

to change the cell cycle speed without altering the uninfected cell count (Fig. 6d). We used this technique to investigate how changing the cell cycle speed affected the equilibrium phase. Fig. 6e shows the results. Cell lifespan increases while the cell cycle speed declines. The equilibrium virus count increased in accordance with slower cell cycle speeds.

#### 4. Discussion

In this study, we investigated the models using two agent-based simulation methods to program a simple virus–host chronic infection model. The same model written in two different programming language systems displayed the same results. The transient phase was unlike that seen in a mathematical simulation with no overshoot in virus count, but rather a smooth transition to the equilibrium phase. The virus count at the start of the simulation only had effect on the rate of infection development. Increases in virus lifespan, uninfected cell lifespan, uninfected cell regeneration rate, virus production count from infected cells, and infection rate all led to increased equilibrium-phase virus count. Rises in the infected cell lifespan–shortening ratio, latent period, and cell cycle speed decreased the equilibrium-phase virus count. The size of the space itself had no innate effect on the equilibrium phase, but a speed of movement of the virus that was twice the size of the space produced the maximum virus count.

Reproducibility is the basis for all scientific study, but there are many problems to prove it in computer simulations, such as programming bugs. As agent-based simulation deals with numerous agents individually, it requires vast amounts of calculations. Accumulation of very small change of values leads to large differences of results. In this study, we investigated two programs based on two programming languages to confirm the reproducibility of our simulation results in different programming languages. The results of two simulations were consistent, but in StarLogo, the lifespan parameters had a tendency to be lower than when they were set while simulations were actually in progress. This may be because the number of digits used in calculations was different between the two programs. RePast performs calculations to at least eight decimal places. In StarLogo, the library settings only enable settings to be made up to five decimal places. It is probable that these small differences accumulate during repeated calculations and are reflected in the simulation. Ultimately, we confirmed that the differences in results obtained by using different libraries and programming languages were not innate and by making the parameters consistent during simulation, consistent results were obtained.

Mathematical models using formulae for HIV therapy was published in 1994, the method has since been applied to HBV and HCV (Ho et al., 1995; Nowak et al., 1996; Neumann et al., 1998), and they were thought to be good reflections of the reality. In the mathematical model, viruses and cells are conceived as individuals in the concept itself, but both of them are perceived *en masse* when calculations are performed. However a feature of the agent-based simulation is that it deals with individual viruses and cells as separate agents. By moving each agent individually, it probes the factors influencing overall shifts from the micro viewpoint. When the space is viewed as a whole, it is possible to watch on the screen the collective movement of groups of agents. Recently, models that provide a visual representation of Epstein-Barr virus and HIV infection have been reported, both of which are useful for an instinctive and intuitive understanding (Duca et al., 2007; Shapiro et al., 2008; Castiglione et al., 2007).

In agent-based simulation model, virus count transit smoothly to the equilibrium phase. On the other hand, virus counts overshoot during transient phase in mathematical model. We think this difference is derived from technicality of different model-

ing. The difference in concepts between mathematical models and agent-based models is the space. The mathematical model has no space in concept, but agents move across the space in the agent-based model. In agent-based models, the densities of virus and cells change overtime especially in the transition phase because of the limited space. These changes of the densities of virus and cells lead to the dynamic change of the encounter rate of viruses and cells. The mathematical model does not make such concept of the density; the encounter rate is constant. This may be the reason for the difference between two models in the transition phase. Since no overshoot of virus counts in transient phase had been reported from in vivo studies of hepatitis C virus and simian immunodeficiency virus (Dahari et al., 2005; Nowak et al., 1997), agent-based model correlates with actual biology in vivo at least for these viruses. The increase of initial virus count at the start of simulation correlates with higher encounter rate of viruses and cells which make the linear increasing of infection forming rate. Mathematical model can only express the infection formation rate as “infected or not”.

The importance of viral passing speed in the agent-based model is also explained by the “space”. Although the virus actually moves through the blood stream in our body and virus could not decide their moving speeds by themselves, there is most appropriate speed for virus to meet the cells on the simulation space by the highest probability. The effect of cell cycle speed should be mentioned by another affection of the space. A fast cell cycle speed means that the lifespan of uninfected cells is short. Then fast cell cycle speed leads to the short lifespan of infected cells. A higher regeneration rate for uninfected cells results in a higher rate of infection among uninfected cells by viruses, but in situations where viruses and cells are dispersed around the space this is ineffective in increasing the infection rate, as the latter depends on the probability that they will encounter one another. As a result, the infected cell count decreases during the equilibrium phase, as does the virus count.

In this study, we confirmed the reproducibility and usability of agent-based models in expressing the interaction between viruses and cells. A feature of this simulation system is that it uses the concept of space as actual space, which means that the existence of the space becomes an additional controlling factor on the simulation results. This is a concept that is absent from mathematical models. The reality is that we have a spatial existence, and an advantage of the agent-based simulation system is the fact that it accounts for the space. Another feature of the simulation system is that it enables the condition to be perceived in visual terms, making it easy to understand. However it may be affected by computer performance and by the limitations of programming languages or the program itself, this system may offer a powerful tool for the future analysis of real virus–host interaction disease.

#### Conflict of interest

No conflicts of interest exist for all authors.

#### References

- Castiglione, F., Pappalardo, F., Bernaschi, M., Motta, S., 2007. Optimization of HAART with genetic algorithms and agent-based models of HIV infection. *Bioinformatics* 23, 3350–3355. doi:10.1093/bioinformatics/btm408.
- Dahari, H., Major, M., Zhang, X., Mihalik, K., Rice, C.M., Perelson, A.S., Feinstone, S.M., Neumann, A.U., 2005. Mathematical modeling of primary hepatitis c infection: noncytolytic clearance and early blockage of virion production. *Gastroenterology* 128, 1056–1066. doi:10.1053/j.gastro.2005.01.049.
- Duca, K.A., Shapiro, M., Delgado-Eckert, E., Hadinoto, V., Jarrah, A.S., Laubenbacher, R., Lee, K., Luzuriaga, K., Polys, N.F., Thorley-Lawson, D.A., 2007. A virtual look at Epstein-Barr virus infection: biological interpretations. *PLoS Pathog.* 3, 1388–1400. doi:10.1371/journal.ppat.0030137.
- Gilbert, N., Banks, S., 2002. Platforms and methods for agent-based modelling. *Proc. Natl. Acad. Sci. U.S.A.* 99 (Suppl. 3), 7197–7198.

- Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., Markowitz, M., 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373, 123–126, doi:10.1038/373123a0.
- Naniche, D., 2009. Human immunology of measles virus infection. *Curr. Top. Microbiol. Immunol.* 330, 151–171.
- Neumann, A.U., Lam, N.P., Dahari, H., Gretch, D.R., Wiley, T.E., Layden, T.J., Perelson, A.S., 1998. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science* 282, 103–107, doi:10.1126/science.282.5386.103.
- Nowak, M.A., Bonhoeffer, S., Hill, A.M., Boehme, R., Thomas, H.C., McDade, H., 1996. Viral dynamics in hepatitis B virus infection. *Proc. Natl. Acad. Sci. U.S.A.* 93, 4398–4402.
- Nowak, M.A., Lloyd, A.L., Vasquez, G.M., Wiltrout, T.A., Wahl, L.M., Biscoberger, N., Williams, J., Kinter, A., Fauci, A.S., Hirsch, V.M., Lifson, J.D., 1997. Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection. *J. Virol.* 71, 7518–7525.
- Shapiro, M., Duca, K.A., Lee, K., Delgado-Eckert, E., Hawkins, J., Jarrar, A.S., Laubacher, R., Polys, N.F., Hadinoto, V., Thorley-Lawson, D.A., 2008. A virtual look at Epstein-Barr virus infection: simulation mechanism. *J. Theor. Biol.* 252, 633–648, doi:10.1016/j.jtbi.2008.01.032.
- See, H., Wark, P., 2008. Innate immune response to viral infection of the lungs. *Paediatr. Respir. Rev.* 9, 243–250, doi:10.1016/j.prrv.2008.04.001.

## Pre-treatment prediction of response to pegylated-interferon plus ribavirin for chronic hepatitis C using genetic polymorphism in *IL28B* and viral factors

Masayuki Kurosaki<sup>1</sup>, Yasuhiro Tanaka<sup>2</sup>, Nao Nishida<sup>3</sup>, Naoya Sakamoto<sup>4</sup>, Nobuyuki Enomoto<sup>5</sup>, Masao Honda<sup>6</sup>, Masaya Sugiyama<sup>2</sup>, Kentaro Matsuura<sup>2</sup>, Fuminaka Sugauchi<sup>2</sup>, Yasuhiro Asahina<sup>1</sup>, Mina Nakagawa<sup>4</sup>, Mamoru Watanabe<sup>4</sup>, Minoru Sakamoto<sup>5</sup>, Shinya Maekawa<sup>5</sup>, Akito Sakai<sup>6</sup>, Shuichi Kaneko<sup>6</sup>, Kiyooki Ito<sup>7</sup>, Naohiko Masaki<sup>7</sup>, Katsushi Tokunaga<sup>3</sup>, Namiki Izumi<sup>1,\*</sup>, Masashi Mizokami<sup>2,7</sup>

<sup>1</sup>Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan; <sup>2</sup>Department of Virology, Liver Unit, Nagoya City University, Graduate School of Medical Sciences, Nagoya, Japan; <sup>3</sup>Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; <sup>4</sup>Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Tokyo, Japan; <sup>5</sup>First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan; <sup>6</sup>Department of Gastroenterology, Kanazawa University, Graduate School of Medicine, Kanazawa, Japan; <sup>7</sup>Research Center for Hepatitis and Immunology, International Medical Center of Japan, Konodai Hospital, Ichikawa, Japan

**Background & Aims:** Pegylated interferon and ribavirin (PEG-IFN/RBV) therapy for chronic hepatitis C virus (HCV) genotype 1 infection is effective in 50% of patients. Recent studies revealed an association between the *IL28B* genotype and treatment response. We aimed to develop a model for the pre-treatment prediction of response using host and viral factors.

**Methods:** Data were collected from 496 patients with HCV genotype 1 treated with PEG-IFN/RBV at five hospitals and universities in Japan. *IL28B* genotype and mutations in the core and IFN sensitivity determining region (ISDR) of HCV were analyzed to predict response to therapy. The decision model was generated by data mining analysis.

**Results:** The *IL28B* polymorphism correlated with early virological response and predicted null virological response (NVR) (odds ratio = 20.83,  $p < 0.0001$ ) and sustained virological response (SVR) (odds ratio = 7.41,  $p < 0.0001$ ) independent of other covariates. Mutations in the ISDR predicted relapse and SVR independent of *IL28B*. The decision model revealed that patients with the minor *IL28B* allele and low platelet counts had the highest NVR (84%) and lowest SVR (7%), whereas those with the major *IL28B* allele and mutations in the ISDR or high platelet counts had the lowest NVR (0–17%) and highest SVR (61–90%). The model had high reproducibility and predicted SVR with 78% specificity and 70% sensitivity.

**Conclusions:** The *IL28B* polymorphism and mutations in the ISDR of HCV were significant pre-treatment predictors of response to PEG-IFN/RBV. The decision model, including these host and viral factors may support selection of optimum treatment strategy for individual patients.

© 2010 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved.

### Introduction

Hepatitis C virus (HCV) infection is the leading cause of cirrhosis and hepatocellular carcinoma worldwide [1]. The successful eradication of HCV, defined as a sustained virological response (SVR), is associated with a reduced risk of developing hepatocellular carcinoma. Currently, pegylated interferon (PEG-IFN) plus ribavirin (RBV) is the most effective standard of care for chronic hepatitis C but the rate of SVR is around 50% in patients with HCV genotype 1 [2,3], the most common genotype in Japan, Europe, the United States, and many other countries. Moreover, 20–30% of patients with HCV genotype 1 have a null virological response (NVR) to PEG-IFN/RBV therapy [4]. The most reliable method for predicting the response is to monitor the early decline of serum HCV-RNA levels during treatment [5] but there is no established method for prediction before treatment. Because PEG-IFN/RBV therapy is costly and often accompanied by adverse effects such as flu-like symptoms, depression and hematological abnormalities, pre-treatment predictions of those patients who are unlikely to benefit from this regimen enables ineffective treatment to be avoided.

Recently, it has been reported through a genome-wide association study (GWAS) of patients with genotype 1 HCV that single nucleotide polymorphisms (SNPs) located near the *IL28B* gene are strongly associated with a response to PEG-IFN/RBV therapy in

**Keywords:** *IL28B*; ISDR; Peg-interferon; Ribavirin; Data mining; Decision tree.  
Received 14 March 2010; received in revised form 22 June 2010; accepted 7 July 2010; available online 19 September 2010

\* Corresponding author. Address: Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, 1-26-1 Kyonan-cho, Musashino-shi, Tokyo 180-8610, Japan. Tel.: +81 422 32 3111; fax: +81 422 32 9551.  
E-mail address: nizumi@musashino.jrc.or.jp (N. Izumi).



## Research Article

Table 1. Baseline characteristics of all patients, and patients assigned to the model building or validation groups.

	All patients n = 496	Model group n = 331	Validation group n = 165
Gender: male	250 (50%)	170 (51%)	80 (48%)
Age (years)	57.1 ± 9.9	56.8 ± 9.7	57.5 ± 10.2
ALT (IU/L)	78.6 ± 60.8	78.1 ± 61.4	79.7 ± 59.6
GGT (IU/L)	59.3 ± 63.6	58.9 ± 62.0	60.2 ± 66.9
Platelets (10 <sup>9</sup> /L)	154 ± 53	153 ± 52	154 ± 56
Fibrosis: F3-4	121 (24%)	80 (24%)	41 (25%)
HCV-RNA: >600,000 IU/ml	409 (82%)	273 (82%)	136 (82%)
ISDR mutation: ≤1	220 (88%)	290 (88%)	145 (88%)
Core 70 (Arg/Gln or His)	293 (59%)/203 (41%)	197 (60%)/134 (40%)	96 (58%)/69 (42%)
Core 91 (Leu/Met)	299 (60%)/197 (40%)	200 (60%)/131 (40%)	99 (60%)/66 (40%)
<i>IL28B</i> : Minor allele	151 (30%)	101 (31%)	50 (30%)
SVR	194 (39%)	129 (39%)	65 (39%)
Relapse	152 (31%)	103 (31%)	49 (30%)
NVR	150 (30%)	99 (30%)	51 (31%)

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase; ISDR, interferon sensitivity determining region; Arg, arginine; Gln, glutamine; His, histidine; Leu, leucine; Met, methionine; Minor, heterozygote or homozygote of minor allele; SVR, sustained virological response; NVR, null virological response.

Japanese [6], European [7], and a multi-ethnic population [8,9]. The last three studies focused on the association of SNPs in the *IL28B* region with SVR [7–9] but we found a stronger association with NVR [6]. In addition to these host genetic factors, we have reported that mutations within a stretch of 40 amino acids in the NS5A region of HCV, designated as the IFN sensitivity determining region (ISDR), are closely associated with the virological response to IFN therapy: a lower number of mutations is associated with treatment failure [10–13]. Amino acid substitutions at positions 70 and 91 of the HCV core region (Core70, Core91) also have been reported to be associated with response to PEG-IFN/RBV therapy: glutamine (Gln) or histidine (His) at Core70 and methionine (Met) at Core91 are associated with treatment resistance [4,14]. The importance of substitutions in the HCV core and ISDR was confirmed recently by a Japanese multicenter study [15]. How these viral factors contribute to response to therapy is yet to be determined. For general application in clinical practice, host genetic factors and viral factors should be considered together.

Data mining analysis is a family of non-parametric regression methods for predictive modeling. Software is used to automatically explore the data to search for optimal split variables and to build a decision tree structure [16]. The major advantage of decision tree analysis over logistic regression analysis is that the results of the analysis are presented in the form of flow chart, which can be interpreted intuitively and readily made available for use in clinical practice [17]. The decision tree analysis has been utilized to define prognostic factors in various diseases [18–25]. We have reported recently its usefulness for the prediction of an early virological response (undetectable HCV-RNA within 12 weeks of therapy) to PEG-IFN/RBV therapy in chronic hepatitis C [26].

This study aimed to define the pre-treatment prediction of response to PEG-IFN/RBV therapy through the integrated analysis of host factors, such as the *IL28B* genetic polymorphism and various clinical covariates, as well as viral factors, such as mutations in the HCV core and ISDR and serum HCV-RNA load. In addition,

for the general application of these results in clinical practice, decision models for the pre-treatment prediction of response were determined by data mining analysis.

### Materials and methods

#### Patients

This was a multicentre retrospective study supported by the Japanese Ministry of Health, Labor and Welfare. Data were collected from a total of 496 chronic hepatitis C patients who were treated with PEG-IFN alpha and RBV at five hospitals and universities throughout Japan. Of these, 98 patients also were included in the original GWAS analysis [6]. The inclusion criteria in this study were as follows (1) infection by genotype 1b, (2) lack of co-infection with hepatitis B virus or human immunodeficiency virus, (3) lack of other causes of liver disease, such as autoimmune hepatitis, and primary biliary cirrhosis, (4) completion of at least 24 weeks of therapy, (5) adherence of more than 80% to the planned dose of PEG-IFN and RBV for the NVR patients, (6) availability of DNA for the analysis of the genetic polymorphism of *IL28B*, and (7) availability of serum for the determination of mutations in the ISDR and substitutions of Core70 and Core91 of HCV. Patients received PEG-IFN alpha-2a (180 µg) or 2b (1.5 µg/kg) subcutaneously every week and were administered a weight adjusted dose of RBV (600 mg for <60 kg, 800 mg for 60–80 kg, and 1000 mg for >80 kg daily) which is the recommended dosage in Japan. Written informed consent was obtained from each patient and the study protocol conformed to the ethical guidelines of the Declaration of Helsinki and was approved by the institutional ethics review committee. The baseline characteristics are listed in Table 1. For the data mining analysis, 67% of the patients (331 patients) were assigned randomly to the model building group and 33% (165 patients) to the validation group. There were no significant differences in the clinical backgrounds between these two groups.

#### Laboratory and histological tests

Blood samples were obtained before therapy and were analyzed for hematologic tests and for blood chemistry and HCV-RNA. Sequences of ISDR and the core region of HCV were determined by direct sequencing after amplification by reverse-transcription and polymerase chain reaction as reported previously [4,11]. Genetic polymorphism in one tagging SNP located near the *IL28B* gene (rs8099917) was determined by the GWAS or DigiTag2 assay [27]. Homozygosity (CG) or heterozygosity (TG) of the minor sequence was defined as having the *IL28B* minor allele, whereas homozygosity for the major sequence (TT) was

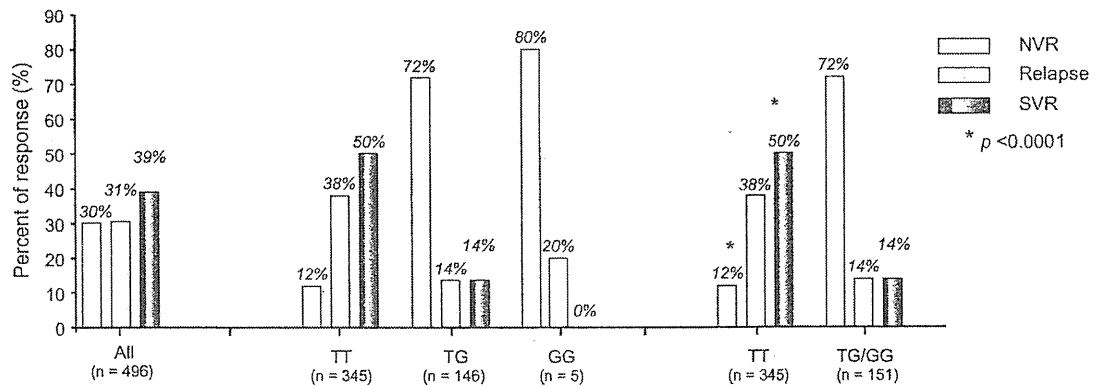


Fig. 1. Association between the *IL28B* genotype (rs8099917) and treatment response. The rates of response to treatment are shown for each rs8099917 genotype. The rate of null virological response (NVR), relapse, and sustained virological response (SVR) is shown. The *p* values are from Fisher's exact test. The rate of NVR was significantly higher ( $p < 0.0001$ ) and the rate of SVR was significantly lower ( $p < 0.0001$ ) in patients with the *IL28B* minor allele compared to those with the major allele.

defined as having the *IL28B* major allele. In this study, NVR was defined as a less than 2 log reduction of HCV-RNA at week 12 and detectable HCV-RNA by qualitative PCR with a lower detection limit of 50 IU/ml (Amplicor, Roche Diagnostic systems, CA) at week 24 during therapy. RVR (rapid virological response) and complete early virological response (cEVR) were defined as undetectable HCV-RNA at 4 weeks and 12 weeks during therapy and SVR was defined as undetectable HCV-RNA 24 weeks after the completion of therapy. Relapse was defined as reappearance of HCV-RNA after the completion of therapy. The stage of liver fibrosis was scored according to the METAVIR scoring system: F0 (no fibrosis), F1 (mild fibrosis: portal fibrosis without septa), F2 (moderate fibrosis: few septa), F3 (severe fibrosis: numerous septa without cirrhosis) and F4 (cirrhosis). Percentage of steatosis was quantified in 111 patients by determining the average proportion of hepatocytes affected by steatosis.

Statistical analysis

Associations between pre-treatment variables and treatment response were analyzed by univariate and multivariate logistic regression analysis. Associations between the *IL28B* polymorphism and sequences of HCV were analyzed by Fisher's exact test. SPSS software v.15.0 (SPSS Inc., Chicago, IL) was used for these analyses. For the data mining analysis, IBM-SPSS Modeler version 13.0 (IBM-SPSS Inc., Chicago, IL) software was utilized as reported previously [26]. The patients used for model building were divided into two groups at each step of the analysis based on split variables. Each value of each variable was considered as a potential split. The optimum variables and cut-off values were determined by a statistical search algorithm to generate the most significant division into two prognostic subgroups that were as homogeneous as possible for the probability of SVR. Thereafter, each subgroup was evaluated again and divided further into subgroups. This procedure was repeated until no additional significant variable was detected or the sample size was below 15. To avoid over-fitting, 10-fold cross validation was used in the tree building process. The reproducibility of the resulting model was tested with the data from the validation patients.

Results

Association between the *IL28B* (rs8099917) genotype and the PEG-IFN/RBV response

The rs8099917 allele frequency was 70% for TT ( $n = 345$ ), 29% for TG ( $n = 146$ ), and 1% for GG ( $n = 5$ ). We defined the *IL28B* major allele as homozygous for the major sequence (TT) and the *IL28B* minor allele as homozygous (GG) or heterozygous (TG) for the minor sequence. The rate of NVR was significantly higher (72% vs. 12%,  $p < 0.0001$ ) and the rate of SVR was significantly lower (14% vs. 50%,  $p < 0.0001$ ) in patients with the *IL28B* minor allele compared to those with the major allele (Fig. 1).

Effect of the *IL28B* polymorphism, substitutions in the ISDR, Core70, and Core91 of HCV on time-dependent clearance of HCV

Patients were stratified according to their *IL28B* allele type, the number of mutations in the ISDR, the amino acid substitutions in Core70 and Core91, and the rate of undetectable HCV-RNA at 4, 8, 12, 24, and 48 weeks after the start of therapy were analyzed (Fig. 2A-D). The rate of undetectable HCV-RNA was significantly higher in patients with the *IL28B* major allele than the minor allele, in patients with two or more mutations in the ISDR compared to none or only one mutation, in patients with arginine (Arg) at Core70 rather than Gln/His, and in patients with leucine (Leu) at Core91 rather than Met. The difference was most significant when stratified by the *IL28B* allele type. The rate of RVR and cEVR was significantly more frequent in patients with the *IL28B* major allele compared with those with the *IL28B* minor allele: 9% vs. 3% for RVR ( $p < 0.005$ ) and 57% vs. 11% for cEVR ( $p < 0.0001$ ). These findings suggest that *IL28B* has the greatest impact on early virological response to therapy.

Association between substitutions in the ISDR and relapse after the completion of therapy

Patients were stratified according to the *IL28B* allele, number of mutations in the ISDR, and amino acid substitutions of Core70 and Core91, and the rate of relapse was analyzed (Fig. 3A and B). Among patients who achieved cEVR, the rate of relapse was significantly lower in patients with two or more mutations in the ISDR compared to those with only one or no mutations (15% vs. 31%,  $p < 0.005$ ) (Fig. 3 B). On the other hand, the relapse rate was not different between the *IL28B* major and minor alleles within patients who achieved RVR (3% vs. 0%) or cEVR (28% vs. 29%) (Fig. 3A). Amino acid substitutions of Core70 and Core91 were not associated with the rate of relapse (data not shown).

Factors associated with response by multivariate logistic regression analysis

By univariate analysis, the minor allele of *IL28B* ( $p < 0.0001$ ), one or no mutations in the ISDR ( $p = 0.03$ ), high serum level of

Research Article

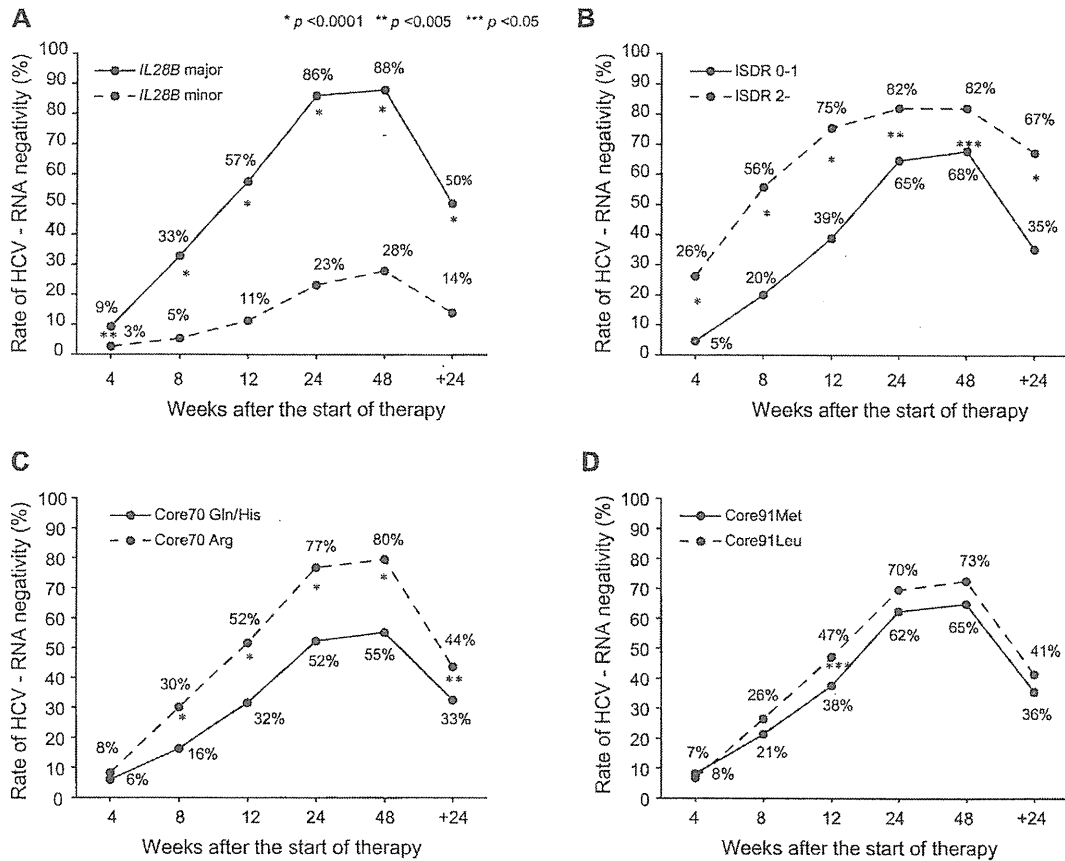


Fig. 2. Effect of *IL28B* mutations in the ISDR, Core70, and Core91 of HCV on time-dependent clearance of HCV. The rate of undetectable HCV-RNA was plotted for serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks) and for 24 weeks after the completion of therapy. Patients were stratified according to (A) the *IL28B* allele (minor allele vs. major allele), (B) the number of mutations in the ISDR (0-1 mutation vs. 2 or more mutations), amino acid substitutions of (C) Core70 (Gln/His vs. Arg), and (D) Core91 (Met vs. Leu). The *p* values are from Fisher's exact test.

HCV-RNA (*p* = 0.035), Gln or His at Core70 (*p* < 0.0001), low platelet counts (*p* = 0.009), and advanced fibrosis (*p* = 0.0002) were associated with NVR. By multivariate analysis, the minor allele of *IL28B* (OR = 20.83, 95%CI = 11.63-37.04, *p* < 0.0001) was associated with NVR independent of other covariates (Table 2). Notably, mutations in the ISDR (*p* = 0.707) and at amino acid Core70 (*p* = 0.207) were not significant in multivariate analysis due to the positive correlation with the *IL28B* polymorphism (*p* = 0.004 for ISDR and *p* < 0.0001 for Core70, Fig. 4).

Genetic polymorphism of *IL28B* also was associated with SVR (OR = 7.41, 95% CI = 4.05-13.57, *p* < 0.0001) independent of other covariates, such as platelet counts, fibrosis, and serum levels of HCV-RNA. Mutation in the ISDR was an independent predictor of SVR (OR = 2.11, 95% CI = 1.06-4.18, *p* = 0.033) but the amino acid at Core70 was not (Table 3).

Factors associated with the *IL28B* polymorphism

Patients with the *IL28B* minor allele had significantly higher serum level of gamma-glutamyltransferase (GGT) and a higher

frequency of hepatic steatosis (Table 4). When the association between the *IL28B* polymorphism and HCV sequences was analyzed, Gln or His at Core70, that is linked to resistance to PEG-IFN and RBV therapy [4,14,15], was significantly more frequent in patients with the minor *IL28B* allele than in those with the major allele (67% vs. 30%, *p* < 0.0001) (Fig. 4). Other HCV sequences with an IFN resistant phenotype also were more prevalent in patients with the minor *IL28B* allele than those with the major allele: Met at Core91 (46% vs. 37%, *p* = 0.047) and one or no mutations in the ISDR (94% vs. 85%, *p* = 0.004) (Fig. 4).

Data mining analysis

Data mining analysis was performed to build a model for the prediction of SVR and the result is shown in Fig. 5. The analysis selected four predictive variables, resulting in six subgroups of patients. Genetic polymorphism of *IL28B* was selected as the best predictor of SVR. Patients with the minor *IL28B* allele had a lower probability of SVR and a higher probability of NVR than those with the major *IL28B* allele (SVR: 14% vs. 50%, NVR: 72% vs.

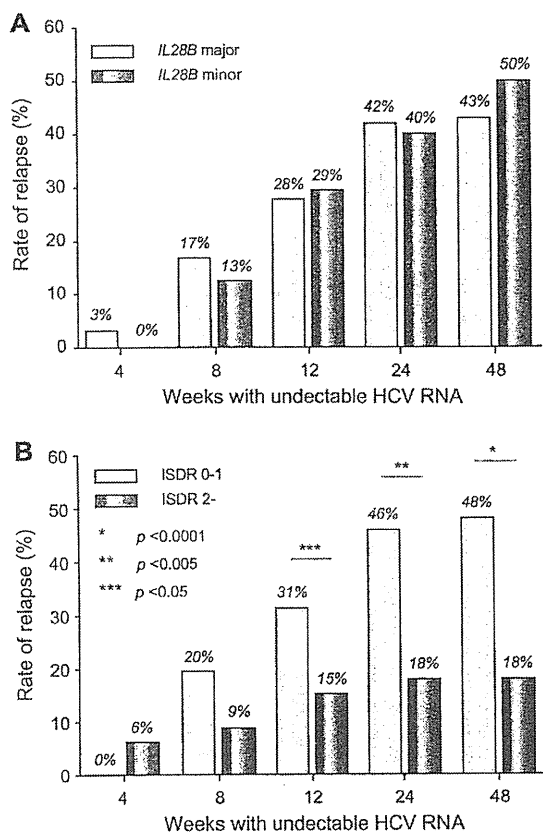


Fig. 3. Association between relapse and the *IL28B* allele or mutations in the ISDR. The rate of relapse was calculated for patients who had undetectable HCV-RNA at serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks). Patients were stratified according to (A) the *IL28B* allele (minor allele vs. major allele) and (B) the number of mutations in the ISDR (0–1 mutation vs. 2 or more mutations). The *p* values are from Fisher's exact test.

12%). After stratification by the *IL28B* allele, patients with low platelet counts ( $<140 \times 10^9/L$ ) had a lower probability of SVR and higher probability of NVR than those with high platelet counts ( $\geq 140 \times 10^9/L$ ): for the minor *IL28B* allele, SVR was 7% vs. 19%, and NVR was 84% vs. 62%, and for the major *IL28B* allele, SVR was 32% vs. 66% and NVR was 16% vs. 8%. Among patients with the major *IL28B* allele and low platelet counts, those with two or more mutations in the ISDR had a higher probability of SVR and lower probability of relapse than those with one or no mutations in the ISDR (SVR: 75% vs. 27%, and relapse: 8% vs. 57%). Among patients with the major *IL28B* allele and high platelet counts, those with a low HCV-RNA titer ( $<600,000$  IU/ml) had a higher probability of SVR and lower probability of NVR and relapse than those with a high HCV-RNA titer (SVR: 90% vs. 61%, NVR: 0% vs. 10%, and relapse: 10% vs. 29%). The sensitivity and specificity of the decision tree were 78% and 70%, respectively. The area under the receiver operating characteristic (ROC) curve of the model was 0.782 (data not shown). The pro-

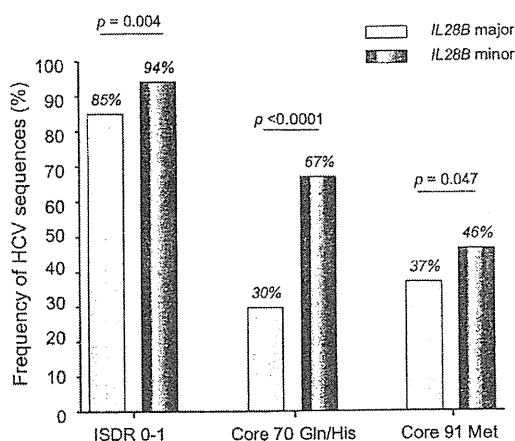


Fig. 4. Associations between the *IL28B* allele and HCV sequences. The prevalence of HCV sequences predicting a resistant phenotype to IFN was higher in patients with the minor *IL28B* allele than those with major allele. (A) 0 or 1 mutation in the ISDR of NS5A, (B) Gln or His at Core70, and (C) Met at Core91. *p* values are from Fisher's exact test.

portion of patients with advanced fibrosis (F3-4) was 39% (84/217) in patients with low platelet counts ( $<140 \times 10^9/L$ ) compared to 13% (37/279) in those with high platelet counts ( $\geq 140 \times 10^9/L$ ).

Validation of the data mining analysis

The results of the data mining analysis were validated with 165 patients who differed from those used for model building. Each patient was allocated to one of the six subgroups for the validation using the flow-chart form of the decision tree. The rate of SVR and NVR in each subgroup was calculated. The rates of SVR and NVR for each subgroup of patients were closely correlated between the model building and the validation patients ( $r^2 = 0.99$  and  $0.98$ ) (Fig. 6).

Discussion

The rate of NVR after 48 weeks of PEG-IFN/RBV therapy among patients infected with HCV of genotype 1 is around 20–30%. Previously, there have been no reliable baseline predictors of NVR or SVR. Because more potent therapies, such as protease and polymerase inhibitor of HCV [28,29] and nitazoxanide [30], are in clinical trials and may become available in the near future, a pre-treatment prediction of the likelihood of response may be helpful for patients and physicians, to support clinical decisions about whether to begin the current standard of care or whether to wait for emerging therapies. This study revealed that the *IL28B* polymorphism was the overwhelming predictor of NVR and is independent of host factors and viral sequences reported previously. The *IL28B* encodes a protein also known as IFN-lambda 3, which is thought to suppress the replication of various viruses including HCV [31,32]. The results of the current study and the findings of the GWAS studies [6–9] may provide the rationale for developing diagnostic testing or an IFN-lambda based therapy for chronic hepatitis C in the future.

## Research Article

Table 2. Factors associated with NVR analyzed by univariate and multivariate logistic regression analysis.

	Univariate			Multivariate		
	Odds ratio	95%CI	p value	Odds ratio	95%CI	p value
Gender: female	0.98	0.67-1.45	0.938	1.29	0.75-2.23	0.363
Age	1.01	0.97-1.01	0.223	0.99	0.97-1.02	0.679
ALT	1.00	1.00-1.00	0.867	1.00	0.99-1.00	0.580
GGT	1.004	1.00-1.01	0.029	1.00	1.00-1.00	0.715
Platelets	0.95	0.91-0.99	0.009	0.92	0.87-0.98	0.006
Fibrosis: F3-4	2.23	1.46-3.42	0.0002	1.97	1.09-3.57	0.025
HCV-RNA: $\geq 600,000$ IU/ml	1.83	1.05-3.19	0.035	2.49	1.17-5.29	0.018
ISDR mutation: $\leq 1$	2.14	1.08-4.22	0.030	0.96	0.78-1.18	0.707
Core 70 (Gln/His)	3.23	2.16-4.78	<0.0001	1.41	0.83-2.42	0.207
Core 91 (Met)	1.39	0.95-2.06	0.093	1.21	0.72-2.04	0.462
IL28B: Minor allele	19.24	11.87-31.18	<0.0001	20.83	11.63-37.04	<0.0001

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase; ISDR, interferon sensitivity determining region; Gln, glutamine; His, histidine; Met, methionine; Minor allele, heterozygote or homozygote of minor allele.

Table 3. Factors associated with SVR analyzed by univariate and multivariate logistic regression analysis.

	Univariate			Multivariate		
	Odds ratio	95%CI	p value	Odds ratio	95%CI	p value
Gender: female	0.81	0.56-1.16	0.253	0.86	0.55-1.35	0.508
Age	0.97	0.95-0.99	0.0003	0.99	0.96-1.01	0.199
ALT	1.00	1.00-1.00	0.337	1.00	1.00-1.01	0.108
GGT	1.00	1.00-1.00	0.273	1.00	1.00-1.00	0.797
Platelets	1.12	1.01-1.16	<0.0001	1.13	1.08-1.19	<0.0001
Fibrosis: F0-2	2.64	1.65-4.22	<0.0001	1.87	1.07-3.28	0.029
HCV-RNA: <600,000 IU/ml	2.49	1.55-3.98	0.0001	2.75	1.55-4.90	0.001
ISDR mutation: $\geq 2$	3.78	2.14-6.68	<0.0001	2.11	1.06-4.18	0.033
Core 70 (Arg)	1.61	1.11-2.28	0.012	0.84	0.52-1.35	0.470
Core 91 (Leu)	1.28	0.88-1.85	0.185	1.26	0.81-1.96	0.300
IL28B: Major allele	6.21	3.75-10.31	<0.0001	7.41	4.05-13.57	<0.0001

ALT, alanine aminotransferase; GGT, Gamma-glutamyltransferase; ISDR, interferon sensitivity determining region; Arg, arginine; Leu, leucine; Major allele, homozygote of major allele.

Among baseline factors, *IL28B* was the most significant predictor of NVR and SVR. Moreover, the *IL28B* allele type was also correlated with early virological response: the rate of RVR and cEVR was significantly high for the *IL28B* major allele compared to the *IL28B* minor allele: 9% vs. 3% for RVR and 57% vs. 11% for cEVR (Fig. 2). On the other hand, the relapse rate was not different between the *IL28B* genotypes within patients who achieved RVR or cEVR (Fig. 3). We believe that optimal therapy should be based on baseline features and a response-guided approach. Our findings suggest that the *IL28B* genotype is a useful baseline predictor of virological response which should be used for selecting the treatment regimen: whether to treat patients with PEG-IFN and RBV or to wait for more effective future therapy including direct acting antiviral drugs. On the other hand, baseline *IL28B* genotype might not be suitable for determining the treatment duration in patients who started PEG-IFN/RBV therapy

and whose virological response is determined because the *IL28B* genotype is not useful for the prediction of relapse. The duration of therapy should be personalized based on the virological response. Future studies need to explore whether the combination of baseline *IL28B* genotype and response-guided approach further improves the optimization of treatment duration.

The SVR rate in patients having the *IL28B* minor allele was 14% in the present study while it was 23% in Caucasians and 9% in African Americans in a study by McCarthy et al. [33]. On the other hand, the SVR rate in patients having the *IL28B* minor allele was 28% in genotypes 1/4 compared to 80% in genotypes 2/3 in a study by Rauch et al. [9]. These data imply that the impact of the *IL28B* polymorphism on response to therapy may be different in terms of race, geographical areas, or HCV genotypes, and that our data need to be validated in future studies including different populations and geographical areas before generalization.



Table 4. Factors associated with IL28B genotype.

	IL28B major allele n = 345	IL28B minor allele n = 151	p value
Gender: male	166 (48%)	84 (56%)	0.143
Age (years)	57 ± 10	57 ± 10	0.585
ALT (IU/L)	79 ± 60	78 ± 62	0.842
Platelets (10 <sup>9</sup> /L)	153 ± 54	155 ± 52	0.761
GGT (IU/L)	51 ± 45	78 ± 91	0.001
Fibrosis: F3-4	76 (22%)	45 (30%)	0.063
Steatosis:			
>10%	16/88 (18%)	13/23 (57%)	0.024
>30%	6/88 (7%)	6/23 (26%)	0.017
HCV-RNA: >600,000 IU/ml	284 (82%)	125 (83%)	1.000

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase.

Four GWAS studies have shown the association between a genetic polymorphism near the *IL28B* gene and response to PEG-IFN plus RBV therapy. The SNPs that showed significant association with response were rs12979860 [8] and rs8099917 [6,7,9]. There is a strong linkage-disequilibrium (LD) between these two SNPs as well as several other SNPs near the *IL28B* gene in Japanese patients [34] but the degree of LD was weaker in Caucasians and Hispanics [8]. Thus, the combination of SNPs is not useful for predicting response in Japanese patients but may improve the predictive value in patients other than Japanese who have weaker LD between SNPs.

Other significant predictors of response independent of *IL28B* genotype were platelet counts, stage of fibrosis, and HCV RVA load. A previous study reported that platelet count is a predictor of response to therapy [35], and the lower platelet count was related with advanced liver fibrosis in the present study. The association between response to therapy and advanced fibrosis independent of the *IL28B* polymorphism is consistent with a recent study by Rauch et al. [9].

There is agreement that the viral genotype is significantly associated with the treatment outcome. Moreover, viral factors such as substitutions in the ISDR of the NS5A region [10] or in the amino acid sequence of the HCV core [4] have been studied in relation to the response to IFN treatment. The amino acid Gln or His at Core70 and Met at Core91 are repeatedly reported to be associated with resistance to therapy [4,14,15] in Japanese patients but these data wait to be validated in different populations or other geographical areas. In this study, we confirmed that patients with two or more mutations in the ISDR had a higher rate of undetectable HCV-RNA at each time point during therapy. In addition, the rate of relapse among patients who achieved cEVR was significantly lower in patients with two or more mutations in ISDR compared to those with only one or no mutations (15% vs. 31%,  $p < 0.05$ ). Thus, the ISDR sequence may be used to predict a relapse among patients who achieved virological response during therapy, while the *IL28B* polymorphism may be used to predict the virological response before therapy. A higher number of mutations in the ISDR are reported to have close association with SVR in Japanese [11-13,15,36] or Asian [37,38] populations but data from Western countries have been controversial [39-42]. A meta-analysis of 1230 patients including 525 patients from Europe has shown that there was a positive correlation

between the SVR and the number of mutations in the ISDR in Japanese as well as in European patients [43] but this correlation was more pronounced in Japanese patients. Thus, geographical factors may account for the different impact of ISDR on treatment response, which may be a potential limitation of our study.

To our surprise, these HCV sequences were associated with the *IL28B* genotype: HCV sequences with an IFN resistant phenotype were more prevalent in patients with the minor *IL28B* allele than those with the major allele. This was an unexpected finding, as we initially thought that host genetics and viral sequences were completely independent. A recent study reported that the *IL28B* polymorphism (rs12979860) was significantly associated with HCV genotype: the *IL28B* minor allele was more frequent in HCV genotype 1-infected patients compared to patients infected with HCV genotype 2 or 3 [33]. Again, patients with the *IL28B* minor allele (IFN resistant genotype) were infected with HCV sequences that are linked to an IFN resistant phenotype. The mechanism for this association is unclear, but may be related to an interaction between the *IL28B* genotype and HCV sequences in the development of chronic HCV infection as discussed by McCarthy et al., since the *IL28B* polymorphism was associated with the natural clearance of HCV [44]. Alternatively, the HCV sequence within the patient may be selected during the course of chronic infection [45,46]. These hypotheses should be explored through prospective studies of spontaneous HCV clearance or by testing the time-dependent changes in the HCV sequence during the course of chronic infection.

How these host and viral factors can be integrated to predict the response to therapy in future clinical practice is an important question. Because various host and viral factors interact in the same patient, predictive analysis should consider these factors in combination. Using the data mining analysis, we constructed a simple decision tree model for the pre-treatment prediction of SVR and NVR to PEG-IFN/RBV therapy. The classification of patients based on the genetic polymorphism of *IL28B*, mutation in the ISDR, serum levels of HCV-RNA, and platelet counts, identified subgroups of patients who have the lowest probabilities of NVR (0%) with the highest probabilities of SVR (90%) as well as those who have the highest probabilities of NVR (84%) with the lowest probability of SVR (7%). The reproducibility of the model was confirmed by the independent validation based on a second group of patients. Using this model, we can rapidly develop an

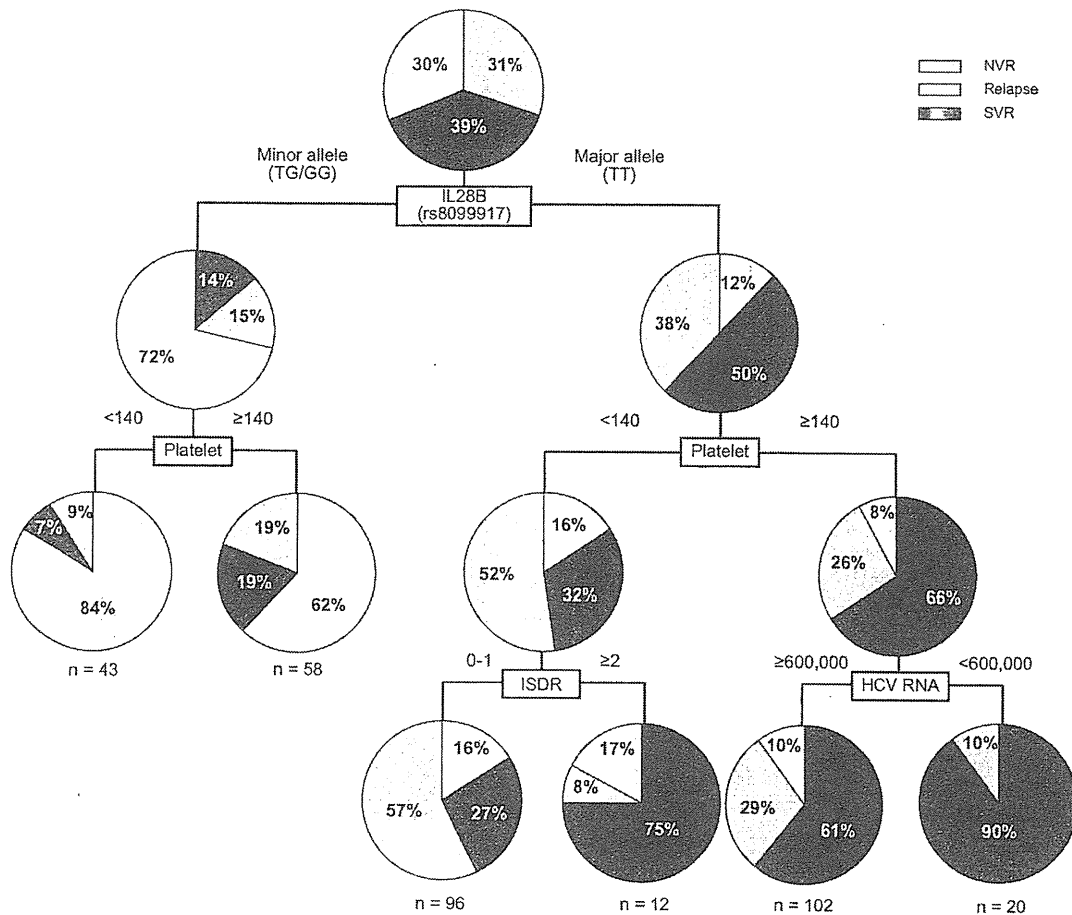
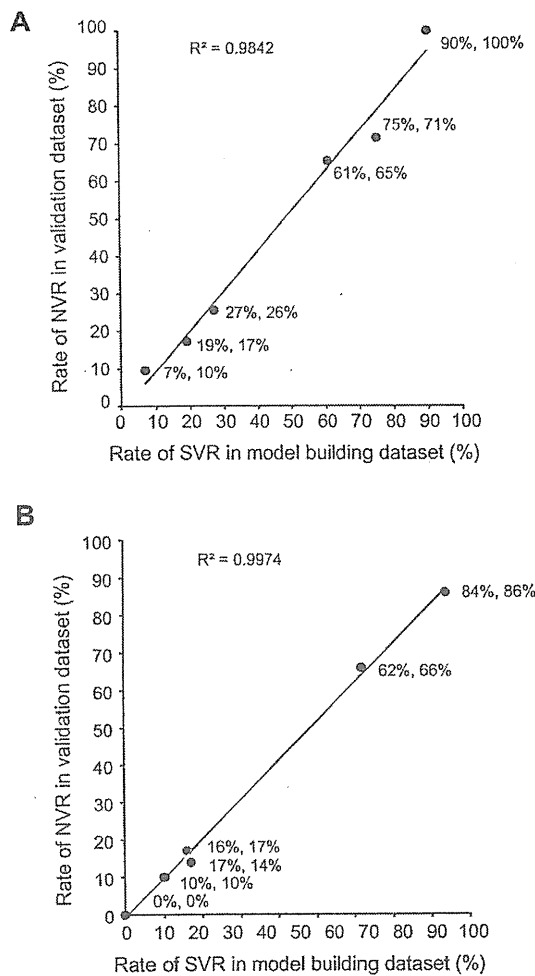


Fig. 5. Decision tree for the prediction of response to therapy. The boxes indicate the factors used for splitting. Pie charts indicate the rate of response for each group of patients after splitting. The rate of null virological response, relapse, and sustained virological response is shown.

estimate of the response before treatment, by simply allocating patients to subgroups by following the flow-chart form, which may facilitate clinical decision making. This is in contrast to the calculating formula, which was constructed by the traditional logistic regression model. This was not widely used in clinical practice as it is abstruse and inconvenient. These results support the evidence based approach of selecting the optimum treatment strategy for individual patients, such as treating patients with a low probability of NVR with current PEG-IFN/RBV combination therapy or advising those with a high probability of NVR to wait for more effective future therapies. Patients with a high probability of relapse may be treated for a longer duration to avoid a relapse. Decisions may be based on the possibility of a response against a potential risk of adverse events and the cost of the therapy, or disease progression while waiting for future therapy.

We have previously reported the predictive model of early virological response to PEG-IFN and RBV in chronic hepatitis C

[26]. The top factor selected as significant was the grade of steatosis, followed by serum level of LDL cholesterol, age, GGT, and blood sugar. The mechanism of association between these factors and treatment response was not clear at that time. To our interest, a recent study by Li et al. [47] has shown that high serum level of LDL cholesterol was linked to the *IL28B* major allele (CC in rs12979860). High serum level of LDL cholesterol was associated with SVR but it was no longer significant when analyzed together with the *IL28B* genotype in multivariate analysis. Thus, the association between treatment response and LDL cholesterol levels may reflect the underlining link of LDL cholesterol levels to *IL28B* genotype. Steatosis is reported to be correlated with low lipid levels [48] which suggest that *IL28B* genotypes may be also associated with steatosis. In fact, there were significant correlations between the *IL28B* genotype and the presence of steatosis in the present study (Table 4). In addition, the serum level of GGT, another predictive factor in our previous study, was signif-



**Fig. 6. Validation of the CART analysis.** Each patient in the validation group was allocated to one of the six subgroups by following the flow-chart form of the decision tree. The rate of (A) sustained virological response (SVR) and (B) null virological response (NVR) in each subgroup was calculated and plotted. The X-axis represents the rate of SVR or NVR in the model building patients and the Y-axis represents those in the validation patients. The rate of SVR and NVR in each subgroup of patients is closely correlated between the model building and the validation patients (correlation coefficient:  $r^2 = 0.98-0.99$ ).

icantly associated with *IL28B* genotype in the present study (Table 4). The serum level of GGT was significantly associated with NVR when examined independently but was no longer significant when analyzed together with the *IL28B* genotype. These observations indicate that some of the factors that we have previously identified may be associated with virological response to therapy through the underlining link to the *IL28B* genotype.

In conclusion, the present study highlighted the impact of the *IL28B* polymorphism and mutation in the ISDR on the pre-treatment prediction of response to PEG-IFN/RBV therapy. A decision model including these host and viral factors has the potential to

support selection of the optimum treatment strategy for individual patients, which may enable personalized treatment.

**Conflict of interest**

The authors who have taken part in this study declare that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

**Financial support**

This study was supported by a grant-in-aid from the Ministry of Health, Labor and Welfare, Japan, (H19-kannen-013), (H20-kannen-006).

**References**

- Ray Kim W. Global epidemiology and burden of hepatitis C. *Microbes Infect* 2002;4 (12):1219-1225.
- Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncales Jr FL, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* 2002;347 (13):975-982.
- 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):958-965.
- Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, et al. Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. *Intervirology* 2005;48 (6):372-380.
- Davis GL, Wong JB, McHutchison JG, Manns MP, Harvey J, Albrecht J. Early virologic response to treatment with peginterferon alfa-2b plus ribavirin in patients with chronic hepatitis C. *Hepatology* 2003;38 (3):645-652.
- Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, et al. Genome-wide association of *IL28B* with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* 2009;41:1105-1109.
- Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, et al. *IL28B* is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 2009;41:1100-1104.
- Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, et al. Genetic variation in *IL28B* predicts hepatitis C treatment-induced viral clearance. *Nature* 2009;461 (7262):399-401.
- Rauch A, Kutalik Z, Descombes P, Cai T, Di Julio J, Mueller T, et al. Genetic variation in *IL28B* is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* 2010;138 (4):1338-1345.
- Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. *J Clin Invest* 1995;96 (1):224-230.
- Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *N Engl J Med* 1996;334 (2):77-81.
- Kurosaki M, Enomoto N, Murakami T, Sakuma I, Asahina Y, Yamamoto C, et al. Analysis of genotypes and amino acid residues 2209 to 2248 of the NS5A region of hepatitis C virus in relation to the response to interferon-beta therapy. *Hepatology* 1997;25 (3):750-753.
- Shirakawa H, Matsumoto A, Joshita S, Komatsu M, Tanaka N, Umemura T, et al. Pretreatment prediction of virological response to peginterferon plus ribavirin therapy in chronic hepatitis C patients using viral and host factors. *Hepatology* 2008;48 (6):1753-1760.
- Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, et al. Predictive factors of early and sustained responses to peginterferon plus ribavirin combination therapy in Japanese patients infected with hepatitis C virus genotype 1b: amino acid substitutions in the core region and low-density lipoprotein cholesterol levels. *J Hepatol* 2007;46 (3):403-410.

## Research Article

- [15] Okanoue T, Itoh Y, Hashimoto H, Yasui K, Minami M, Takehara T, et al. Predictive values of amino acid sequences of the core and NS5A regions in antiviral therapy for hepatitis C: a Japanese multi-center study. *J Gastroenterol* 2009;44 (9):952-963.
- [16] Segal MR, Bloch DA. A comparison of estimated proportional hazards models and regression trees. *Stat Med* 1989;8 (5):539-550.
- [17] LeBlanc M, Crowley J. A review of tree-based prognostic models. *Cancer Treat Res* 1995;75:113-124.
- [18] 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 (19):4322-4329.
- [19] 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: implications on the future of cancer staging. *Surgery* 2002;132 (4):589-602.
- [20] 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 (18):3660-3667.
- [21] Valera VA, Walter BA, Yokoyama N, Koyama Y, Iai 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 (1):34-40.
- [22] 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 (15):5440-5443.
- [23] Thabane M, Simunovic M, Akhtar-Danesh N, Marshall JK. Development and validation of a risk score for post-infectious irritable bowel syndrome. *Am J Gastroenterol* 2009;104 (9):2267-2274.
- [24] Wu BU, Johannes RS, Sun X, Tabak Y, Conwell DL, Banks PA. The early prediction of mortality in acute pancreatitis: a large population-based study. *Gut* 2008;57 (12):1698-1703.
- [25] Fonarow GC, Adams Jr KF, Abraham WT, Yancy CW, Boscardin WJ. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *Jama* 2005;293 (5):572-580.
- [26] 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 (3):251-260.
- [27] Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K. Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. *Anal Biochem* 2007;364 (1):78-85.
- [28] Hezode C, Forestier N, Dusheiko G, Ferenci P, Pol S, Goefer T, et al. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N Engl J Med* 2009;360 (18):1839-1850.
- [29] McHutchison JG, Everson GT, Gordon SC, Jacobson IM, Sulkowski M, Kauffman R, et al. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. *N Engl J Med* 2009;360 (18):1827-1838.
- [30] Rossignol JF, Elfert A, El-Gohary Y, Keeffe EB. Improved virologic response in chronic hepatitis C genotype 4 treated with nitazoxanide, peginterferon, and ribavirin. *Gastroenterology* 2009;136 (3):856-862.
- [31] Marcello T, Grakoui A, Barba-Spaeth G, Machlin ES, Kottenko SV, MacDonald MR, et al. Interferons alpha and lambda inhibit hepatitis C virus replication with distinct signal transduction and gene regulation kinetics. *Gastroenterology* 2006;131 (6):1887-1898.
- [32] Robek MD, Boyd BS, Chisari FV. Lambda interferon inhibits hepatitis B and C virus replication. *J Virol* 2005;79 (6):3851-3854.
- [33] McCarthy JJ, Li JH, Thompson A, Suchindran S, Lao XQ, Patel K, et al. Replicated association between an IL28B Gene Variant and a Sustained Response to Pegylated Interferon and Ribavirin. *Gastroenterology* 2010;138:2307-2314.
- [34] Tanaka Y, Nishida N, Sugiyama M, Tokunaga K, Mizokami M. A-interferons and the single nucleotide polymorphisms: a milestone to tailor-made therapy for chronic hepatitis C. *Hepatol Res* 2010;40:449-460.
- [35] Backus LI, Boothroyd DB, Phillips BR, Mole LA. Predictors of response of US veterans to treatment for the hepatitis C virus. *Hepatology* 2007;46 (1):37-47.
- [36] Mori N, Imamura M, Kawakami Y, Saneto H, Kawaoka T, Takaki S, et al. Randomized trial of high-dose interferon-alpha-2b combined with ribavirin in patients with chronic hepatitis C: correlation between amino acid substitutions in the core/NS5A region and virological response to interferon therapy. *J Med Virol* 2009;81 (4):640-649.
- [37] Hung CH, Lee CM, Lu SN, Lee JF, Wang JH, Tung HD, et al. Mutations in the NS5A and E2-PePHD region of hepatitis C virus type 1b and correlation with the response to combination therapy with interferon and ribavirin. *J Viral Hepat* 2003;10 (2):87-94.
- [38] Yen YH, Hung CH, Hu TH, Chen CH, Wu CM, Wang JH, et al. Mutations in the interferon sensitivity-determining region (nonstructural 5A amino acid 2209-2248) in patients with hepatitis C-1b infection and correlating response to combined therapy of pegylated interferon and ribavirin. *Aliment Pharmacol Ther* 2008;27 (1):72-79.
- [39] Zeuzem S, Lee JH, Roth WK. Mutations in the nonstructural 5A gene of European hepatitis C virus isolates and response to interferon alpha. *Hepatology* 1997;25 (3):740-744.
- [40] Squadrito G, Leone F, Sartori M, Nalpas B, Berthelot P, Raimondo G, et al. Mutations in the nonstructural 5A region of hepatitis C virus and response of chronic hepatitis C to interferon alpha. *Gastroenterology* 1997;113 (2):567-572.
- [41] Sarrazin C, Berg T, Lee JH, Teuber G, Dietrich CF, Roth WK, et al. Improved correlation between multiple mutations within the NS5A region and virological response in European patients chronically infected with hepatitis C virus type 1b undergoing combination therapy. *J Hepatol* 1999;30 (6):1004-1013.
- [42] Murphy MD, Rosen HR, Marousek GI, Chou S. Analysis of sequence configurations of the ISDR, PKR-binding domain, and V3 region as predictors of response to induction interferon-alpha and ribavirin therapy in chronic hepatitis C infection. *Dig Dis Sci* 2002;47 (6):1195-1205.
- [43] Pascu M, Martus P, Hohne M, Wiedenmann B, Hopf U, Schreier E, et al. Sustained virological response in hepatitis C virus type 1b infected patients is predicted by the number of mutations within the NS5A-ISDR: a meta-analysis focused on geographical differences. *Gut* 2004;53 (9):1345-1351.
- [44] Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* 2009;461 (7265):798-801.
- [45] Kurosaki M, Enomoto N, Marumo F, Sato C. Evolution and selection of hepatitis C virus variants in patients with chronic hepatitis C. *Virology* 1994;205 (1):161-169.
- [46] Enomoto N, Kurosaki M, Tanaka Y, Marumo F, Sato C. Fluctuation of hepatitis C virus quasispecies in persistent infection and interferon treatment revealed by single-strand conformation polymorphism analysis. *J Gen Virol* 1994;75 (Pt 6):1361-1369.
- [47] Li JH, Lao XQ, Tillmann HL, Rowell J, Patel K, Thompson A, et al. Interferon-lambda genotype and low serum low-density lipoprotein cholesterol levels in patients with chronic hepatitis C infection. *Hepatology* 1904;51 (6):1904-1911.
- [48] Serfaty L, Andreani T, Giral P, Carbonell N, Chazouilleres O, Poupon R. Hepatitis C virus induced hypobetalipoproteinemia: a possible mechanism for steatosis in chronic hepatitis C. *J Hepatol* 2001;34 (3):428-434.

## Expression of Keratin 19 Is Related to High Recurrence of Hepatocellular Carcinoma after Radiofrequency Ablation

Kaoru Tsuchiya<sup>a</sup> Mina Komuta<sup>b</sup> Yutaka Yasui<sup>a</sup> Nobuharu Tamaki<sup>a</sup>  
Takanori Hosokawa<sup>a</sup> Ken Ueda<sup>a</sup> Teiji Kuzuya<sup>a</sup> Jun Itakura<sup>a</sup>  
Hiroyuki Nakanishi<sup>a</sup> Yuka Takahashi<sup>a</sup> Masayuki Kurosaki<sup>a</sup> Yasuhiro Asahina<sup>a</sup>  
Nobuyuki Enomoto<sup>c</sup> Michiie Sakamoto<sup>b</sup> Namiki Izumi<sup>a</sup>

<sup>a</sup>Department of Gastroenterology and Hepatology, Musashino Red Cross Hospital, and <sup>b</sup>Department of Pathology, School of Medicine, Keio University, Tokyo, and <sup>c</sup>First Department of Internal Medicine, Yamanashi University School of Medicine, Yamanashi, Japan

### Key Words

Hepatocellular carcinoma · Radiofrequency ablation · Recurrence · Keratin · Carcinogenesis · Needle biopsy · Hepatic progenitor cell

### Abstract

**Objective:** Keratin (K) 19 positivity has been reported to be a useful predictive marker for recurrence in patients with hepatocellular carcinoma (HCC) who have undergone hepatic resection. We investigated the clinical usefulness of K19 positivity in patients who had received curative radiofrequency ablation (RFA). **Methods:** We retrospectively evaluated the clinicopathological features, including imaging and K19 expression, in 246 patients with HCC who were within the Milan criteria and had received curative RFA. Using a two-step insertion method, tumor biopsies were obtained just prior to RFA and were evaluated histologically. **Results:** Tumor seeding due to liver biopsy and RFA was not observed. Ten patients (4.1%) had K19-positive HCC. Imaging findings were similar between K19-positive and -negative HCC ( $p = 0.187$ ). Nine out of 10 patients (90%) who had K19-positive HCC had

recurrence of HCC after RFA, and intrahepatic recurrences were observed within 12 months in 6 out of 10 (60.0%). K19 positivity was a significant risk factor for recurrence ( $p < 0.0001$ ) and early recurrence ( $<1$  year after RFA;  $p = 0.012$ ). K19 expression ( $p = 0.016$ ) was an independent risk factor for tumor status exceeding the Milan criteria after RFA. **Conclusion:** Expression of K19 is related to high recurrence of HCC after curative RFA.

Copyright © 2011 S. Karger AG, Basel

### Introduction

Radiofrequency ablation (RFA) is regarded as an important treatment modality for hepatocellular carcinoma (HCC) [1–4], and its efficacy, especially for tumors  $<2$  cm in diameter, is better than that of ethanol and nearly comparable to that of surgical resection [5]. In addition, RFA

Kaoru Tsuchiya and Mina Komuta contributed equally to this work. Michiie Sakamoto and Namiki Izumi contributed equally to this work.

### KARGER

Fax +41 61 306 12 34  
E-Mail karger@karger.ch  
www.karger.com

© 2011 S. Karger AG, Basel  
0030-2414/11/0804-0278\$38.00/0

Accessible online at:  
www.karger.com/ocl

Namiki Izumi, MD, PhD  
Department of Gastroenterology and Hepatology  
Musashino Red Cross Hospital  
1-26-1 Kyonan-cho, Musashino-shi, Tokyo 180-8610 (Japan)  
Tel. +81 422 32 3111, E-Mail nizumi@musashino.jrc.or.jp

is considered to be a bridge to liver transplantation because the prolonged waiting time for cadaveric livers leads to dropouts from the waiting list [6]. Tumor recurrence after curative RFA has been a problem, as it is after hepatic resection. Tumor size (>3 cm in diameter) [7], time after treatment (>1 year) [7], the number of HCC nodules [8] and hepatitis C virus (HCV) infection [8] have been reported to be risk factors for intrahepatic tumor recurrence after curative RFA. Moreover, primary technical failure is reported to be a risk factor for tumor progression beyond the Milan criteria after RFA [9].

Keratin (K) 19, which is considered to be a biliary/hepatic progenitor cell marker [10], has attracted attention as a useful predictive marker for detecting the more aggressive HCCs after curative resection, because tumors with K19 expression have a poorer prognosis [11, 12] and higher rates of recurrence [13, 14] and lymph node metastasis [12] than K19-negative HCC. In these previous studies, surgical specimens were investigated and K19 positivity was defined as expression in >5% of tumor cells [11-14].

As a result, one would expect that K19 expression might be a useful predictive marker for detecting HCC with a worse outcome after RFA, especially regarding tumor recurrence. To the best of our knowledge, the correlation between clinicopathological features and K19 expression has not been investigated in HCC patients treated by RFA. Therefore, we performed a clinicopathological study on 246 HCC cases treated with RFA and investigated the relationship between the K19 expression and recurrence and prognosis after treatment.

## Methods

### Patients

Between April 1999 and February 2010, 1,284 patients were admitted to the Musashino Red Cross Hospital for the first treatment of HCC. A total of 684 patients were treated with RFA as the initial therapy for HCC. Ablation therapy was chosen either because the patients were considered not to be suitable for resection ( $n = 323$ ), when considering impairment of liver function, number and distribution of the tumors as well as cardiopulmonary dysfunction, or because they preferred ablation and provided informed consent ( $n = 361$ ), despite surgery also being feasible. From the outset, 172 patients were excluded because RFA was performed without tumor biopsy. Therefore, 512 consecutive patients, on whom tumor biopsies had been performed before RFA, were included and we evaluated these specimens retrospectively. The result of retrospective analysis was that there were 57 patients with no residual samples, 119 patients with no tumorous lesion and 9 patients with no definitive histological diagnosis because of a small and/or fragmented specimen. The remaining specimens

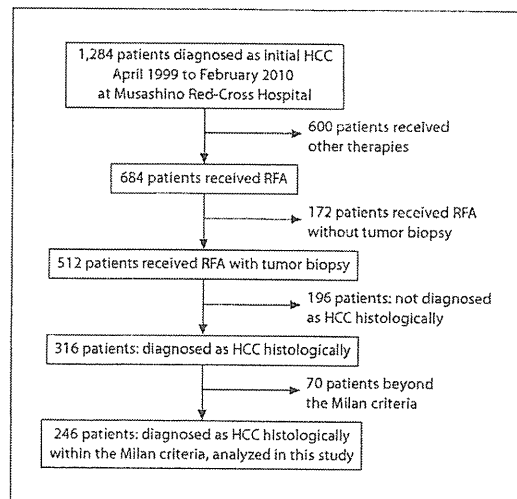


Fig. 1. Flow chart summarizing the patient selection for the study.

were diagnosed as HCC in 316 patients, as dysplastic nodule in 6 patients, as adenocarcinoma in 4 patients and as neuroendocrine tumor in 1 patient. Seventy patients were excluded, because their states of HCC were beyond the Milan criteria ( $\leq 3$  cm and up to 3 nodules, or  $\leq 5$  cm and a single nodule). Therefore, 246 consecutive patients, on whom tumor biopsies had been performed before RFA and diagnosed as HCC retrospectively, were included in the study (fig. 1). The inclusion criteria for receiving RFA were as follows: total bilirubin concentration  $<3.0$  mg/dl, platelet count  $>3 \times 10^9/\text{mm}^3$ , prothrombin activity  $>50\%$  (approximately equal to an international normalized ratio of 1.5) and Child-Pugh score  $<8$  points. Ascites were controlled by administration of diuretics before RFA. Patients with macroscopic vascular invasion or extrahepatic metastases were excluded. The criteria of the International Union against Cancer were used for TNM classification [15]. Written informed consent was obtained from all patients, and the study was approved by the ethics committee at Musashino Red Cross Hospital, in accordance with the Declaration of Helsinki.

### Diagnosis of HCC

All the patients were diagnosed as having HCC on the basis of tumor markers and a combination of typical imaging findings on ultrasonography (US) and dynamic computed tomography (CT), according to the American Association for the Study of Liver Diseases and the Japan Society of Hepatology guidelines [1, 16]. When patients had 2 or 3 HCC nodules, a needle biopsy was taken from the main nodule. The histological diagnosis of HCC was based on the World Health Organization criteria [17].

For the evaluation of vascularity and Kupffer cell activity of the target nodule, CT during arteriography (CTHA) and CT dur-

ing arteriportography (CTAP) were performed in 188 (76.4%) patients, superparamagnetic iron oxide-enhanced magnetic resonance imaging (SPIO-MRI) was performed in 194 (78.8%) patients and gadolinium-ethoxybenzyl-diethylenetriamine pentaacetic acid magnetic resonance imaging (Gd-EOB-DTPA) was performed in 47 patients (19.1%), from March 2008. For triple-phase dynamic CT scans, arterial, portal and equivalent phases were 35, 70 and 150 s, respectively, after injection of contrast agent. Spiral CT scans were obtained from 3- to 5-mm-thick sections. Board-certified radiologists diagnosed HCC on the basis of typical patterns, such as an early-phase hyperattenuation area and late-phase hypoattenuation on dynamic CT. According to previous studies, the sensitivity of the diagnosis of HCC in CTHA/CTAP is higher than that of spiral CT. The diagnosis of HCC in CTHA/CTAP is hyperattenuation area in CTHA and hypoattenuation area in CTAP. It has been reported that the presence of Kupffer cells could be evaluated, and this was defined by a hyper-intensity area in the T2\* image of SPIO-MRI as a typical imaging finding of HCC. Gd-EOB-DTPA MRI is a liver-specific contrast-enhanced agent, and hypointensity in the hepatobiliary phase is a typical imaging finding. We started to perform Gd-EOB-DTPA MRI instead of SPIO-MRI from March 2008, because it was reported that the sensitivity of Gd-EOB-DTPA MRI was superior to SPIO-MRI for the diagnosis of HCC.

#### *Tumor Biopsy and RFA*

There are 24 operators who participated in this study. They are specialized liver physicians who have great experiences in performing percutaneous ethanol injection for HCC, percutaneous tumor biopsy for liver tumor, percutaneous liver biopsy for hepatitis, percutaneous hepatobiliary drainage for obstructive jaundice, or percutaneous liver abscess drainage. A needle-guiding technique was used, consisting of an initial guided needle and a secondary outer needle (two-step insertion method). This method was reported by another center previously [18] and involves the initial insertion of a 21-gauge needle (Silux, Saitama, Japan) just adjacent to the tumor under real-time US guidance, and using this to insert a 14-gauge Daimon outer needle (Silux), also just adjacent to the tumor. After removal of the inner needle, an 18-gauge biopsy needle was inserted to obtain the tumor tissue sample. After removal of the biopsy needle, a 17-gauge cooled-tip electrode was inserted into the targeted tumor. The electrode, with a 2- or 3-cm exposed tip, was connected to a 480-kHz RF Generator (Radionics, Burlington, Mass., USA), which produces 200 W at 50  $\Omega$  of impedance [19, 20]. The equipment also allows the measurement of power output, tissue impedance and electrode tip temperature. A tip temperature of 10–20°C was maintained by infusion of chilled water through a peristaltic pump. After insertion of the electrode into the tumor, ablation was performed at 60 W for the 3-cm exposed tip and 40 W for the 2-cm exposed tip. The power was increased to 140 W at a rate of 10–20 W/min. When a rapid increase in impedance was observed during thermal ablation, the output was reduced. The duration of a single ablation was 12 min. After RF exposure, the pump was stopped and the temperature of the needle tip was measured. When the temperature of the electrode tip was >60°C, ablation was defined as being sufficient. When the target nodule was >2 cm in diameter, multiple needle insertions and ablations were performed in 1 nodule to achieve complete necrosis. A session was defined as a single intervention consisting of  $\geq 1$  ablations performed on  $\geq 1$  tumors at

the same time. After completion of nodule ablation, the intrahepatic needle track was treated by thermocoagulation to avoid needle track seeding. Finally, a mixture of gelatin sponge particles (Gelfoam<sup>®</sup>; Upjohn, Kalamazoo, Mich., USA) was injected into the puncture route. All procedures were completed within 15–20 min. After each session of RFA, a dynamic CT scan (section thickness 5 mm) was performed to evaluate the efficacy of ablation. Complete ablation of HCC was defined as non-enhancement of the lesion, including the whole surrounding liver parenchyma. The ablative margin was shown as the boundary between the low density area as ablated area and the isodensity area as surrounding normal liver parenchyma. The residual portion of the tumor was treated by additional RFA within a few days of the post-treatment CT scan. Follow-up consisted of monthly serial measurements of tumor markers [ $\alpha$ -fetoprotein (AFP) and des- $\gamma$ -carboxy prothrombin (DCP)], US examination every 2 months and dynamic CT every 3 months. We checked various complications of RFA with conventional contrast-enhanced CT and blood examination at day 1 after RFA.

#### *Tumor Recurrence*

Recurrence of HCC was defined as an early enhancement area on dynamic CT, concomitant with late wash out. Two types of recurrence, local tumor progression and distant intrahepatic recurrence, were identified. Local tumor progression was defined as an enhancing area located adjacent to the ablated area [21], while distant intrahepatic recurrence referred to the appearance of a new tumor in the liver, distant from the ablated area. Early recurrence was defined as a recurrence within 12 months of the initial RFA.

#### *Immunohistochemistry*

Immunohistochemistry using antibodies against K19 (1:100, BA17, Dakocytomation, Glostrup, Denmark) was performed on paraffin-embedded sections from 246 needle biopsy specimens. The slides were reviewed by 2 independent pathologists (M. Komuta and M. Sakamoto). Expression of K19 was considered positive if >5% of tumor cells were stained according to the expected pattern of reactivity.

#### *Statistical Analysis*

Categorical variables were compared with the  $\chi^2$  test and continuous variables with the Mann-Whitney test; a p value <0.05 was considered statically significant. Continuous variables were expressed as the mean  $\pm$  standard deviation. The imaging findings were compared with the  $\chi^2$  test between K19-positive and -negative patients. Overall survival was defined as the interval between treatment and death or the date of the last follow-up or the date of the most recent follow-up visit. Probability of recurrence-free survival was defined as the interval between treatment and the date of HCC recurrence.

Univariate analysis was performed to identify clinical and biological parameters (sex, age, etiology, prothrombin activity, albumin, bilirubin levels, Child-Pugh class, serum AFP level, serum DCP level) and tumor factors (size, number, tumor stage, tumor differentiation, K19 expression) predicting overall survival, recurrence-free survival and the interval beyond the Milan criteria.

Survival curves were computed according to the Kaplan-Meier method and compared by the log-rank test. All variables with a p value <0.05 were subjected to multivariate analysis by Cox's

**Table 1.** Comparison of clinicopathological features of patients (n = 246) with HCC with and without K19 expression

Features	K19 >5% (n = 10)	K19 ≤5% (n = 236)	p value
Mean age ± SD, years	70 ± 8	68 ± 8	0.541
Sex, male/female	2/8	146/90	0.016
<i>Clinical and laboratory data</i>			
Mean AFP, ng/ml	489 [52.1]	12 [16.2]	0.062
Mean DCP, mAU/ml	42 [25]	321 [22]	0.773
Child-Pugh score A/B	8/2	200/36	0.655
Total bilirubin, mg/dl	0.9 ± 0.5	0.8 ± 0.4	0.480
Albumin, g/dl	3.4 ± 0.7	3.6 ± 0.5	0.137
PT, %	97 ± 12	92 ± 15	0.375
<i>Pathology</i>			
Tumor size, mm	24 ± 7	22 ± 8	0.392
Tumor number	1.3 ± 0.7	1.2 ± 0.6	0.891
Vascular invasion, yes/no	0/10	0/236	
Tumor differentiation well/moderate/poor	0/8/2	108/126/2	<0.0001
TNM stage I/II	8/2	183/53	0.855
Lymph node involvement yes/no	0/10	0/236	
Metastasis, yes/no	0/10	0/236	
<i>Major associated liver diseases</i>			
HBsAg+	1 (10)	24 (10.1)	0.895
HCV Ab+	9 (90)	189 (80.1)	
ALD	0	8 (3.4)	
NASH	0	2 (0.8)	
Unknown etiology	0	13 (5.6)	

Figures in parentheses are percentages; figures in brackets are medians. PT = Prothrombin time; HBsAg = hepatitis B surface antigen; HCV Ab = HCV antibody; ALD = alcoholic liver disease; NASH = non-alcoholic steatohepatitis.

proportional hazards model to assess their value as independent predictors.

All statistical analyses were performed using StatView (version 5.0) software (Abacus Concepts, Berkeley, Calif., USA).

## Results

### *Proportion of HCCs Expressing K19*

The biopsy number was 272, and the median length of our biopsy specimens was 8.2 ± 4.0 mm. In 117 cases, the specimens were <1 cm, and ≥1 cm in 155 cases. Pathological diagnosis and K19 staining were practicable in all specimens <1 cm. Expression of K19 in >5% of tumor

**Table 2.** Comparison of the image findings of patients with HCC with and without K19 expression

	K19 positive >5% (n = 10)	K19 negative (n = 236)	p value
CECT arterial phase high density	10/10	200/235	0.187
CTHA high density	7/7	159/181	0.326
CTAP low density	7/7	179/181	0.779
SPIO-MRI T2*	10/10	175/184	0.473
EOB-MRI			
Hepatobiliary phase low intensity	-	46/47	-

cells was observed in HCCs from 10 of 246 patients (4.1%). Two of the 10 HCCs (20.0%) were poorly differentiated, and 8 (80.0%) were moderately differentiated. None of the well-differentiated HCCs showed K19 positivity. Among the 10 patients with K19-positive HCCs, 2 had a HCC nodule >3 cm and 8 had HCC nodules ≤3 cm in diameter. The 8 HCC nodules with K19 positivity ≤3 cm in diameter were moderately (n = 7) and poorly differentiated HCCs (n = 1).

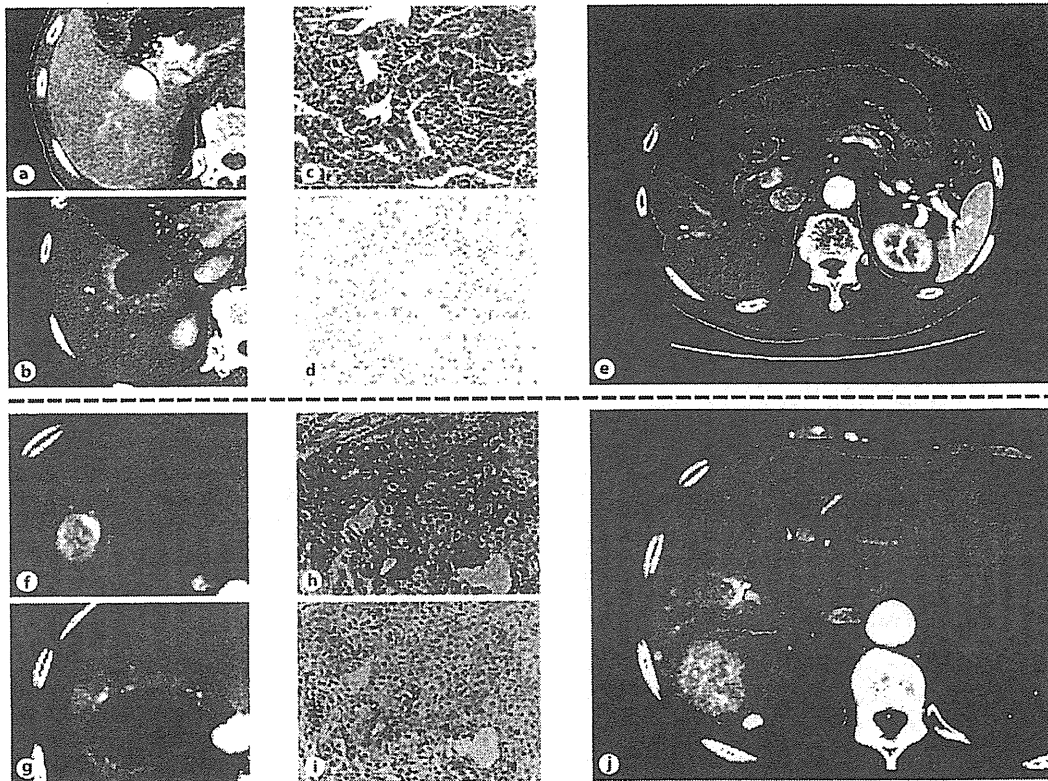
### *Clinicopathological Characteristics of Patients with HCC in Relation to Expression of K19*

The clinicopathological characteristics of the patients in relation to K19 expression in HCCs are shown in table 1. The proportion of well-differentiated HCCs was significantly lower among K19-positive HCC patients (p < 0.0001). K19 expression was more frequent among female than among male patients (p = 0.016). There were no significant differences in age, clinical laboratory data, tumor size, number of tumor nodules, tumor stage in TNM classification or etiology between K19-positive and -negative HCC patients. There was no significant difference in tumor location (near the major vessels, bile ducts and organs) between K19-positive and -negative patients. The number of RFA sessions did not differ significantly between K19-positive and -negative HCC patients. Serum AFP before initial RFA was not evaluated in 1 patient.

### *Imaging Characteristics of HCCs in Relation to Expression of K19*

Comparison of the various imaging findings, according to vascular profiling, and in relation to K19 expres-





**Fig. 2. a–e** A patient with K19-negative HCC: a 70-year-old man with chronic hepatitis (anti-HCV positive). The HCC (25 mm in diameter, in segment 6) showed an early enhancement area by dynamic CT (a). Dynamic CT at 1 day after RFA (b). On histological investigation, the tumor showed moderately differentiated HCC on H&E staining (c), and K19 expression was negative in tumor cells (d). The HCC did not show early enhancement on dynamic CT 4 years and 10 months after curative RFA (e). **f–j** A patient with

K19-positive HCC: a 72-year-old female with chronic hepatitis (anti-HCV positive). The HCC (25 mm in diameter, in segment 8) showed an early enhancement area by dynamic CT (f). CT 1 day after RFA (g). On histological investigation, the tumor showed moderately differentiated HCC on H&E staining (h), and K19-positive cells were seen in the tumor (i). Five months after RFA, the HCC showed intrahepatic recurrence beyond the Milan criteria (j).

sion, is shown in table 2. These imaging findings were consistent with the histological diagnosis, as determined by pretreatment needle biopsy.

All K19-positive HCCs showed typical HCC images, such as hypervascularity at the arterial phase, hypovascularity at the portal and equilibrium phases in dynamic CT, and hyperintensity at the T2\* image in SPIO-MRI. There was no significant difference between K19-positive and -negative patients in terms of the imaging findings.

#### *Recurrence of HCC after RFA*

The median follow-up period was 34.0 months (range 65 days to 10.3 years). A recurrence of HCC was diagnosed at least once during the follow-up period in 156 patients (63.4%). The cumulative recurrence-free survival at 1, 3 and 5 years was 69.9, 26.6 and 12.2%, respectively. Among the 156 patients with recurrent HCC, 14 (8.9%) had local tumor progression and 142 (91.1%) had distant intrahepatic recurrences. Five of 14 patients (35%) who had local tumor progression had K19-positive HCC and

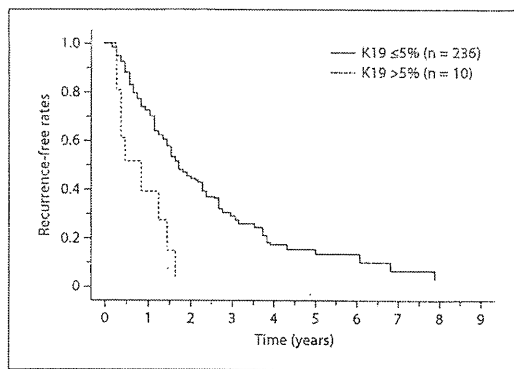


Fig. 3. The cumulative recurrence-free survival rate in patients with K19-positive (>5%) HCC was significantly lower than that in patients with K19-negative HCC ( $p = 0.0001$ ).

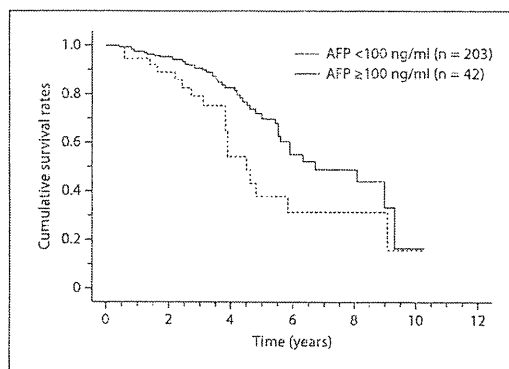


Fig. 4. The cumulative overall survival rate in patients with AFP  $\geq 100$  ng/ml was significantly lower than that in patients with AFP <100 ng/ml ( $p = 0.026$ ).

3 of 5 patients with K19-positive HCC (60%) showed vascular invasion at the local tumor progression. Nine of 10 patients (90.0%) with K19-positive HCC had recurrences after initial treatment and 6 of 10 (60.0%) were detected within 1 year of initial curative RFA. On the other hand, 147 of 236 patients (62.2%) with K19-negative HCC had recurrences, and only 58 patients (24.5%) had recurrences within 1 year after RFA. There were no patients with K19-negative HCC who showed vascular invasion at the local tumor progression. Patients with K19-positive HCC were more likely to have an early recurrence of HCC (<1 year after RFA) than patients with K19-negative HCC ( $p = 0.012$ ). The typical cases are shown in figure 2. The median recurrence-free survival in patients with K19-positive HCC was 194 days (range 93–635), while in patients with K19-negative HCC it was 446 days (range 65–2,978). Patients with K19-positive HCC had a significantly shorter recurrence-free survival than patients with K19-negative HCC ( $p = 0.0001$ ) (fig. 3). The recurrence type, local tumor progression or distant intrahepatic recurrence differed between K19-positive and -negative patients. Local tumor progression was significantly higher in K19-positive patients than in K19-negative patients ( $p < 0.0001$ ). Table 3 shows the results of univariate and multivariate analyses of prognostic factors for recurrence-free survival. In the multivariate analysis, K19 expression, the number of HCC nodules and total bilirubin  $\geq 2$  mg/dl were significant independent risk factors for HCC recurrence in all patients.

The percentage of distant metastasis and major portal invasion (VP3–4) was significantly higher in K19-positive than in K19-negative patients ( $p < 0.0001$ ). Distal metastasis was detected in the lung (2 patients) and lymph node (1 patient), and major portal invasion was detected in 3 patients.

#### Risk Factors for Poor Prognosis

There was no patient who received liver transplantation in this study. Fifty-seven of 246 patients (23.1%) died during the follow-up period. The cause of death was progression of HCC in 37 patients, hepatic failure in 16 patients and causes unrelated to the liver in 4 patients. The overall survival rates for all patients were 97.2, 88.7 and 63.4% at 1, 3 and 5 years, respectively. A serum AFP level  $\geq 100$  ng/ml ( $p = 0.034$ ), a total bilirubin level  $\geq 2$  mg/dl ( $p < 0.0001$ ) and female sex ( $p = 0.018$ ) were identified as risk factors for a poor prognosis in HCC in both univariate and multivariate analyses (table 4). Patients with high serum AFP levels ( $\geq 100$  ng/ml) had significantly lower overall survival rates than patients with low serum AFP levels ( $p = 0.026$ ) (fig. 4).

On the other hand, age ( $\geq 65$  years), albumin concentration ( $\leq 3.5$  g/dl), prothrombin time ( $\leq 70\%$ ), DCP ( $\geq 100$  mAU/ml), tumor size, the number of HCC nodules and K19 expression were not significant risk factors for poor prognosis in the univariate analysis (table 4).

**Table 3.** Risk factors associated with recurrence-free survival in 246 patients with HCC after complete ablation by RFA

Risk factor	Univariate			Multivariate		
	RR	95% CI	p	RR	95% CI	p
Age <65 years	1.43	1.02–2.02	0.037	1.28	0.90–1.81	0.163
Sex, female	1.24	0.90–1.71	0.162			
Total bilirubin $\geq$ 2 mg/dl	2.50	1.02–6.25	0.034	2.70	1.08–6.66	0.032
Albumin $\leq$ 3.5 g/dl	1.12	0.81–1.56	0.492			
PT $\leq$ 70%	1.28	0.73–2.22	0.394			
AFP $\geq$ 100 ng/ml	1.42	0.95–2.12	0.087			
DCP $\geq$ 100 mAU/ml	1.08	0.68–1.69	0.790			
Tumor size $>$ 3.0 cm	1.08	0.70–1.69	0.713			
2 or 3 tumor nodules	2.29	1.58–3.33	$<$ 0.0001	2.28	1.56–3.32	$<$ 0.0001
K19 positive ( $>$ 5%)	3.57	1.75–7.14	0.0004	3.44	1.72–7.14	0.0005

RR = Risk ratio; CI = confidence interval; PT = prothrombin time.

**Table 4.** Risk factors associated with poor prognosis in 246 patients with HCC after complete ablation by RFA

Risk factor	Univariate			Multivariate		
	RR	95% CI	p	RR	95% CI	p
Age <65 years	1.19	0.68–2.09	0.527			
Sex, female	2.03	1.18–3.46	0.009	1.92	1.11–3.30	0.018
Total bilirubin $\geq$ 2 mg/dl	12.5	4.54–33.3	$<$ 0.0001	10.0	3.70–33.3	$<$ 0.0001
Albumin $\leq$ 3.5 g/dl	1.25	0.71–2.17	0.450			
PT $\leq$ 70%	1.49	0.59–3.84	0.674			
AFP $\geq$ 100 ng/ml	1.88	1.06–3.44	0.030	1.88	1.05–3.33	0.034
DCP $\geq$ 100 mAU/ml	1.06	0.53–2.12	0.880			
Tumor size $>$ 3.0 cm	1.12	0.44–1.78	0.730			
2 or 3 tumor nodules	1.23	0.67–2.26	0.492			
K19 positive ( $>$ 5%)	1.29	0.46–3.57	0.632			

RR = Risk ratio; CI = confidence interval; PT = prothrombin time.

*Risk Factors for Exceeding the Milan Criteria after RFA*

Patients with K19-positive HCC exceeded the Milan criteria within 16.8 months. Multivariate analyses showed that K19 expression, high levels of DCP ( $\geq$ 100 mAU/ml), tumor number and total bilirubin  $\geq$ 2 mg/dl were significant risk factors for tumor status exceeding the Milan criteria after curative RFA (table 5; fig. 5).

*Complications*

Most patients had mild pain or discomfort during RFA. Intraperitoneal hemorrhage and biloma were not

seen in any patient. None of the patients developed dissemination of HCC, or skin or peritoneal metastases. There was no fatal complication.

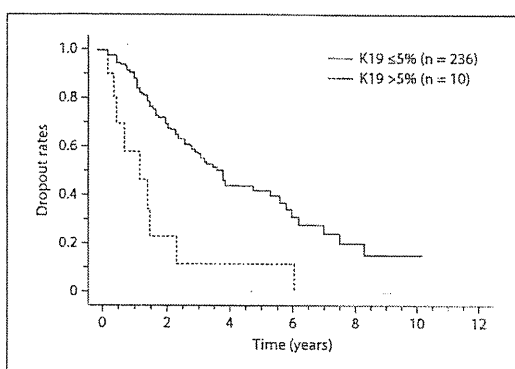
*Percentage of K19 Stain*

We also analyzed another percentage of K19 stain ( $>$ 1%). Thirteen of 246 patients had K19-positive ( $>$ 1%) HCC and 12 of 13 patients with K19-positive ( $>$ 1%) HCC had recurrences beyond the Milan criteria. Nine of 12 (75.0%) were detected with recurrence of HCC within 1 year of initial curative RFA. The final results were the same for K19 positivity ( $>$ 5 and  $>$ 1%, respectively). The

**Table 5.** Risk factors associated with exceeding the Milan criteria in 246 patients with HCC after complete ablation by RFA

Risk factor	Univariate			Multivariate		
	RR	95% CI	p	RR	95% CI	p
Age <65 years	1.63	1.08–2.45	0.018	1.17	0.75–1.83	0.463
Sex, female	1.16	0.78–1.72	0.457			
Total bilirubin $\geq 2$ mg/dl	2.94	1.05–8.33	0.039	3.57	1.25–10.0	0.017
Albumin $\leq 3.5$ g/dl	0.97	0.64–1.47	0.857			
PT $\leq 70\%$	0.89	0.41–1.96	0.763			
AFP $\geq 100$ ng/ml	2.17	1.38–3.44	0.0008	1.56	0.96–2.50	0.077
DCP $\geq 100$ mAU/ml	2.32	1.42–3.70	0.0007	2.08	1.26–3.44	0.004
Tumor size $>3.0$ cm	1.03	0.61–1.72	0.914			
2 or 3 tumor nodules	2.98	1.91–4.64	$<0.0001$	3.05	1.91–4.88	$<0.0001$
K19 positive ( $>5\%$ )	3.70	1.81–7.69	0.0003	2.47	1.19–5.18	0.016

RR = Risk ratio; CI = confidence interval; PT = prothrombin time.



**Fig. 5.** The cumulative rate of exceeding the Milan criteria in patients with K19-positive HCC was significantly higher than that in patients with K19-negative HCC ( $p < 0.0001$ ).

rate of recurrence and dropout from the Milan criteria were significantly higher in the patients with K19-positive ( $>1\%$ ) than in the patients with K19-negative HCC (data not shown).

### Discussion

RFA therapy for HCC has been shown to achieve excellent results in appropriately selected patients [2–5]. However, recurrence of tumors is a serious impediment to im-

proving the prognosis for patients treated with curative RFA. Therefore, several factors have been investigated as potential predictive markers for recurrence after curative RFA [7–9]. Recently, K19 was proposed as an independent prognostic factor for HCC [11–14]. However, these investigations were performed on surgically resected cases only and not on tumor biopsies. Although tumor biopsy is controversial because of potential complications such as tumor seeding [22], it would be beneficial to clinicians and patients to predict the individual tumor characteristics from a biopsy. Until now, the relationship between K19 expression and tumor recurrence after RFA treatment has not been assessed. Therefore, we have investigated the relationship between K19 expression in tumor biopsies and the clinicopathological findings in HCC. In this study, we investigated K19 expression in biopsy specimens taken just prior to the RFA session, and K19 expression ( $>5\%$ ) was demonstrated in 10 of 246 patients (4.1%). Because most of our patients were in early stage (within the Milan criteria) and 108 of 246 patients (43.9%) had well-differentiated HCC, the positive rate of K19 stain in our study was lower than that in surgical specimens.

We also analyzed another percentage of K19 stain ( $>1\%$ ) and the final results were the same for K19 positivity ( $>5$  and  $>1\%$ , respectively). K19 expression ( $>1\%$ ) was a statistically significant independent predictor for recurrence of HCC after RFA. Although the amount of tissue obtained by tumor biopsy is small compared to resected material, present data suggest that even biopsy can provide meaningful data on tumor recurrence irrespective of the percentage of K19 positivity (1 or 5%) (online sup-