has an intrinsic risk of showing relationships that are relevant to the original dataset but are not reproducible across different populations. Although internal validation showed that our model had high reproducibility, we recognize that further validation using a larger external validation cohort, especially in populations other than Japanese, is necessary to verify the reliability of our model.

In conclusion, we built the decision-tree model for predicting severe anemia caused by PEG-IFN alpha-2b plus ribavirin combination therapy in chronic hepatitis C with genotype 1b and high viral load. Because this decision-tree model was composed of simple variables, it can be easily applied to clinical practice. This model may have the potential to support decisions concerning ribavirin dose reduction during PEG-IFN alpha-2b plus ribavirin combination therapy and contribute to increasing the rate of SVR.

Acknowledgments This study was supported by a grant-in-aid from the Ministry of Health, Labour and Welfare of Japan (H20-kannen-006).

Conflict of interest All authors have no financial relationship relevant to this study to disclose.

References

- Strader DB, Wright T, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C. *Hepatology*. 2004;39(4):1147–71.
- Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncalves FL Jr, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med*. 2002;347(13):975–82.
- Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, Reindollar R, et al. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet*. 2001;358(9286):958–65.
- Nomura H, Tanimoto H, Kajiwara E, Shimono J, Maruyama T, Yamashita N, et al. Factors contributing to ribavirin-induced anemia. *J Gastroenterol Hepatol*. 2004;19:1312–7.
- Takaki S, Tsubota A, Hosaka T, Akuta N, Someya T, Kobayashi M, et al. Factors contributing to ribavirin dose reduction due to anemia during interferon alfa2b and ribavirin combination therapy for chronic hepatitis C. *J Gastroenterol*. 2004;39:668–73.
- Van Vlierbergh H, Delanghe JR, De Vos M, Leroux-Roel G, BASL Steering Committee. Factors influencing ribavirin-induced hemolysis. *J Hepatol*. 2001;34:911–6.
- Jen JF, Glue P, Gupta S, Zambas D, Hajian G. Population pharmacokinetic and pharmacodynamic analysis of ribavirin in patients with chronic hepatitis C. *Ther Drug Monit*. 2000;22:555–65.
- Lindhal K, Schvarcz R, Bruchfeld A, Ståhle L. Evidence that plasma concentration rather than dose per kilogram body weight predicts ribavirin-induced anemia. *J Viral Hepat*. 2004;11:84–7.
- Kamar N, Chatelut E, Manolis E, Lafont T, Izopet J, Rostaing L. Ribavirin pharmacokinetics in renal and liver transplant patients: evidence that it depends on renal function. *Am J Kidney Dis*. 2004;43:140–6.
- Hiramatsu N, Kurashige N, Oze T, Takehara T, Tamura S, Kasahara A, et al. Early decline of hemoglobin can predict progression of hemolytic anemia during pegylated interferon and ribavirin combination therapy in patients with chronic hepatitis C. *Hepatol Res*. 2008;38:52–9.
- Oze T, Hiramatsu N, Kurashige N, Tsuda N, Yakushijin T, Kanto T, et al. Early decline of hemoglobin correlates with progression of ribavirin-induced hemolytic anemia during interferon plus ribavirin combination therapy in patients with chronic hepatitis C. *J Gastroenterol*. 2006;41:862–87.
- Breiman L, Friedman RA, Olshen CJ, Stone CM. *Classification and regression trees*. Monterey: Wadsworth; 1980.
- Garzotto M, Beer TM, Hudson RG, Peters L, Hsieh YC, Barrera E, et al. Improved detection of prostate cancer using classification and regression tree analysis. *J Clin Oncol*. 2005;23:4322–9.
- Miyaki K, Takei I, Watanabe K, Nakashima H, Omae K. Novel statistical classification model of type 2 diabetes mellitus patients for tailor-made prevention using data mining algorithm. *J Epidemiol*. 2002;12:243–8.
- Averbook BJ, Fu P, Rao JS, Mansour EG. A long-term analysis of 1018 patients with melanoma by classic Cox regression and tree-structured survival analysis at a major referral center. *Surgery*. 2002;132:589–602.
- Leiter U, Buettner PG, Eigentler TK, Garbe C. Prognostic factors of thin cutaneous melanoma: an analysis of the central malignant melanoma registry of the German Dermatological Society. *J Clin Oncol*. 2004;22:3660–7.
- Valera VA, Walter BA, Yokoyama N, Koyama Y, Iiai T, Okamoto H, et al. Prognostic groups in colorectal carcinoma patients based on tumor cell proliferation and classification and regression tree (CART) survival analysis. *Ann Surg Oncol*. 2007;14:34–40.
- Zlobec I, Steele R, Nigam N, Compton CC. A predictive model of rectal tumor response to preoperative radiotherapy using classification and regression tree methods. *Clin Cancer Res*. 2005;11:5440–3.
- Baquerizo A, Anselmo D, Shackleton C, Chen TW, Cao C, Weaver M, et al. Phosphorus as an early predictive factor in patients with acute liver failure. *Transplantation*. 2003;75:2007–14.
- LeBlanc M, Crowley J. A review of tree-based prognostic models. *Cancer Treat Res*. 1995;75:113–24.
- Kurosaki M, Matsunaga K, Hirayama I, Tanaka T, Sato M, Yasui Y, et al. A predictive model of response to peginterferon ribavirin in chronic hepatitis C using classification and regression tree analysis. *Hepatol Res*. 2010;40:251–60.
- Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, Hiramatsu N, et al. Pretreatment prediction of response to peginterferon plus ribavirin therapy in genotype 1 chronic hepatitis C using data mining analysis. *J Gastroenterol*. 2011;46:401–9.
- Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, Hiramatsu N, et al. Sequences in the interferon sensitivity-determining region and core region of hepatitis C virus impact pretreatment prediction of response to peg-interferon plus ribavirin: data mining analysis. *J Med Virol*. 2011;83:445–52.
- Kurosaki M, Tanaka Y, Nishida N, Sakamoto N, Enomoto N, Honda M, et al. Pre-treatment prediction of response to

- pegylated-interferon plus ribavirin for chronic hepatitis C using genetic polymorphism in *IL28B* and viral factors. *J Hepatol*. 2011;54:439–48.
25. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron*. 1976;16:31–41.
 26. Segal MR, Bloch DA. A comparison of estimated proportional hazards models and regression trees. *Stat Med*. 1989;8:539–50.
 27. Hiramatsu N, Oze T, Yakushijin T, Inoue Y, Igura T, Mochizuki K, et al. Ribavirin dose reduction raises relapse rate dose-dependently in genotype 1 patients with hepatitis C responding to pegylated interferon alfa-2b plus ribavirin. *J Viral Hepat*. 2009; 16:586–94.
 28. Fellay J, Thompson AJ, Ge D, Gumbs CE, Urban TJ, Shianna KV, et al. ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C. *Nature*. 2010;464:405–8.
 29. Thompson AJ, Fellay J, Patel K, Tillmann HL, Naggie S, Ge D, et al. Variants in the ITPA gene protect against ribavirin-induced hemolytic anemia and decrease the need for ribavirin dose reduction. *Gastroenterology*. 2010;139:1181–9.
 30. Ochi H, Maekawa T, Abe H, Hayashida Y, Nakano R, Kubo M, et al. ITPA polymorphism affects ribavirin-induced anemia and outcomes of therapy-A genome-wide study of Japanese HCV virus patients. *Gastroenterology*. 2010;139:1190–7.