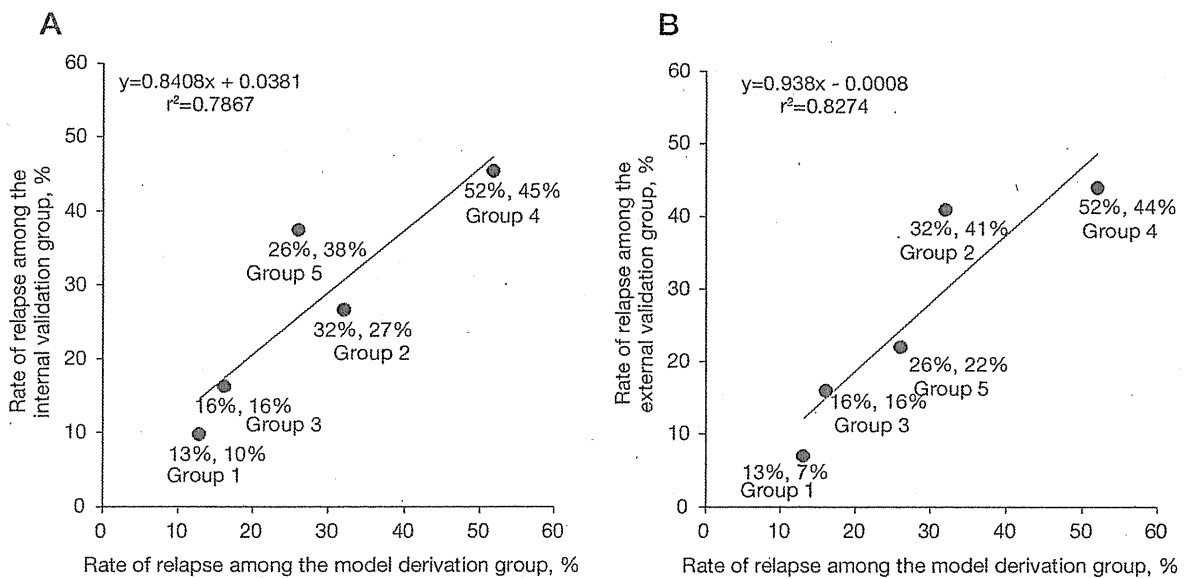


model derivation and internal validation ($n=718$). The relapse rate decreased with an increase in RBV dose (Figure 3A). When patients were stratified into two groups according to age, the relapse rate decreased with an increase in RBV dose in patients <60 years. The relapse rate was lowest (11%) in patients <60 years who received ≥ 4.0 g/kg of body weight of RBV. By contrast, among patients ≥ 60 years, the relapse rate decreased with an increase in RBV dose up to 3.0 g/kg of body weight, but remained relatively stable despite a further increase in the RBV dose beyond 3.0 g/kg of body weight. The rate of relapse was 31% to 33% in patients who received ≥ 3.0 g/kg of body weight.

Patients ≥ 60 years had higher relapse rate compared with patients <60 years after stratification by RBV dose ($P=0.044$ for RBV <2.5 g/kg, $P=0.009$ for RBV 2.5–2.9 g/kg, $P=0.150$ for RBV 3.0–3.4 g/kg, $P=0.036$ for RBV 3.5–3.9 g/kg and $P=0.006$ for RBV ≥ 4.0 g/kg).

To exclude the effect of the duration of therapy, patients who received 42–54 weeks of therapy were selected ($n=544$). Again, the relapse rate decreased with an increase in RBV dose in patients <60 years but remained stable despite a further increase in the RBV dose beyond 3.0 g/kg of body weight in patients ≥ 60 years (Figure 3B); in addition, patients ≥ 60 years had a higher relapse rate compared with younger patients after stratification by

Figure 2. Internal and external validation of the decision-tree model: subgroup-stratified comparison of the rate of relapse between the model derivation and validation groups



Each patient in the internal and external validation population was allocated to groups 1 to 5 following the flowchart of the decision tree. The rates of relapse were then calculated for each group and a graph was plotted. The rate of relapse in the (A) internal and (B) external validation groups are shown. The rates of relapse are shown as percentages below data points: the value on the left is from the model derivation group and on the right is from the validation group. The rates of relapse in each group of patients correlated closely between the model derivation group and the validation group (correlation coefficient: $r^2=0.79$ and 0.83 , respectively).

Table 2. Multivariable analysis of factors associated with relapse among patients with RVR/cEVR

Factor	OR	95% CI	P-value
No-RVR	4.07	2.57–6.43	<0.0001
Total RBV dose <3.0 g/kg body weight	2.19	1.58–3.03	<0.0001
Creatinine <0.7 g/dl	1.67	1.22–2.29	0.001
Age ≥ 60 years	2.37	1.73–3.24	<0.0001

cEVR, complete early virological response (HCV-RNA-positive at week 4, but negative at week 12); RBV, ribavirin; RVR, rapid virological response (HCV-RNA-negative at week 4).

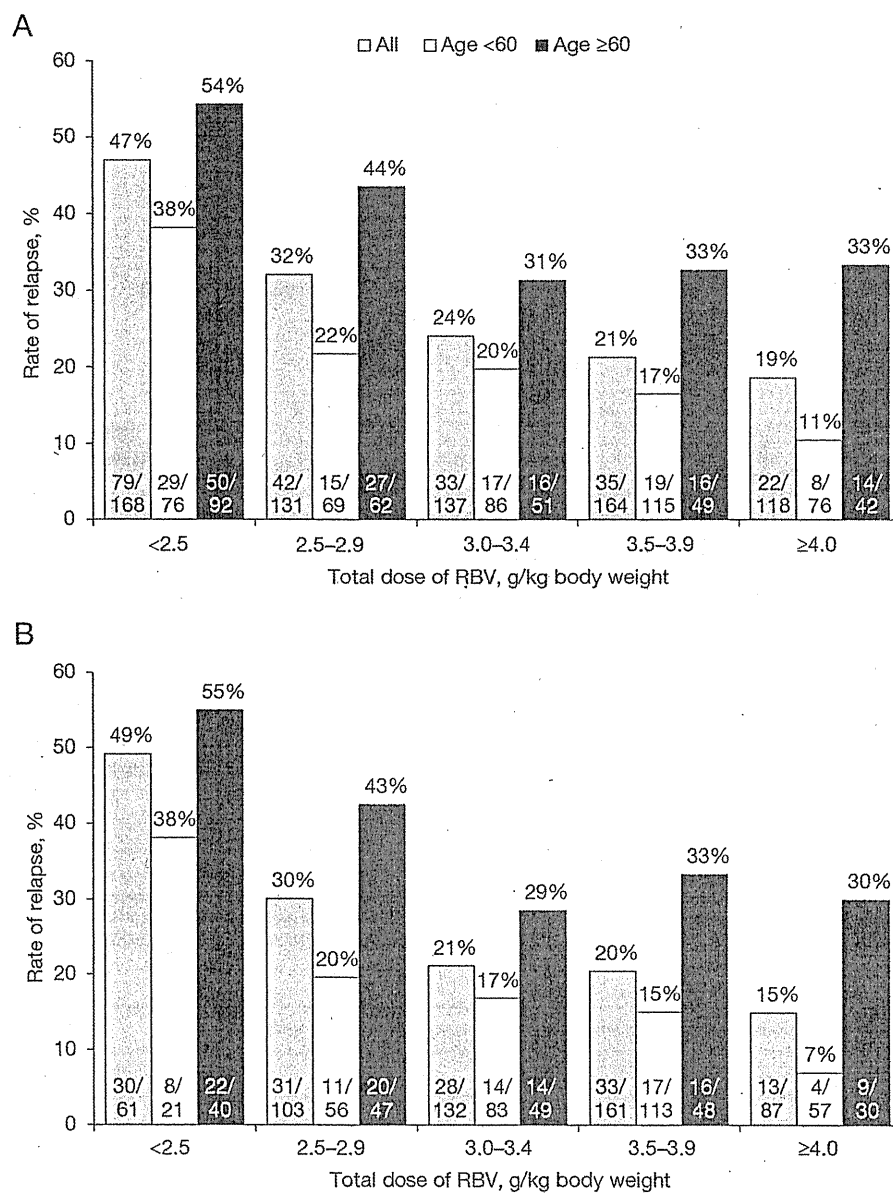
RBV dose ($P=0.283$ for RBV <2.5 g/kg, $P=0.017$ for RBV 2.5–2.9 g/kg, $P=0.127$ for RBV 3.0–3.4 g/kg, $P=0.011$ for RBV 3.5–3.9 g/kg and $P=0.009$ for RBV ≥ 4.0 g/kg).

Total dose of RBV was associated with relapse independently of PEG-IFN dose. The cutoff value of 58 $\mu\text{g}/\text{kg}$ of PEG-IFN was selected, which corresponds to the 80% of 1.5 $\mu\text{g}/\text{kg}$ dose for 48 weeks. In patients who received <58 $\mu\text{g}/\text{kg}$ of body weight of PEG-IFN,

the rate of relapse for patients who received ≥ 3.0 g/kg or <3.0 g/kg of body weight of RBV was 24% and 42%, respectively. In patients who received ≥ 58 $\mu\text{g}/\text{kg}$ of body weight of PEG-IFN, the rate of relapse for patients who received ≥ 3.0 g/kg or <3.0 g/kg of body weight of RBV was 21% and 38%, respectively.

The data mining analysis procedure did not select further split variables among RVR patients. However,

Figure 3. Correlation between the rate of relapse and total RBV dose among patients with cEVR after stratification by age



Association between the total ribavirin (RBV) dose and the rate of relapse among patients with complete early virological response (cEVR) is shown. (A) Higher dose of RBV was associated with reduced rate of relapse. (B) These associations were also confirmed in selected patients who received 42–54 weeks of therapy.

when analysed separately, the rate of relapse was also associated with age and total RBV dose among patients with RVR. The rate of relapse for patients who received ≥ 3.0 g/kg or < 3.0 g/kg of body weight of RBV was 5% and 14%, respectively. The rate of relapse for patients < 60 and ≥ 60 years was 9% and 18%, respectively. Collectively, the rate of relapse for patients < 60 years who received ≥ 3.0 g/kg or < 3.0 g/kg of body weight of RBV was 2% and 11%, respectively, whereas the rate of relapse for patients ≥ 60 years who received ≥ 3.0 g/kg or < 3.0 g/kg of body weight of RBV was 12% and 20%, respectively.

Discussion

The result of the present study shows that older age and insufficient dose of RBV are significant and independent risk factors for relapse among patients with cEVR to PEG-IFN plus RBV. Older patients (≥ 60 years) who received a total RBV dose < 3.0 g/kg of body weight had the highest risk of relapse (52%), whereas younger patients who received a total RBV dose ≥ 3.0 g/kg of body weight had the lowest risk of relapse (16%). The rate of relapse decreased depending on the total RBV dose in younger patients, but remained stable in older patients despite a further increase in the RBV dose beyond 3.0 g/kg of body weight. These findings imply that the target dose of total RBV can be set at 3.0 g/kg of body weight in patients who achieved cEVR, and further increase in RBV dose up to 4.0 g/kg of body weight or greater may be recommended in patients < 60 years.

The associations between the drug adherence and virological response had been reported with inconsistent results. In an earlier study, patients who received $> 80\%$ of the planned dose of PEG-IFN plus RBV for $> 80\%$ of the planned duration of therapy had a higher rate of SVR compared to those who received a lesser dose (51% versus 34%) [31]. Consistent results were obtained in a study reporting that patients who received $> 80\%$ of the planned dose of PEG-IFN and RBV within the first 12 weeks of therapy had a higher rate of EVR compared with those who received a lesser dose of both drugs (80% versus 33%) [4]. By contrast, a large-scale multicentre study showed that reducing the PEG-IFN dose during the first 20 weeks reduced SVR; however, reducing RBV did not affect SVR as long as RBV was not prematurely discontinued [32]. The reason for these inconsistencies is unclear. One reason may be the differences in the backgrounds of patients enrolled in the study, and hence the last study was limited to patients with advanced fibrosis and prior non-responders to PEG-IFN therapy. Because the probability of SVR is affected by virological response and relapse after response, the effect of drug dosing should be analysed separately with respect to these two factors.

In the present study, we focused on factors predictive of relapse after early virological response. According to the decision tree model, relapse was less likely in patients with RVR compared with cEVR. Among patients with cEVR, older patients (≥ 60 years) had a higher risk of relapse compared to younger patients (41% versus 22%). In addition, our results emphasized the effect of RBV dose for the prevention of relapse. In our study, a total RBV dose of ≥ 3.0 g/kg of body weight was repeatedly associated with a suppressed rate of relapse in the model derivation and validation groups. The rate of relapse in patients < 60 years who received an RBV dose of < 3.0 versus ≥ 3.0 g/kg of body weight in the model derivation, internal validation and external validation groups were 32% versus 16%, 27% versus 16%, and 41% versus 16%, respectively. The rate of relapse in patients ≥ 60 years who received an RBV dose of < 3.0 versus ≥ 3.0 g/kg of body weight in the model derivation, internal validation and external validation groups were 52% versus 26%, 45% versus 38%, and 44% versus 22%, respectively. It has been reported that the rate of relapse is suppressed in 48 weeks of IFN plus RBV combination therapy compared to IFN monotherapy, indicating that RBV contributes to the increase in SVR by reducing relapse [2,3]. Another study, focused on the associations between the drug dose reduction and relapse in patients with virological response, found that maintaining RBV dose ≥ 12 mg/kg/day during 48 weeks of treatment, which can be translated into a total dose of 4.0 g/kg of body weight, suppressed relapse [33]. Results of the present study are in accordance with this report.

The importance of drug dosing on reduction in relapse is also supported by the findings that extending therapy from 48 to 72 weeks in patients with delayed virological response improved SVR rates by reducing relapse [9–13]. Apart from these clinical studies, in the real world of clinical practice, duration of therapy is extended – even in patients with cEVR – at the physician's discretion. The relationship between duration of therapy or RBV dose, and relapse among patients with cEVR and treated with various lengths of therapy has not been examined. In the combined group of our study, extending the duration of therapy was not associated with a reduction in relapse rate. Rather, the rate of relapse decreased depending on the total RBV dose. These findings suggest that acquiring a sufficient total RBV dose, either within 48 weeks or by extending the duration of therapy, is essential to prevent relapse among patients with cEVR. The limitation of the present study was that the mean duration of therapy was only 56.3 weeks in patients whose duration of therapy was extended beyond 48 weeks. It is probable that extended duration of therapy was not long enough for the prevention of relapse. Further studies with

longer durations of therapy are necessary to confirm the effect of extended duration of therapy on reduction of relapse among patients with cEVR.

Previous reports did not consider the effects of age in setting the optimal dose of RBV. In the present study, the relapse rate decreased with an increase in RBV dose from <2.5 to 3.0–3.5 g/kg of body weight, but remained relatively stable despite a further increase in the RBV dose in older patients. Thus, a total RBV dose ≥ 3.0 g/kg of body weight should be the target dose for patients ≥ 60 years with cEVR. By contrast, ≥ 3.0 g/kg of body weight of RBV was associated with lower risk of relapse in patients <60 with cEVR (16% versus 32%), and a further increase in RBV dose led to a more profound reduction in relapse rates, as low as 11% in patients who received ≥ 4.0 g/kg of body weight. Thus, a total dose of ≥ 4.0 g/kg of body weight or even greater should be the target dose in patients <60 years.

In the near future, more potent therapies, such as direct antiviral agents [34,35], may become available. These drugs require RBV and PEG-IFN in combination. However, not all patients may be able to tolerate this triple combination therapy due to adverse drug reactions, such as severe anaemia or skin eruption. In particular, it may be difficult to administer a full dose of triple drugs to older patients. Thus, personalizing the PEG-IFN and RBV combination therapy based on this model may be beneficial to patients who were intolerant to triple combination therapy.

In the present study creatinine was an independent predictor of relapse by multivariable logistic regression analysis. However creatinine was not selected as a splitting variable in decision tree, which may be due to the unique property of data mining analysis. In data mining analysis, limitation is imposed to stop the analysis when the number of patients is <20. This limitation is used to avoid dividing patients into too small subgroups which lead to the generation of rules that only apply to the model derivation population and not reproduced when applied to other populations. This phenomenon is called the over-fitting of the model. Due to this limitation, the variables selected in the data mining analysis are not necessarily identical to the variables that are significant by ordinary multivariable analysis. In a separate analysis, lower level of creatinine was associated with higher rate of relapse in each subgroup of patients with cEVR. The reason for this association is not clear, but lower creatinine level may be related to more efficient clearance of RBV leading to lower serum level of RBV. Further research is needed to confirm this speculation.

A potential limitation of the present study is that data mining analysis has an intrinsic risk of showing relationships that fit to the original dataset, but

are not reproducible in different groups. Although internal and external validations showed that our model had high reproducibility, we recognized that further validation on a larger external validation cohort, especially in groups other than Japanese, may be necessary to further verify the reliability of our model.

In conclusion, we built a decision tree model for the prediction of relapse among patients with EVR to PEG-IFN plus RBV. The result of the present study shows that older age and insufficient dose of RBV are significant and independent risk factors for relapse. The target dose of total RBV can be set at 3.0 g/kg of body weight in patients who achieved cEVR. A further increase in RBV dose up to 4.0 g/kg of body weight may be warranted in patients <60 years.

Acknowledgements

This study was supported by a Grant-in-Aid from the Ministry of Health, Labor and Welfare, Japan (H20-kañen-006).

Disclosure statement

The authors declare no competing interests.

References

1. Strader DB, Wright T, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C. *Hepatology* 2004; 39:1147–1171.
2. Fried MW, Shiffman ML, Reddy KR, *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* 2002; 347:975–982.
3. Manns MP, McHutchison JG, Gordon SC, *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; 358:958–965.
4. 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; 38:645–652.
5. Lee SS, Ferenci P. Optimizing outcomes in patients with hepatitis C virus genotype 1 or 4. *Antivir Ther* 2008; 13 Suppl 1:9–16.
6. Namiki I, Nishiguchi S, Hino K, *et al.* Management of hepatitis C: report of the consensus meeting at the 45th annual meeting of the Japan Society of Hepatology (2009). *Hepatol Res* 2010; 40:347–368.
7. Jensen DM, Morgan TR, Marcellin P, *et al.* Early identification of HCV genotype 1 patients responding to 24 weeks peginterferon alpha-2a (40 kd)/ribavirin therapy. *Hepatology* 2006; 43:954–960.
8. Yu ML, Dai CY, Huang JF, *et al.* Rapid virological response and treatment duration for chronic hepatitis C genotype 1 patients: a randomized trial. *Hepatology* 2008; 47:1884–1893.
9. Berg T, von Wagner M, Nasser S, *et al.* Extended treatment duration for hepatitis C virus type 1: comparing 48 versus 72 weeks of peginterferon-alfa-2a plus ribavirin. *Gastroenterology* 2006; 130:1086–1097.

10. Sánchez-Tapias JM, Diago M, Escartin P, *et al*. Peginterferon-alfa2a plus ribavirin for 48 versus 72 weeks in patients with detectable hepatitis C virus RNA at week 4 of treatment. *Gastroenterology* 2006; 131:451–460.
11. Ferenci P, Laferl H, Scherzer TM, *et al*. Peginterferon alfa-2a/ribavirin for 48 or 72 weeks in hepatitis C genotypes 1 and 4 patients with slow virologic response. *Gastroenterology* 2010; 138:503–512.e1.
12. Buri M, Lurie Y, Zakharova NG, *et al*. Randomized trial of peginterferon alfa-2b and ribavirin for 48 or 72 weeks in patients with hepatitis C virus genotype 1 and slow virologic response. *Hepatology* 2010; 52:1201–1207.
13. Pearlman BL, Ehleben C, Saifee S. Treatment extension to 72 weeks of peginterferon and ribavirin in hepatitis C genotype 1-infected slow responders. *Hepatology* 2007; 46:1688–1694.
14. Tanaka Y, Nishida N, Sugiyama M, *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.
15. Suppiah V, Moldovan M, Ahlenstiel G, *et al*. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 2009; 41:1100–1104.
16. Ge D, Fellay J, Thompson AJ, *et al*. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009; 461:399–401.
17. Kurosaki M, Tanaka Y, Nishida N, *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; 54:439–448.
18. Breiman L, Friedman RA, Olshen CJ, Stone CM. *Classification and regression trees*. 1980. Belmont, CA: Wadsworth.
19. Garzotto M, Park Y, Mongoue-Tchokote S, *et al*. Recursive partitioning for risk stratification in men undergoing repeat prostate biopsies. *Cancer* 2005; 104:1911–1917.
20. 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; 12:243–248.
21. 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; 132:589–604.
22. 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; 22:3660–3667.
23. Valera VA, Walter BA, Yokoyama N, *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.
24. 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; 11:5440–5443.
25. Baquerizo A, Anselmo D, Shackleton C, *et al*. Phosphorus as an early predictive factor in patients with acute liver failure. *Transplantation* 2003; 75:2007–2014.
26. Kurosaki M, Matsunaga K, Hirayama I, *et al*. A predictive model of response to peginterferon ribavirin in chronic hepatitis C using classification and regression tree analysis. *Hepatol Res* 2010; 40:251–260.
27. Kurosaki M, Sakamoto N, Iwasaki M, *et al*. Pretreatment prediction of response to peginterferon plus ribavirin therapy in genotype 1 chronic hepatitis C using data mining analysis. *J Gastroenterol* 2011; 46:401–409.
28. Kurosaki M, Sakamoto N, Iwasaki M, *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; 83:445–452.
29. LeBlanc M, Crowley J. A review of tree-based prognostic models. *Cancer Treat Res* 1995; 75:113–124.
30. Segal MR, Bloch DA. A comparison of estimated proportional hazards models and regression trees. *Stat Med* 1989; 8:539–550.
31. McHutchison JG, Manns M, Patel K, *et al*. Adherence to combination therapy enhances sustained response in genotype-1-infected patients with chronic hepatitis C. *Gastroenterology* 2002; 123:1061–1069.
32. Shiffman ML, Ghany MG, Morgan TR, *et al*. Impact of reducing peginterferon alfa-2a and ribavirin dose during retreatment in patients with chronic hepatitis C. *Gastroenterology* 2007; 132:103–112.
33. Hiramatsu N, Oze T, Yakushijin T, *et al*. Ribavirin dose reduction raises relapse rate dose-dependently in genotype 1 patients with hepatitis C responding to pegylated interferon alpha-2b plus ribavirin. *J Viral Hepat* 2009; 16:586–594.
34. Hézode C, Forestier N, Dusheiko G, *et al*. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N Engl J Med* 2009; 360:1839–1850.
35. McHutchison JG, Everson GT, Gordon SC, *et al*. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. *N Engl J Med* 2009; 360:1827–1838.

Accepted 18 April 2011; published online 21 October 2011

Data mining model using simple and readily available factors could identify patients at high risk for hepatocellular carcinoma in chronic hepatitis C

Masayuki Kurosaki¹, Naoki Hiramatsu², Minoru Sakamoto³, Yoshiyuki Suzuki⁴, Manabu Iwasaki⁵, Akihiro Tamori⁶, Kentaro Matsuura⁷, Sei Kakinuma⁸, Fuminaka Sugauchi⁹, Naoya Sakamoto⁸, Mina Nakagawa⁸, Namiki Izumi^{1,*}

¹Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan; ²Department of Gastroenterology and Hepatology, Osaka University Graduate School of Medicine, Osaka, Japan; ³First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan; ⁴Department of Hepatology, Toranomon Hospital, Tokyo, Japan; ⁵Department of Computer and Information Science, Seikei University, Tokyo, Japan; ⁶Department of Hepatology, Osaka City University Medical School, Osaka, Japan; ⁷Department of Gastroenterology and Metabolism, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; ⁸Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Tokyo, Japan; ⁹Department of Gastroenterology, Nagoya Koseiin Medical Welfare Center, Nagoya, Japan

Background & Aims: Assessment of the risk of hepatocellular carcinoma (HCC) development is essential for formulating personalized surveillance or antiviral treatment plan for chronic hepatitis C. We aimed to build a simple model for the identification of patients at high risk of developing HCC.

Methods: Chronic hepatitis C patients followed for at least 5 years ($n = 1003$) were analyzed by data mining to build a predictive model for HCC development. The model was externally validated using a cohort of 1072 patients (472 with sustained virological response (SVR) and 600 with nonSVR to PEG-interferon plus ribavirin therapy).

Results: On the basis of factors such as age, platelet, albumin, and aspartate aminotransferase, the HCC risk prediction model identified subgroups with high-, intermediate-, and low-risk of HCC with a 5-year HCC development rate of 20.9%, 6.3–7.3%, and 0–1.5%, respectively. The reproducibility of the model was confirmed through external validation ($r^2 = 0.981$). The 10-year HCC development rate was also significantly higher in the high- and intermediate-risk group than in the low-risk group (24.5% vs. 4.8%; $p < 0.0001$). In the high- and intermediate-risk group, the incidence of HCC development was significantly reduced in patients with SVR compared to those with nonSVR (5-year rate, 9.5% vs. 4.5%; $p = 0.040$).

Conclusions: The HCC risk prediction model uses simple and readily available factors and identifies patients at a high risk of HCC development. The model allows physicians to identify patients requiring HCC surveillance and those who benefit from IFN therapy to prevent HCC.

Keywords: Decision tree; Prediction; Pegylated interferon; Ribavirin; Risk.
Received 27 May 2011; received in revised form 8 August 2011; accepted 4 September 2011

* 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).



Journal of Hepatology 2011 vol. xxx | xxx–xxx

© 2011 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved.

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common cancer worldwide [1] and its incidence is increasing in many countries [2]. Chronic viral hepatitis is responsible for 80% of all HCC cases [2]. The need to conduct HCC surveillance should be determined according to the risk of HCC development because this surveillance is cost-effective only in populations with an annualized cancer development rate of $\geq 1.5\%$ [3]. The annualized rate of developing HCC from type C liver cirrhosis is 2–8% [4–6], indicating that this population with type C liver cirrhosis needs surveillance. However, the annualized rate of HCC development is $< 1.5\%$ in patients with chronic hepatitis C but without cirrhosis and the benefit of surveillance for all patients with chronic hepatitis has not yet been established [3]. HCC surveillance may be needed for patients with advanced fibrosis because the risk of HCC development increases in parallel with the progression of liver fibrosis [7,8]. Liver biopsy is the most accurate means of diagnosing fibrosis, but a single liver biopsy cannot indicate long-term prognosis because liver fibrosis progresses over time. Serial liver biopsies are not feasible because of the procedure's invasiveness. Moreover, factors other than fibrosis, such as advanced age, obesity, sex, lower albumin, and low platelet counts, also contribute to the development of HCC from chronic hepatitis C [8–11]. Therefore, these factors must be considered while assessing the risk of HCC development.

A meta-analysis of controlled trials [12] has shown that interferon (IFN) therapy reduced the rate of HCC development in patients with type C liver cirrhosis. However, there was a marked heterogeneity in the magnitude of the prevention effect

Please cite this article in press as: Kurosaki M et al. Data mining model using simple and readily available factors could identify patients at high risk for hepatocellular carcinoma in chronic hepatitis C. J Hepatol (2011), doi:10.1016/j.jhep.2011.09.011

Research Article

of IFN on HCC development among the studies, probably due to the large differences in the baseline rate of HCC development among the different trials [12]. Whether the incidence of HCC development could be reduced in all patients with chronic hepatitis C, especially in those without liver cirrhosis, remains to be elucidated.

Data mining analysis, unlike conventional statistical analysis, is performed in an exploratory manner without considering a predefined hypothesis. Decision tree analysis, the major component of data mining analysis, is used to extract relevant factors from among various factors. These relevant factors are then combined in an orderly sequence to identify rules for predicting the incidence of the target outcome [13]. Data mining analysis has been used to define prognostic factors in various diseases [14–20]. In the field of hepatic diseases, data mining analysis has proven to be a useful tool for predicting early response [21], sustained virological response (SVR) [22–25], relapse [26], and adverse events [27] in patients with chronic hepatitis C treated with pegylated interferon (PEG-IFN) plus ribavirin (RBV). The findings of data mining analysis are expressed as flowcharts and are therefore easily understood [28] and readily available for clinical use, even by physicians without a detailed understanding of statistics.

In the present study, data mining analysis was used to identify risk factors for HCC development in a cohort of patients with chronic hepatitis C who had been followed for at least 5 years. An HCC risk prediction model was constructed on the basis of simple and generally available tests because the goal was to make the model easy to use in the clinic. The suitability, reproducibility, and generalizability of the results were validated using the data of an external cohort that was independent of the model derivation cohort.

Materials and methods

Patients

The model derivation cohort consisted of 1003 chronic hepatitis C patients without cirrhosis who had a non-sustained virological response (nonSVR) to previous IFN administered at the Musashino Red Cross Hospital and were followed for at least 5 years. Patients who had SVR or those who were followed for less than 5 years were not included. An analytical database on age, body mass index, albumin, aspartate aminotransferase (AST) levels, alanine aminotransferase (ALT) levels, γ -glutamyltransferase (GGT) levels, total bilirubin levels, total cholesterol levels, hemoglobin levels, and platelet count at the start of the observation was created. Histological data such as fibrosis stage, activity grade, or degree of steatosis was not included in the database because the goal of the present study was to make the model on the basis of simple and generally available tests. The patients who developed HCC more than 5 years after the start of the observation were considered not to have developed HCC by the 5-year point because the model was intended to predict HCC development within 5 years. The 1072 chronic hepatitis C patients included in the external validation cohort were treated with PEG-IFN and RBV at the University of Yamaguchi, Tokyo Medical and Dental University, Osaka University, Osaka City University, Nagoya City University, or Toranomon Hospital and followed for at least 5 years. Among them, 600 had nonSVR and 472 had SVR. Data from nonSVR patients in this external cohort were used for external validation of the HCC prediction model. To assess the preventive effect of PEG-IFN plus RBV therapy on HCC development, the cumulative HCC development rate was compared between SVR and nonSVR patients in the external validation cohort after stratification by the risk of HCC development as determined by data mining analysis. Informed 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 committees of all concerned hospitals.

HCC surveillance and diagnosis

HCC surveillance was conducted by performing abdominal ultrasonography every 4–6 months. Contrast-enhanced computer tomography, magnetic resonance imaging, or angiography were performed when abdominal ultrasonography suggested a new lesion suspicious for HCC. Classical HCC was diagnosed for tumors showing vascular enhancement with washout on at least two types of diagnostic imaging. Tumor biopsy was used to diagnose tumors with non-classical imaging findings.

Statistical analysis

The IBM-SPSS Modeler 13 (IBM SPSS Inc., Chicago, IL, USA) was used for decision tree analysis. The statistical methods used have been described previously [21,22,24–27]. In brief, the software searched the analytical database for the factor that most effectively predicted HCC development and for its cutoff value. The patients were divided into two groups according to that predictor. Each divided group was repeatedly assessed and divided according to this 2-choice branching method. Branching was stopped when the number of patients decreased to ≤ 20 to avoid over fitting. Finally, an HCC risk prediction model was created through this analysis. The model classified patients into subgroups with different HCC development rates in a flowchart form. For model validation, nonSVR patients from an external cohort were individually fitted into the model and classified into the subgroups and the HCC development rates of those subgroups were then calculated. The suitability and reproducibility of the model were validated by comparing the subgroup HCC development rates of the model derivation group to those of the validation group.

On univariate analysis, Student's *t*-test was used for continuous variables and Fisher's exact test was used for categorical data. Logistic regression was used for multivariate analysis. A log-rank test for Kaplan–Meier analysis was used to statistically test HCC development rates over time. *p*-Values of < 0.05 were considered significant. SPSS Statistics 18 (IBM SPSS Inc.) was used for these analyses.

Results

Univariate and multivariate analysis of factors associated with HCC development

The baseline characteristics of patients are shown in Table 1. The 5-year HCC development rate in the model derivation group was 6.2%, which did not differ significantly from the rate of 6.0% in the nonSVR group of the external cohort, but the rate of 2.0% in the SVR group of the external cohort was significantly lower than that in the model derivation group ($p = 0.0003$) and the nonSVR group of the external cohort ($p = 0.0012$). On univariate analysis, the factors found to be associated with HCC development in the model derivation cohort were age, AST levels, albumin levels, total cholesterol levels, and platelet count. On multivariate analysis, age (odds ratio 1.086), albumin levels (odds ratio 0.248), and platelet count (odds ratio 0.842) were significant predictors of HCC development (Table 2).

HCC risk prediction model by data mining analysis

The results of decision tree analysis are presented in Fig. 1. Age was selected as the first predictor. The 5-year HCC development rate was 3.4% in younger patients (< 60 years) and 8.6% in older patients (≥ 60 years). The second predictor for younger patients (< 60 years) was platelet count. The HCC development rate was 6.9% in patients with a lower platelet count ($< 150 \times 10^9/L$) and 0.8% in patients with a higher count ($\geq 150 \times 10^9/L$). The second predictor for older patients (≥ 60 years) was also platelet count. The HCC development rate was 13.1% in patients with a lower platelet count ($< 150 \times 10^9/L$) and 1.8% in patients with a higher count ($\geq 150 \times 10^9/L$). The third predictor was albumin levels,

Table 1. Baseline characteristics of patients for model derivation and external validation.

	Model derivation (n = 1003)	External cohort, non-SVR (n = 600)	External cohort, SVR (n = 472)
Sex: Male/Female*	463 (46%)/540 (54%)	306 (51%)/294 (49%)	299 (63%)/173 (37%)
Age (yr)	57.3 (11.1)	55.9 (9.6)	51.4 (10.6)
Body mass index (kg/m ²)	23.5 (3.2)	23.4 (3.3)	23.3 (3.1)
Albumin (g/dl)	4.1 (0.3)	4.0 (0.4)	4.0 (0.3)
AST (IU/L)	64.2 (36.5)	67.3 (43.8)	62.5 (48.3)
ALT (IU/L)	80.6 (55.1)	81.2 (62.3)	88.6 (82.1)
GGT (IU/L)	59.3 (50.5)	67.6 (65.1)	55.7 (71.2)
Total cholesterol (mg/dl)	172.1 (31.5)	168.2 (31.0)	174.3 (33.7)
Platelet (10 ⁹ /L)	154.0 (53.0)	153.7 (53.2)	176.6 (49.7)
Hemoglobin (g/dl)	13.3 (1.5)	14.2 (1.5)	14.4 (1.4)
HCC development within 5 years: n (%)*	62 (6.2%)	36 (6.0%)	10 (2.0%)

Data expressed as mean (standard deviation) unless otherwise indicated.

AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase; HCC, hepatocellular carcinoma; SVR, sustained virological response.

*Data expressed as number of patients (percentage).

whose cutoff value was 3.75 g/dl in patients with a higher platelet count ($\geq 150 \times 10^9/L$). The HCC development rate was 6.3% when albumin levels were lower (< 3.75 g/dl) and 1.5% when levels were higher (≥ 3.75 g/dl). The cutoff value for albumin levels was 4.0 g/dl in patients with a lower platelet count ($< 150 \times 10^9/L$). The HCC development rate was 20.9% when albumin levels were lower (< 4.0 g/dl) and 6.4% when levels were higher (≥ 4.0 g/dl). The fourth and final predictor was AST levels. The HCC development rate was 7.3% when AST levels were at least 40 IU/L and 0% when the levels were < 40 IU/L. On the basis of this analysis, seven subgroups with a 5-year HCC development rate of 0–20.9% were identified. The area under the receiver operating characteristic curve according to the HCC risk prediction model was 0.817.

External validation of the HCC risk prediction model with an independent external cohort

Six hundred nonSVR patients from an external cohort were fitted into the HCC risk prediction model and classified into the seven subgroups. The 5-year HCC development rate of these subgroups was 0–17.9%. The HCC development rate in the individual subgroups of the model derivation group was closely correlated to that in the corresponding subgroups of the external validation group (Fig. 2; correlation coefficient $r^2 = 0.981$). The HCC development rate in the subgroup of patients with the highest risk of HCC development (high-risk group) according to the model older age (≥ 60 years) with a lower platelet count ($< 150 \times 10^9/L$) and lower albumin levels (< 4.0 g/dl) was 20.9% in the model derivation

group and 17.9% in the external validation group. The intermediate-risk group or the patients with an HCC development rate of at least 5% consisted of the following three subgroups: (1) older age (≥ 60 years), lower platelet count ($< 150 \times 10^9/L$), higher albumin levels (≥ 4.0 g/dl), and higher AST levels (≥ 40 IU/L); (2) older age (≥ 60 years), higher platelet count ($\geq 150 \times 10^9/L$), and lower albumin levels (< 3.75 g/dl); and (3) younger age (< 60 years) and lower platelet count ($< 150 \times 10^9/L$). In these intermediate-risk groups, the 5-year HCC development rate was 6.3–7.3% in the model derivation group and 5.3–7.9% in the external validation group. The low-risk group consisted of the following three subgroups: (1) younger age (< 60 years) and higher platelet count ($\geq 150 \times 10^9/L$); (2) older age (≥ 60 years), lower platelet count ($< 150 \times 10^9/L$), higher albumin levels (≥ 4.0 g/dl), and lower AST levels (< 40 IU/L); and (3) older age (≥ 60 years), higher platelet count ($\geq 150 \times 10^9/L$), and higher albumin levels (≥ 3.75 g/dl). In these low-risk groups, the 5-year HCC development rate was 0–1.5% in the model derivation group and 0–2.9% in the external validation group.

Predictability of the HCC risk prediction model on HCC development rate beyond 5 years

Cumulative HCC development rates in the high-, intermediate-, and low-risk groups were compared over time using the Kaplan-Meier method. The 10-year rates were 28.9% in the high-risk group, 22.9% in the intermediate-risk group, and 4.8% in the low-risk group (Fig. 3A). The high and intermediate-risk group created by pooling data from the high- and intermediate-risk groups had a significantly higher cumulative HCC development rate than the low-risk group beyond 5 years (Fig. 3B; 5-year rate, 11.6% vs. 1.0%; 10-year rate, 24.5% vs. 4.8%; $p < 0.0001$).

Effect of response to PEG-IFN plus RBV therapy in the reduction of HCC development: analysis stratified by the HCC risk prediction model

The 600 nonSVR patients and 472 SVR patients in the external cohort were fitted into the HCC risk prediction model and

Table 2. Multivariable analysis of factors associated with subsequent development of HCC within 5 years.

	Odds ratio	95% CI	p value
Age	1.086	1.029-1.146	0.003
Albumin	0.248	0.100-0.613	0.003
Platelet	0.842	0.769-0.921	< 0.0001

CI, confidence interval.

Research Article

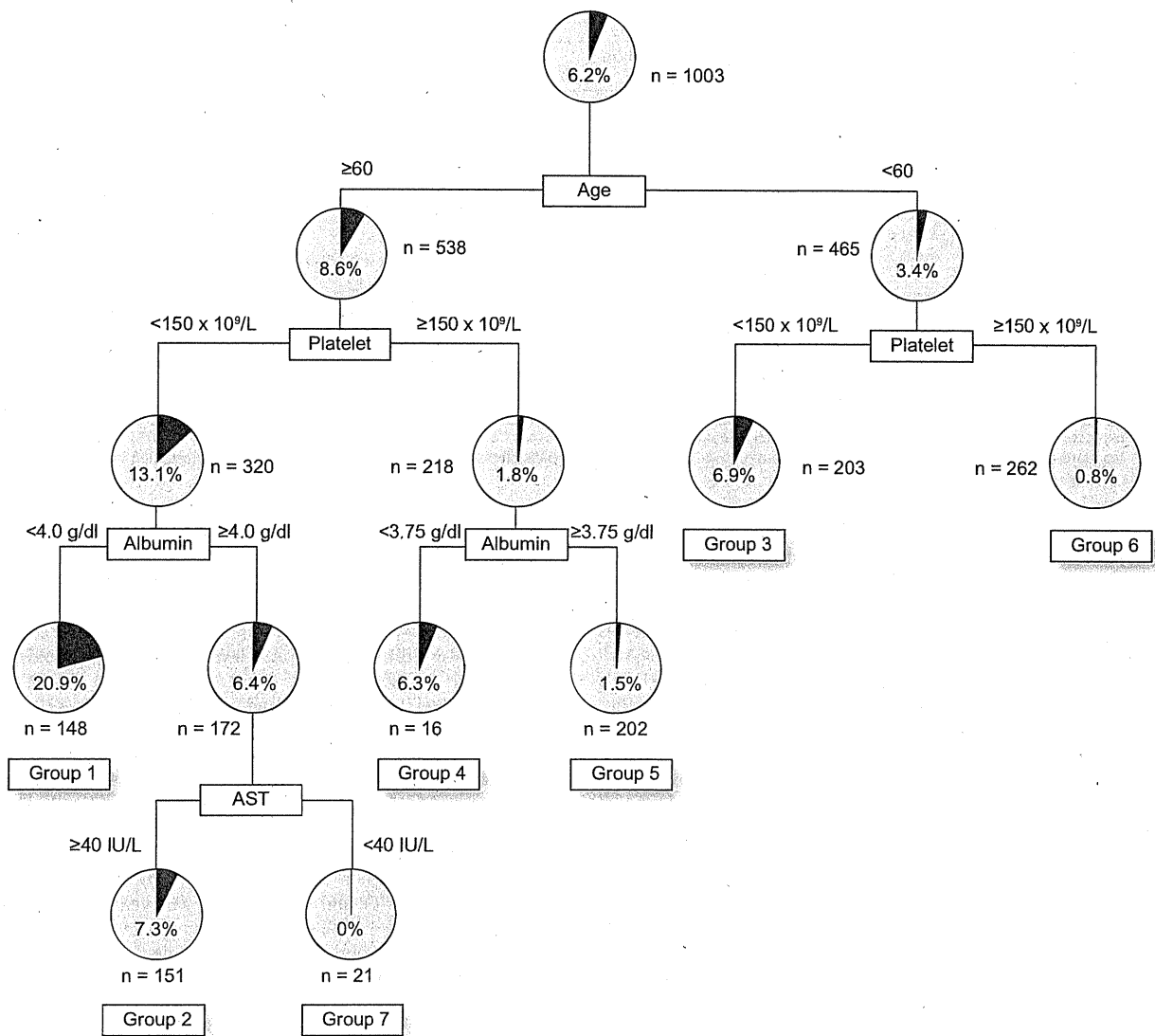


Fig. 1. The decision tree model of HCC development within 5 years. Boxes indicate the factors used to differentiate patients and the cutoff values for those different groups. Pie charts indicate the HCC development rate within 5 years for each group of patients after differentiation. Terminal groups of patients differentiated by analysis are numbered from 1 to 7.

classified into the high- and intermediate-risk group or the low-risk group, as defined above. The HCC development rate was significantly lower in SVR patients than in nonSVR patients in the high- and intermediate-risk group (5-year HCC rate, 9.5% vs. 4.5%; $p = 0.040$, log-rank test). In the low-risk group, the 5-year rate was 1.8% in nonSVR patients and 0.9% in SVR patients. Both rates were low and not significantly different ($p = 0.331$, log-rank test) (Fig. 4).

Discussion

An awareness of the risk of HCC development in the context of routine care for chronic hepatitis C is essential for formulating

an HCC surveillance plan personalized for individual patients. The risk of developing HCC from chronic hepatitis is lower than that from cirrhosis [7]; therefore, across-the-board surveillance for chronic hepatitis C is not recommended [3]. A method to easily determine this risk, without performing serial liver biopsies, would be extremely significant clinically. In the present study, an HCC risk prediction model that included the factors such as age, platelet count, albumin levels, and AST levels was constructed. The model was found to have excellent reproducibility when validated with an external cohort. This model could identify subgroups of chronic hepatitis C patients at high risk of HCC development; the 5-year HCC development rate for the high- and intermediate-risk groups was 11.6%, yielding an annual incidence of 2.3%. This HCC risk prediction model requires only

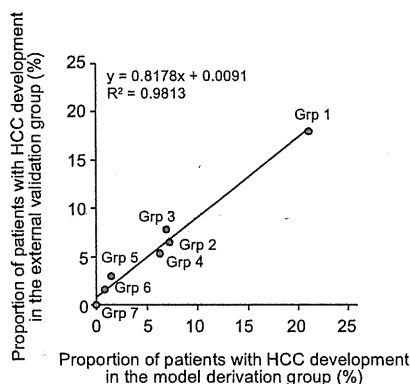


Fig. 2. External validation of the decision tree model with an independent cohort. Each patient in the external validation group was allocated to groups 1–7 following the flowchart of the decision tree. The HCC development rates were then calculated for each group and the graph plotted. The x-axis represents the HCC development rate in the model derivation group, and the y-axis represents the HCC development rate in the external validation group. The HCC development rates in each subgroup of patients are closely correlated between the model derivation group and the external validation group (correlation coefficient: $R^2 = 0.981$).

simple test values that are readily obtained in routine care and can therefore be easily used at the patient bedside. The model can be used to identify patients with a high risk of HCC development and therefore requiring surveillance, thereby allowing the formulation of surveillance plans personalized for individual patients.

Advanced fibrosis has been reported as independent risk factors for HCC development [7,8]. Platelet counts and albumin levels, which were factors selected for discrimination of the risk of HCC development, are closely related to the stage of fibrosis. Their correlation with the HCC risk has been repeatedly demonstrated [9–11,29–31]. The present study confirmed the impact of old age and advanced fibrosis, as reflected by low platelet counts and albumin levels. These results are consistent with our previous report [32]. What is unique to the present study was the study design to build a simple and reliable model for

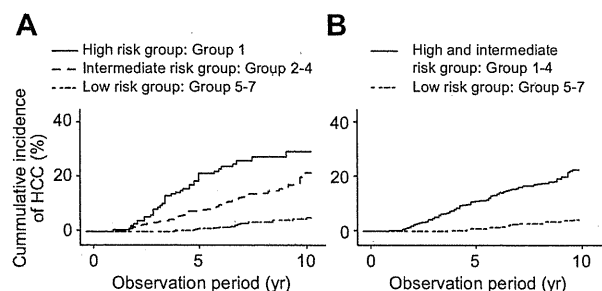


Fig. 3. Cumulative incidence of HCC development beyond 5 years in subgroups of patients defined by the decision tree model. Cumulative incidences of HCC in the groups classified by the decision tree model are compared. (A) The cumulative HCC development rate beyond 5 years is higher in the high- (group 1) and intermediate-risk (groups 2–4) groups compared to the low-risk group (groups 5–7). (B) The high and intermediate-risk groups have a significantly higher cumulative HCC development rate than the low-risk group (5-year rate, 11.6% vs. 1.0%; 10-year rate, 24.5% vs. 4.8%; $p < 0.0001$).

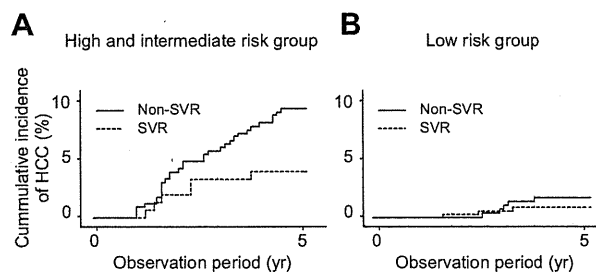


Fig. 4. Sustained virological response to PEG-IFN plus RBV therapy reduces the incidence of HCC development after stratification by the HCC risk. The 600 nonSVR patients and the 472 SVR patients in the external cohort were fitted into the HCC risk prediction model and classified into the high and intermediate-risk group or the low-risk group. The HCC development rate is significantly lower in SVR patients than in nonSVR patients in the high and intermediate-risk group (groups 1–4) (5-year HCC rate, 9.5% vs. 4.5%; $p = 0.040$). In the low-risk group (groups 5–7), the 5-year rate is 1.8% in nonSVR patients and 0.9% in SVR patients. Both rates are low and not significantly different ($p = 0.331$).

the prediction of HCC development that could be easily used in the clinic. For this purpose, a novel statistical method was used, histological factors were excluded in the analysis, the model derivation cohort was restricted to those who had nonSVR and had a long follow-up period duration (5 years), and the reproducibility of the model was independently validated by an external cohort. These are the major differences of the present study compared to our previous report. Many researchers have put a lot of efforts to formulate regression models for HCC prediction [9,10,33]. These prediction models are useful for identifying high-risk patients but are somewhat complicated to use at the bedside because they require calculations to be performed. Our prediction model is used simply by incorporating patients' data obtained through simple tests into the decision tree and following the flowchart. These prediction models based on factors easily accessible in routine clinical settings help physicians identify high-risk patients out of chronic hepatitis.

Viral eradication is the short-term goal of IFN therapy, but the ultimate goal is the prevention of HCC occurrence. Previous reports have shown that SVR to IFN therapy suppresses HCC occurrence in patients with type C liver cirrhosis and chronic hepatitis [7,12,30,34,35]. However, there is a marked heterogeneity in the magnitude of the treatment effect on the risk of HCC among studies, probably due to differences in the baseline risk of HCC among different trials [12]. Thus, the question remains whether the preventive effect of IFN therapy on HCC development could apply to all patients with chronic hepatitis C, especially those without liver cirrhosis. The result of the present study indicated that among high- and intermediate-risk patients, as assessed with our HCC risk prediction model, the cumulative HCC development rate was significantly reduced in SVR patients compared with nonSVR patients. This finding suggests that patients with chronic hepatitis, in whom disease has not yet progressed to hepatic cirrhosis but who are at a high risk of HCC development, benefit from antiviral treatment. The preventive effect of IFN on HCC development was not evident in low-risk patients within 5 years of observation. A longer observation term may be required to analyze the possible effect of antiviral therapy in these patients. Application of the present model on treatment decision may have limitations in that effect to prevent HCC development may differ in newer therapeutic agents such as protease

Research Article

inhibitors [36,37], and that low-risk patients may also benefit from therapy after a longer term observation period such as 15–20 years.

Patients with chronic hepatitis often have no subjective symptoms accompanying their disease and therefore have a low consciousness of the disease. The broad array of adverse reactions and the high cost of IFN therapy are frequent hurdles in motivating patients to undergo therapy. However, patients may be convinced to undergo therapy or remain motivated for continued therapy if they are made aware of their risk of HCC development and the preventive effect of IFN on HCC development.

In conclusion, a reproducible HCC risk prediction model, which includes the factors such as age, platelet count, albumin levels, and AST levels, was constructed to predict the 5-year HCC development rate in patients with chronic hepatitis C. The model requires only a combination of readily available test values and can therefore be easily used at the bedside. The information provided by the model allows the physician to identify patients requiring IFN therapy for the prevention of HCC and formulate plans for imaging HCC surveillance.

Conflict of interest

The authors who have taken part in this study declared 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 (H20-kanen-006).

References

- [1] Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005;55:74–108.
- [2] Bosch FX, Ribes J, Diaz M, Cleries R. Primary liver cancer: worldwide incidence and trends. *Gastroenterology* 2004;127:S5–S16.
- [3] Bruix J, Sherman M. Management of hepatocellular carcinoma. *Hepatology* 2005;42:1208–1236.
- [4] Fattovich G, Giustina G, Degos F, Tremolada F, Diodati G, Almasio P, et al. Morbidity and mortality in compensated cirrhosis type C: a retrospective follow-up study of 384 patients. *Gastroenterology* 1997;112:463–472.
- [5] Niederau C, Lange S, Heintges T, Erhardt A, Buschkamp M, Hurter D, et al. Prognosis of chronic hepatitis C: results of a large, prospective cohort study. *Hepatology* 1998;28:1687–1695.
- [6] Degos F, Christidis C, Ganne-Carrie N, Farmachidi JP, Degott C, Guettier C, et al. Hepatitis C virus related cirrhosis: time to occurrence of hepatocellular carcinoma and death. *Gut* 2000;47:131–136.
- [7] Yoshida H, Tateishi R, Arakawa Y, Sata M, Fujiyama S, Nishiguchi S, et al. Benefit of interferon therapy in hepatocellular carcinoma prevention for individual patients with chronic hepatitis C. *Gut* 2004;53:425–430.
- [8] Kurosaki M, Hosokawa T, Matsunaga K, Hirayama I, Tanaka T, Sato M, et al. Hepatic steatosis in chronic hepatitis C is a significant risk factor for developing hepatocellular carcinoma independent of age, sex, obesity, fibrosis stage and response to interferon therapy. *Hepatology* 2010;40:870–877.
- [9] Velazquez RF, Rodriguez M, Navascues CA, Linares A, Perez R, Sotorrios NG, et al. Prospective analysis of risk factors for hepatocellular carcinoma in patients with liver cirrhosis. *Hepatology* 2003;37:520–527.
- [10] Lok AS, Seeff LB, Morgan TR, di Bisceglie AM, Sterling RK, Curto TM, et al. Incidence of hepatocellular carcinoma and associated risk factors in hepatitis C-related advanced liver disease. *Gastroenterology* 2009;136:138–148.
- [11] Cardoso AC, Moucari R, Figueiredo-Mendes C, Ripault MP, Giully N, Castelnau C, et al. Impact of peginterferon and ribavirin therapy on hepatocellular carcinoma: incidence and survival in hepatitis C patients with advanced fibrosis. *J Hepatol* 2010;52:652–657.
- [12] Craxi A, Camma C. Prevention of hepatocellular carcinoma. *Clin Liver Dis* 2005;9:329–346, viii.
- [13] Breiman LJH, Friedman RA, Olshen CJ, Stone CM. Classification and regression trees. Calif: Wadsworth; 1980.
- [14] Garzotto M, Park Y, Mongoue-Tchokote S, Bledsoe J, Peters L, Blank BH, et al. Recursive partitioning for risk stratification in men undergoing repeat prostate biopsies. *Cancer* 2005;104:1911–1917.
- [15] 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;12:243–248.
- [16] 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;132:589–602.
- [17] 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;22:3660–3667.
- [18] 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.
- [19] 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;11:5440–5443.
- [20] 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;75:2007–2014.
- [21] 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. *Hepatology* 2010;40:251–260.
- [22] 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;83:445–452.
- [23] Izumi N, Asahina Y, Kurosaki M. Predictors of virological response to a combination therapy with pegylated interferon plus ribavirin including virus and host factors. *Hepat Res Treat* 2010;2010:703602.
- [24] 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;46:401–409.
- [25] 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;54:439–448.
- [26] Kurosaki M, Hiramatsu N, Sakamoto M, Suzuki Y, Iwasaki M, Tamori A, et al. Age and total ribavirin dose is an independent predictor of relapse among early virological responders to peg-interferon plus ribavirin therapy in chronic hepatitis C revealed by data mining analysis. *Antivir Ther*, in press.
- [27] Hiramatsu N, Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, et al. Pretreatment prediction of anemia progression by pegylated interferon alpha-2b plus ribavirin combination therapy in chronic hepatitis C infection: decision-tree analysis. *J Gastroenterol* 2011;46:1111–1119.
- [28] LeBlanc M, Crowley J. A review of tree-based prognostic models. *Cancer Treat Res* 1995;75:113–124.
- [29] Lu SN, Wang JH, Liu SL, Hung CH, Chen CH, Tung HD, et al. Thrombocytopenia as a surrogate for cirrhosis and a marker for the identification of patients at high-risk for hepatocellular carcinoma. *Cancer* 2006;107:2212–2222.
- [30] Bruno S, Stroffolini T, Colombo M, Bollani S, Benvegnu L, Mazzella G, et al. Sustained virological response to interferon-alpha is associated with improved outcome in HCV-related cirrhosis: a retrospective study. *Hepatology* 2007;45:579–587.
- [31] Bonis PA, Tong MJ, Blatt LM, Conrad A, Griffith JL. A predictive model for the development of hepatocellular carcinoma, liver failure, or liver transplantation for patients presenting to clinic with chronic hepatitis C. *Am J Gastroenterol* 1999;94:1605–1612.

- [32] Asahina Y, Tsuchiya K, Tamaki N, Hirayama I, Tanaka T, Sato M, et al. Effect of aging on risk for hepatocellular carcinoma in chronic hepatitis C virus infection. *Hepatology* 2010;52:518-527.
- [33] Ganne-Carrie N, Chastang C, Chapel F, Munz C, Pateron D, Sibony M, et al. Predictive score for the development of hepatocellular carcinoma and additional value of liver large cell dysplasia in Western patients with cirrhosis. *Hepatology* 1996;23:1112-1118.
- [34] Nishiguchi S, Kuroki T, Nakatani S, Morimoto H, Takeda T, Nakajima S, et al. Randomised trial of effects of interferon-alpha on incidence of hepatocellular carcinoma in chronic active hepatitis C with cirrhosis. *Lancet* 1995;346:1051-1055.
- [35] Shiratori Y, Ito Y, Yokosuka O, Imazeki F, Nakata R, Tanaka N, et al. Antiviral therapy for cirrhotic hepatitis C: association with reduced hepatocellular carcinoma development and improved survival. *Ann Intern Med* 2005;142:105-114.
- [36] Hezode C, Forestier N, Dusheiko G, Ferenci P, Pol S, Goeser T, et al. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N Engl J Med* 2009;360:1839-1850.
- [37] 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;360:1827-1838.

Characterization of naturally occurring protease inhibitor-resistance mutations in genotype 1b hepatitis C virus patients

Hiroko Shindo · Shinya Maekawa · Kazuki Komase · Ryota Sueki · Milka Miura · Makoto Kadokura · Kuniaki Shindo · Fumitake Amemiya · Takatoshi Kitamura · Yasuhiro Nakayama · Taisuke Inoue · Minoru Sakamoto · Shun-ichi Okada · Yasuhiro Asahina · Namiki Izumi · Masao Honda · Shuichi Kaneko · Nobuyuki Enomoto

Received: 21 April 2011 / Accepted: 22 July 2011
© Asian Pacific Association for the Study of the Liver 2011

Abstract

Background and aims Protease inhibitor (PI)-resistant hepatitis C virus (HCV) variants may be present in substantial numbers in PI-untreated patients according to recent reports. However, influence of these viruses in the clinical course of chronic hepatitis C has not been well characterized.

Methods The dominant HCV nonstructural 3 (NS3) amino acid sequences were determined in 261 HCV genotype 1b-infected Japanese patients before pegylated interferon plus ribavirin (PEG-IFN/RBV) therapy, and investigated the patients' clinical characteristics as well as treatment responses including sustained virological response (SVR) rate. HCV-NS3 sequences were also determined in 39 non-SVR patients after completion of the therapy.

Results Four single mutations (T54S, Q80K, I153V, and D168E) known to confer PI resistance were found in 35 of 261 patients (13.4%), and double mutations (I153V plus

T54S/D168E) were found in 6 patients (2.3%). Responses to PEG-IFN/RBV therapy did not differ between patients with and without PI-resistance mutations (mutation group, SVR 48%; wild-type group, SVR 40%; $P = 0.38$). On the other hand, two mutations appeared in two non-SVR patients after PEG-IFN/RBV therapy (I153V and E168D, 5.1%).

Conclusions PI-resistance-associated NS3 mutations exist in a substantial proportion of untreated HCV-1b-infected patients. The impact of these mutations in the treatment of PIs is unclear, but clinicians should pay attention to avoid further development of PI resistance.

Keywords HCV · Protease inhibitor · Naturally occurring viral resistance mutations

Introduction

Hepatitis C virus (HCV) infects more than 170 million persons worldwide and thus represents a global health problem. At least 130 million infected individuals are chronic carriers of HCV and are at significant risk of developing liver cirrhosis and hepatocellular carcinoma [1]. The current standard treatment with pegylated interferon plus ribavirin (PEG-IFN/RBV) is complicated by frequent adverse reactions, and a sustained virologic response (SVR) can be achieved only in 50% of patients infected with the most prevalent genotype 1 [2]. In Japan, since 70% of patients are infected with intractable genotype 1b HCV, more effective treatments are urgently required.

A promising approach is the development of specifically targeted antiviral therapies for hepatitis C (STAT-C). HCV-specific protease inhibitors (PIs) target an essential step in HCV replication by blocking the nonstructural 3/4A (NS3/4A) protease-dependent cleavage of the HCV polyprotein

Electronic supplementary material The online version of this article (doi:10.1007/s12072-011-9306-7) contains supplementary material, which is available to authorized users.

H. Shindo · S. Maekawa (✉) · K. Komase · R. Sueki · M. Miura · M. Kadokura · K. Shindo · F. Amemiya · T. Kitamura · Y. Nakayama · T. Inoue · M. Sakamoto · S. Okada · N. Enomoto

First Department of Internal Medicine, University of Yamanashi, 1110, Shimokata, Chuo, Yamanashi 409-3898, Japan
e-mail: maekawa@yamanashi.ac.jp

Y. Asahina · N. Izumi
Division of Gastroenterology and Hepatology,
Musashino Red Cross Hospital, Tokyo, Japan

M. Honda · S. Kaneko
Department of Gastroenterology, Kanazawa University Graduate
School of Medicine, Kanazawa, Japan

[1]. Among these NS3/4A PIs, telaprevir, boceprevir, SCH446211, danoprevir (ITMN-191), naldaprevir (SCH900 518), and TMC435 are now under clinical trials [1, 3–7]. In PROVE1 and PROVE2 studies [3, 4] undertaken in North America and Europe, the SVR rate was favorable (67 and 69%, respectively) in a triple therapy regimen including telaprevir. In addition, some studies have suggested that shortening of treatment duration may be possible for patients who achieve a rapid virologic response (RVR) [8, 9].

However the sole use of STAT-C drugs, such as PIs, promotes production and selection of drug-resistant variants in patients experiencing viral rebound during treatment [3, 10, 11] as well as in HCV replicon experiments [11, 12]. Therefore, these drugs should be used in combination with the PEG-IFN/RBV to prevent the appearance of drug-resistant variants. However, Kuntzen et al. [13] demonstrated the presence of these drug-resistant variants in high frequencies (8.6–16.2%) by population-based sequencing in patients not treated with the drugs [1, 13]. Gaudieri et al. [14] have suggested that regions of NS3 protease and NS5B polymerase are likely to be under HLA immune pressure and therapeutic selection, and that drug-resistant variants may occur naturally to escape the immune system. These observations seem quite astonishing and troubling, since a substantial number of patients may not respond to the new therapies such as STAT-C drugs.

In the present study, to assess the prevalence of NS3 mutations conferring PI resistance in HCV genotype 1b-infected Japanese patients who had not been previously treated with PIs, as well as to assess the influence of those mutations in response to PEG-IFN/RBV therapy, the dominant HCV-NS3 sequences were determined in 261 HCV-1b patients before starting the PEG-IFN/RBV therapy.

Methods

Patients

Serum samples were acquired from 261 HCV genotype 1b-infected adult Japanese patients before combination therapy with PEG-IFN (PEGINTRON[®], Schering-Plough, Tokyo, Japan) plus RBV (REBETOL[®], Schering-Plough) between 2004 and 2008 at the University of Yamanashi, Musashino Red Cross Hospital and Kanazawa University. The therapy was administered according to the standard PEG-IFN/RBV treatment protocol established for Japanese patients by a hepatitis study group of the Ministry of Health, Labor, and Welfare, Japan. Specifically, the patients were subcutaneously administered PEG-IFN α -2b, 1.5 μ g/kg body weight, once weekly and RBV 600–800 mg daily per os for 48 weeks. These patients were not infected with human immunodeficiency virus (HIV). The study was

approved by the ethics committees of all participating universities and the hospital, and the protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki as reflected in a priori approval by the Institutional Review Board at Massachusetts General Hospital. Written informed consent was obtained from each study participant.

Amplification and sequencing of full-length HCV genomes

Viral loads were determined using the Amplicor HCV RNA kit, version 2.0 (Roche Diagnostics, Tokyo, Japan) or the Cobas TaqMan test (Roche Diagnostics). HCV RNA was extracted from pretreatment serum samples by the AGPC method using Isogen (Wako, Osaka, Japan) according to the manufacturer's protocol. Complementary DNA was synthesised using Superscript II (Invitrogen, Tokyo, Japan) and random primers (Invitrogen), and then amplified by two-step nested PCR using the primers listed in Supplementary Table 1. All samples were initially denatured at 95°C for 7 min, followed by 40 cycles of amplification with denaturation at 95°C for 15 s, annealing at 55°C for 15 s, and extension at 72°C for 45 s using the BD Advantage[™] 2 PCR Enzyme system (BD Biosciences Clontech, CA, USA). PCR amplicons were directly sequenced using BigDye Terminator version 3.1 (ABI, Tokyo, Japan) and universal M13 forward/reverse primers using an ABI prism 3130 sequencer (ABI).

Sequence alignment and analysis

Sequences were determined in both directions, particularly for the ambiguous stretches, were assembled using the Vector NTI software (Invitrogen), and base-calling errors were corrected following the inspection of chromatograms. If mixed bases were detected as two different chromatogram peaks at the same residue, only the dominant base was called after evaluation of all overlapping fragments. A consensus sequence was generated from the alignment on the basis of the most common amino acid at each site.

Determination of PI resistance mutations

Multiple viral NS3 mutations were observed in amino acid positions reported to confer PI resistance among 261 patients: V36, Q41, F43, T54, V55, Q80, R109, I153, R155, A156, D168, V170, and M175. NS3 amino acid mutations with proven PI resistance in previously published studies (Table 1) were designated as resistance proven mutations (e.g., V36M/A). Mutations in the PI-resistance site not known to confer drug resistance were designated resistance unproven mutations (e.g., V36I). Patients were allocated to two groups according to the presence of PI-resistance

mutations (including resistance unproven mutations), and clinical characteristics including HCV RNA levels and responses to PEG-IFN/RBV therapy were compared. To assess the influence of PEG-IFN/RBV therapy on NS3 mutational status, posttreatment HCV-NS3 sequences in 39 of 58 non-SVR patients were also examined.

Statistical analysis

Statistical differences in the data, including all available patients' demographic, biochemical, hematologic, and virologic data such as sequence variation factors, were determined among the various groups by Student's *t* test or Mann-Whitney *U* test for numerical variables and Fisher's exact probability test for categorical variables.

Results

Prevalence of dominant PI-resistance-associated nonstructural 3 mutations in untreated patients

Figure 1 shows the frequency of substitutions in 261 patients for each of 181 NS3 protease amino acid residues

compared to the consensus sequence. A total of 41 resistance proven mutations were detected in 35 (13.4%) patients: T54S (14 patients, 5.4%), Q80K (1 patient, 0.4%), I153V (22 patients, 8.4%), D168E (4 patients, 1.5%), T54S plus I153V double mutation (4 patients, 1.5%), and I153V plus D168E double mutation (2 patients, 0.8%). The mutation number increased to 54 in 47 (18.0%) patients when resistance unproven mutations were included: V36I (2 patients, 0.8%), I153L (11 patients, 4.2%), and I153V plus V36I double mutation (2 patients, 1.5%). Double mutations were found in 7 patients (2.7%) (Table 1). Q80L was observed in 47 (18%) patients but these were excluded from consideration because a previous study demonstrated that this mutation does not confer resistance [15]. All mutations observed in this study would confer low- to moderate-level PI resistance according to previous studies [6, 15–19]. No mutations conferring high-level resistance such as R155 or A156 [11, 17, 19–22] were observed.

Clinical characteristics of patients with PI-resistance mutations

Table 2 presents the characteristics of patients classified according to the presence of PI-resistance mutations

Table 1 Prevalence of PI-resistance-associated NS3 mutations

Drug-resistance mutations described in the literature				References	Detected resistance mutations Genotype 1b (<i>N</i> = 261), (%)
NS3 residue	Resistance mutations	Drugs			
V36	A, M, L, G, C	Telaprevir, Boceprevir	[1, 3, 4, 10, 11, 19, 31, 37]	I × 2 (0.8)	
Q41	R	ITMN-191, Boceprevir	[19]		
F43	S, C	ITMN-191, Boceprevir, Telaprevir, TMC435	[15, 19]		
T54	A, S	Telaprevir, Boceprevir, SCH900518	[1, 3, 10, 11, 19, 20, 31, 38]	S × 14 (5.4)	
V55	A	Boceprevir	[1]		
Q80	R, K	TMC435	[6, 15]	K × 1 (0.4)	
R109	K	SCH446211	[17]		
I153	V	SCH446211	[17]	V × 22 (8.4), L × 11 (4.2)	
R155	K, T, I, M, G, L, S, Q	Telaprevir, Boceprevir, ITMN-191, BILN2061, TMC435	[1, 3, 4, 6, 10, 11, 15, 19, 20]		
A156	S, T, V, I, G	Telaprevir, Boceprevir, ITMN-191, BILN2061, SCH446211, TMC435, SCH900518	[1, 3, 4, 10, 11, 15, 17, 19, 20, 38]		
D168	A, V, E, N, T, H	BILN2061, ITMN-191, TMC435	[6, 15, 20]	E × 4 (1.5)	
V170	A	Telaprevir, Boceprevir	[1, 19, 20]		
M175	L	Boceprevir	[39]		
Total number (%) of patients with resistance proven mutations				35 (13.4)	
Total number (%) of patients with resistance proven and unproven mutations				47 (18.0)	

Amino acid mutations conferring PI resistance in the literatures and those observed in PI-treatment-naïve patients in this study are indicated. Bold indicates resistance proven mutations, and the others indicate resistance unproven mutations

Double mutations found were as follows: V36I and I153V × 1, T54S and I153V × 4, I153V and D168E × 2

(including resistance unproven mutations). Age, sex ratio, body mass index, alanine aminotransferase (ALT) levels, serum albumin, platelet count, and fibrosis stage did not differ between the NS3 mutation and wild-type groups. No significant difference was observed between the two groups in the parameters of PEG-IFN/RBV treatment response, HCV sequence variations in interferon sensitivity determining region (ISDR), Core 70, interferon plus ribavirin resistance-determining region (IRRDR), or interleukin 28B (IL28B) single nucleotide polymorphism (SNP) (rs8099917; T/G and G/G vs. T/T) [23–30]. These clinical variables were also compared between the mutation group defined as resistance proven mutations and the wild-type group, but no notable differences were observed.

Unimpaired in vivo fitness of viral strains with resistance mutations

Because most PI-resistance mutations described till date have been associated with reduced replicative capacity of varying degrees [1, 10, 11, 13, 17, 20–22, 31, 32], we examined viral replication levels in patients with drug-resistance mutations (Fig. 2). The estimated *P* value indicated no significant difference between the mutation (median 1,500 KIU/ml) and wild-type (median 1,800 KIU/ml) groups (*P* = 0.69). The results indicate that drug-resistant HCVs were not necessarily impaired in their ability to replicate in vivo. However, patients with double mutations (*N* = 7) tended to have low viral loads (median 1,200 KIU/ml) (*P* = 0.09).

Resistance mutations and virologic response to PEG-IFN/RBV therapy

To determine the difference in virologic response to PEG-IFN/RBV therapy according to the PI mutation, frequency of HCV RNA levels below detection at 4 weeks (rapid viral response, RVR) and 12 weeks (complete early viral response, cEVR), and SVR rate (%) were investigated in

each group. The frequency of HCV RNA levels below detection at 4 and 12 weeks was 14 and 50%, respectively, in the mutation group, and was 11 and 46%, respectively, in the wild-type group. The SVR rate was 48 and 40% in the mutation and wild-type groups, respectively (*P* = 0.38). No significant difference was observed between the two groups in any of the indexes investigated (Table 2). The time-dependent viral clearance rate during PEG-IFN/RBV therapy was estimated in 133 patients including 25 patients (19%) with PI-resistance mutations available for the analysis. Kaplan–Meier analysis demonstrated that HCV clearance did not differ between the two groups with and without resistance mutations (log-rank test, *P* = 0.30) (Fig. 3).

Changes in nonstructural 3 amino acid sequence diversity during PEG-IFN/RBV therapy

Full-length NS3 protease sequences were determined in 39 non-SVR patients after PEG-IFN/RBV therapy. A single amino acid change at resistance-associated sites in two patients was observed. In one patient, isoleucine (Ile) at position 153 changed to valine (Val), and glutamic acid (Glu) changed to aspartic acid (Asp) at position 168 in the second (Fig. 4). At the nucleotide level, ATC (Ile) changed to GTC (Val) in I153V, and GAA (Glu) changed to GAC (Asp) in E168D. Both mutations were caused by one nucleotide exchange. No other changes were observed in the other 37 patients.

Discussion

Here we report that in 18% (47/261) HCV genotype 1b-infected patients who had not been previously treated with NS3 PIs, the viral genome contained dominant amino acid mutations within the NS3 PI-resistance sites. Even after confining the data to established PI-resistance mutations, the mutation rate was still significant in 13.4% (35/261). No clinical differences were observed between patients

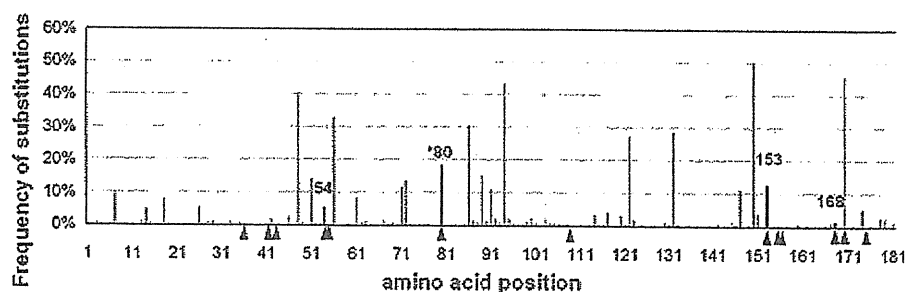


Fig. 1 Frequency of polymorphic mutations for each of the 181 NS3 protease amino acid residues in 261 patients. *Arrowheads* indicate the sites reported to confer PI resistance. *Dark bars* denote the amino acid

variations at the resistant sites in this study. *80, we detected one resistant mutation (Q80K) and 47 (18%) non-resistant variations (Q80L) at the 80th residue

Table 2 Characteristics of patients with or without HCV genomes harboring drug-resistance mutations

Characteristics	Mutation type (N = 47)	Wild-type (N = 214)	P value
Patients' characteristics			
Age, median (range)	59 (46–72)	57 (19–77)	0.17
Male, no. (%)	26 (55)	112 (52)	0.70
BMI, median (range)	23.2 (15.5–31.9)	22.8 (16.1–31.9)	0.41
ALT IU/ml	81.3 ± 72.6 ^a	74.8 ± 51.9	0.93
Serum albumin g/dl	4.00 ± 0.37	4.01 ± 0.36	0.81
Platelet count × 10 ⁴ /μl	15.8 ± 4.3	14.5 ± 4.8	0.18
HCV RNA KIU/ml, median (range)	1,500 (58–6,310)	1800 (28–15,849)	0.69
Fibrosis, no. (%)			0.97
F0	0 (0)	7 (3)	
F1	23 (50)	89 (42)	
F2	9 (20)	52 (24)	
F3	9 (20)	40 (19)	
F4	5 (11)	26 (12)	
IFN pre-treatment no. (%)	15/40 (38) ^b	66/172 (38)	1.00
IL28B (rs8099917) T/G or G/G no. (%)	6/20 (30)	19/67 (28)	1.00
Response to PEG-IFN/RBV therapy			
SVR total cases no. (%)	22/46 (48)	83/210 (40)	0.38
RVR in total cases no. (%)	6/44 (14)	22/195 (11)	0.83
cEVR in total cases no. (%)	22/44 (50)	92/200 (46)	0.75
SVR 48w treatment no. (%)	16/29 (55)	55/130 (42)	0.29
End of treatment response no. (%)	26/41 (63)	123/202 (61)	0.91
HCV genome sequence variation			
ISDR mutation ≤1 no. (%)	32/46 (70)	167/210 (80)	0.21
Core70 R no. (%)	26/44 (59)	136/210 (65)	0.56
IRRDR mutation >3 no. (%)	25/38 (66)	107/190 (56)	0.34

^a Mean ± SD^b Number/total number (%)

harboring viruses with and without these mutations. Moreover, no differences were observed in the responses of either group to PEG-IFN/RBV therapy.

Recent studies reported that significant number of patients who were never treated with PI possess viral sequences with PI-resistance-associated NS3 mutations. In these studies, the prevalence of PI-resistance mutations was determined to be 8.6–16.2% [13, 14], in HCV genotype 1- and 3-infected patients in European–American populations. These patients were often coinfecting with HIV. Analysis of the public HCV databases (EuHCVdb and Los Alamos) also reported the presence of naturally occurring PI-resistance-associated NS3 mutations in worldwide isolates [33]. However, in vivo and in vitro studies demonstrated that most of the mutations observed conferred only low- to moderate-level PI resistance [7, 13, 14, 34, 35]. Regarding viral fitness, PI-resistant HCVs show lower fitness at varying degrees as revealed by in vitro studies [1, 10, 11, 17, 20–22, 31, 32], but HCV RNA levels in a clinical study did not differ significantly. The response to PEG-IFN/RBV therapy was almost comparable to that in HCV-infected patients without PI-resistance mutations either in HCV replicon experiments or in a clinical study of small number of treated patients [34].

The prevalence of 13.4% for PI-resistance-proven patients observed in the present study was almost comparable to the results of previous studies. Although HIV is known to increase HCV replication in coinfection with HCV [36], and HIV patients are often treated with the HIV-specific PIs, the HIV infection might not affect the natural occurrence of HCV-specific PI-resistance mutations since our studied patients were all proven to be free from coinfection with HIV infection. As shown in Table 1 and Fig. 1, I153 V (22/261, 8.4%), T54S (14/261, 5.4%), and D168E (4/261, 1.5%) were among the most prevalent PI-resistance-proven mutations in the present study. The most frequent mutation detected in our study I153V was reported to appear secondarily to the occurrence of R109K mutations in a HCV replicon system [17]. Although the role of this mutation is not understood, the I153V mutation on its own conferred SCH446211 resistance to the HCV replicon to a lesser degree [17]. Interestingly, I153V was often found in double mutations in our study, as shown in Fig. 2. This suggests analogy between in vitro and in vivo data. T54S and D168E, the other frequent mutations, have been also reported to occur as single dominant mutations in previous in vitro or in vivo studies in HCV genotype 1

Fig. 2 In vivo fitness of HCV with PI-resistance-associated NS3 mutations. HCV RNA levels were compared between patients with and without NS3 PI-resistance-associated mutations (a) and between patients with each resistance mutation (b). The estimated *P* value (Mann–Whitney *U* test) indicates no significant difference between the wild-type and other groups (wild-type vs. mutation type, wild-type vs. single mutation type, and wild-type vs. double mutation type), (Wild-type, *N* = 214; mutation type, *N* = 47; single mutation type, *N* = 40; double mutation type, *N* = 7; V36I, *N* = 2; T54S, *N* = 14; Q80K, *N* = 1; I153L, *N* = 11; I153V, *N* = 22; D168E, *N* = 4; E176A, *N* = 1; V36I + I153V, *N* = 1; T54S + I153V, *N* = 4, and I153V + D168E, *N* = 2)

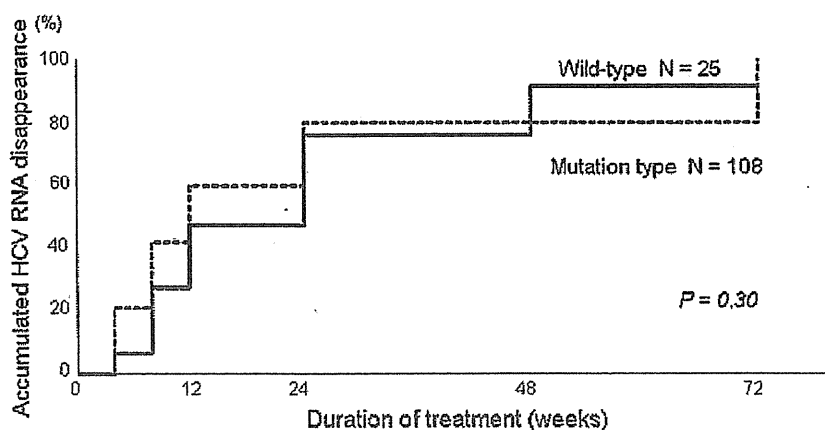
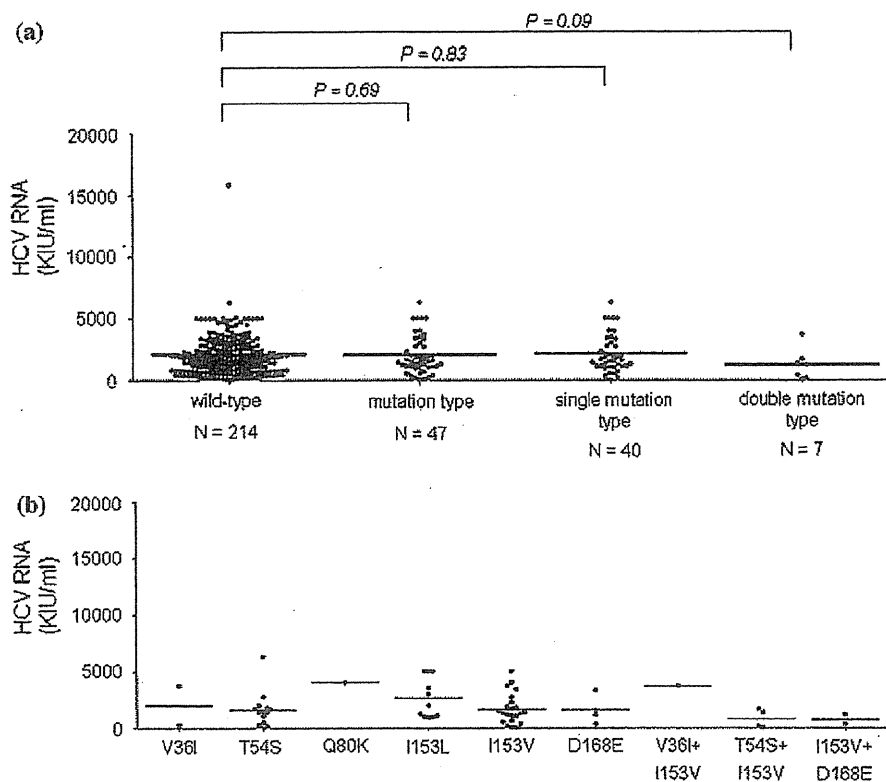


Fig. 3 Comparison of virologic response to PEG-IFN/RBV therapy between HCV-infected patients with and without PI-resistance-associated NS3 mutations. Time-dependent HCV clearance rate analysis was based on serum HCV RNA positivity during PEG-IFN/RBV therapy for HCV isolates with resistance mutations or wild-

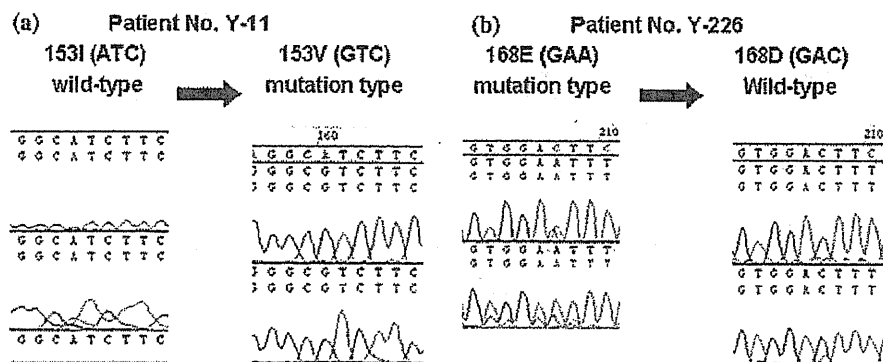
type sequences. A total of 133 patients for whom the limit of viral genome detection could be determined were analyzed. Among this group, NS3 mutations were detected in 25 patients (19%). The estimated *P* value (log-rank test) shows no significant difference between the two groups (*P* = 0.30)

infections showing moderate degrees of resistance [16, 18, 19].

Most PI-resistance mutations described to date have been associated with varying degrees of reduced replicative

capacity [10, 11, 17, 20–22, 31, 32]. In the present study, HCV RNA levels of those patients with low- to moderate-level resistance mutations were similar to those in patients in the wild-type groups, suggesting that in vitro viral fitness

Fig. 4 Appearance of PI-resistance-associated NS3 mutations during the PEG-IFN/RBV therapy. Chromatograms show part of the HCV NS3 sequence demonstrating PI-resistance mutations in two patients receiving therapy. **a** Site 153 isoleucine (Ile) (ATC) changed to valine (Val) (GTC), **b** Site 168 glutamic acid (Glu) (GAA) changed to aspartic acid (Asp) (GAC)



does not necessarily reflect in vivo viral fitness. This, however, does not rule out the possibility that some unknown compensatory viral mutations might have resulted in upregulation of reduced viral fitness. Interestingly, although the replicative capacity conferred by a single mutation seemed to be the same, the HCV RNA levels of double mutations were frequently low, suggesting that double mutations might weaken viral fitness.

In previous studies, clinical characteristics representing the state of liver disease other than HCV RNA levels were not studied in patients with PI-resistance mutations. In this study, we show that those clinical characteristics did not differ according to the presence of viral NS3 mutations. As shown in Table 2, age, sex ratio, fibrosis stage, ALT levels, serum albumin, platelet count, and past history of IFN pretreatment did not differ according to the presence of NS3 mutations. These results suggest that NS3 mutations occur independently of disease progression. Moreover, no evident differences were observed between viral and host factors known to affect IFN-based treatment responses. However, viral amino acid variations in the core and NS5A or the allelic frequency of IL28B SNPs, which were recently reported for the close relationship of responses to PEG-IFN/RBV therapy, did not differ between the two groups.

A significant outcome of the present study is the demonstration that PI-resistance mutations might not affect responses to PEG-IFN/RBV therapy. Previous in vitro studies demonstrated that HCV replicons harboring PI-resistance mutations were also sensitive to IFN treatment [31]. In addition, recent clinical studies also indicated that PI-resistance mutations were sensitive to the PEG-IFN/RBV [10, 34]. However, our analysis was more comprehensive because viral and host factors that contribute to treatment responses were simultaneously analyzed. A unique aspect of the present study is that we investigated the influence of the PEG-IFN/RBV treatment on the occurrence of new PI mutations by direct nucleotide sequencing, and were able to show that the PEG-IFN/RBV might not induce amino acid mutations.

Will the pre-existence of naturally occurring PI-resistance mutations have an influence on future treatment of HCV infections? Since new PIs are on the verge of clinical use, all clinicians should bear in mind the substantial numbers of HCV-infected patients with PI-resistance mutations. Although the degree of resistance is considered to be low or moderate in untreated patients, weak resistance might progress to more potent resistance with additional mutations, when PIs become widely used. Therefore, all clinicians need to be sufficiently prepared for the possibility of later onset of PI-resistance mutations that confer greater drug resistance and concomitant poorer responses to therapy. In SPRINT-1 study, the lead-in therapy was associated with a modestly lower rate of breakthrough than with no lead in [7]. Considering that PEG-IFN/RBV was equally effective for PI-resistant viruses, sufficient "lead-in" therapy before the administration of PIs could be an option in the forthcoming triple therapy modality.

In conclusion, we demonstrate here that PI-resistance-associated NS3 mutations exist in a substantial proportion of untreated HCV-1b-infected patients. Although the degree of resistance might not be strong, clinicians will need to consider this upon the introduction of triple therapy.

Acknowledgements This work was supported in part by a grant-in-aid scientific research fund from the Ministry of Education, Science, Sports, and Culture [grant number 20390206] and in part by a grant-in-aid from the Ministry of Health, Labor, and Welfare of Japan [grant number H19-kanen-002].

References

1. Susser S, Welsch C, Wang Y, et al. Characterization of resistance to the protease inhibitor boceprevir in hepatitis C virus-infected patients. *Hepatology* 2009;50:1709–1718
2. Hadziyannis SJ, Sette H Jr, Morgan TR, et al. Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Ann Intern Med* 2004;140:346–355

3. Hezode C, Forestier N, Dusheiko G, et al. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N Engl J Med* 2009;360:1839–1850
4. McHutchison JG, Everson GT, Gordon SC, et al. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. *N Engl J Med* 2009;360:1827–1838
5. McHutchison JG, Manns MP, Muir AJ, et al. Telaprevir for previously treated chronic HCV infection. *N Engl J Med* 2010;362:1292–1303
6. Reesink HW, Fanning GC, Farha KA, et al. Rapid HCV-RNA decline with once daily TMC435: a phase I study in healthy volunteers and hepatitis C patients. *Gastroenterology* 2010;138:913–921
7. Kwo PY, Lawitz EJ, McCone J, et al. Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naive patients with genotype 1 hepatitis C infection (SPRINT-1): an open-label, randomised, multicentre phase 2 trial. *Lancet* 2010;376:705–716
8. Forestier N, Reesink HW, Weegink CJ, et al. Antiviral activity of telaprevir (VX-950) and peginterferon alfa-2a in patients with hepatitis C. *Hepatology* 2007;46:640–648
9. Suzuki F, Akuta N, Suzuki Y, et al. Rapid loss of hepatitis C virus genotype 1b from serum in patients receiving a triple treatment with telaprevir (MP-424), pegylated interferon and ribavirin for 12 weeks. *Hepatol Res* 2009;39:1056–1063
10. Kieffer TL, Sarrazin C, Miller JS, et al. Telaprevir and pegylated interferon-alpha-2a inhibit wild-type and resistant genotype 1 hepatitis C virus replication in patients. *Hepatology* 2007;46:631–639
11. Sarrazin C, Kieffer TL, Bartels D, et al. Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir. *Gastroenterology* 2007;132:1767–1777
12. Lin C, Lin K, Luong YP, et al. In vitro resistance studies of hepatitis C virus serine protease inhibitors, VX-950 and BILN 2061: structural analysis indicates different resistance mechanisms. *J Biol Chem* 2004;279:17508–17514
13. Kuntzen T, Timm J, Berical A, et al. Naturally occurring dominant resistance mutations to hepatitis C virus protease and polymerase inhibitors in treatment-naive patients. *Hepatology* 2008;48:1769–1778
14. Gaudieri S, Rauch A, Pfafferott K, et al. Hepatitis C virus drug resistance and immune-driven adaptations: relevance to new antiviral therapy. *Hepatology* 2009;49:1069–1082
15. Lenz O, Verbinnen T, Lin TI, et al. (2010) In vitro resistance profile of the hepatitis C virus NS3/4A protease inhibitor TMC435. *Antimicrob Agents Chemother* 54:1878–1887
16. Kieffer TL, Kwong AD, Picchio GR. Viral resistance to specifically targeted antiviral therapies for hepatitis C (STAT-Cs). *J Antimicrob Chemother* 2010;65:202–212
17. Yi M, Tong X, Skelton A, et al. Mutations conferring resistance to SCH6, a novel hepatitis C virus NS3/4A protease inhibitor. Reduced RNA replication fitness and partial rescue by second-site mutations. *J Biol Chem* 2006;281:8205–8215
18. Welsch C, Domingues FS, Susser S, et al. Molecular basis of telaprevir resistance due to V36 and T54 mutations in the NS3-4A protease of the hepatitis C virus. *Genome Biol* 2008;9:R16
19. Tong X, Bogen S, Chase R, et al. Characterization of resistance mutations against HCV ketoamide protease inhibitors. *Antiviral Res* 2008;77:177–185
20. He Y, King MS, Kempf DJ, et al. Relative replication capacity and selective advantage profiles of protease inhibitor-resistant hepatitis C virus (HCV) NS3 protease mutants in the HCV genotype 1b replicon system. *Antimicrob Agents Chemother* 2008;52:1101–1110
21. Tong X, Chase R, Skelton A, Chen T, Wright-Minogue J, Malcolm BA. Identification and analysis of fitness of resistance mutations against the HCV protease inhibitor SCH 503034. *Antiviral Res* 2006;70:28–38
22. Lu L, Pilot-Matias TJ, Stewart KD, et al. Mutations conferring resistance to a potent hepatitis C virus serine protease inhibitor in vitro. *Antimicrob Agents Chemother* 2004;48:2260–2266
23. Akuta N, Suzuki F, Hirakawa M, et al. Amino acid substitutions in the hepatitis C virus core region of genotype 1b affect very early viral dynamics during treatment with telaprevir, peginterferon, and ribavirin. *J Med Virol* 2010;82:575–582
24. Ge D, Fellay J, Thompson AJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009;461:399–401
25. Tanaka Y, Nishida N, Sugiyama M, 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
26. Suppiah V, Moldovan M, Ahlenstiel G, et al. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 2009;41:1100–1104
27. El-Shamy A, Nagano-Fujii M, Sasase N, Imoto S, Kim SR, Hotta H. Sequence variation in hepatitis C virus nonstructural protein 5A predicts clinical outcome of pegylated interferon/ribavirin combination therapy. *Hepatology* 2008;48:38–47
28. Enomoto N, Sakuma I, Asahina Y, et al. Mutations in the non-structural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *N Engl J Med* 1996;334:77–81
29. Akuta N, Suzuki F, Kawamura 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;46:403–410
30. Akuta N, Suzuki F, Sezaki H, et al. Predictive factors of virological non-response to interferon-ribavirin combination therapy for patients infected with hepatitis C virus of genotype 1b and high viral load. *J Med Virol* 2006;78:83–90
31. Zhou Y, Bartels DJ, Hanzelka BL, et al. Phenotypic characterization of resistant Val36 variants of hepatitis C virus NS3-4A serine protease. *Antimicrob Agents Chemother* 2008;52:110–120
32. Mo H, Lu L, Pilot-Matias T, et al. Mutations conferring resistance to a hepatitis C virus (HCV) RNA-dependent RNA polymerase inhibitor alone or in combination with an HCV serine protease inhibitor in vitro. *Antimicrob Agents Chemother* 2005;49:4305–4314
33. Lopez-Labrador FX, Moya A, Gonzalez-Candelas F. Mapping natural polymorphisms of hepatitis C virus NS3/4A protease and antiviral resistance to inhibitors in worldwide isolates. *Antivir Ther* 2008;13:481–494
34. Bartels DJ, Zhou Y, Zhang EZ, et al. Natural prevalence of hepatitis C virus variants with decreased sensitivity to NS3.4A protease inhibitors in treatment-naive subjects. *J Infect Dis* 2008;198:800–807
35. Cubero M, Esteban JI, Otero T, et al. Naturally occurring NS3-protease-inhibitor resistant mutant A156T in the liver of an untreated chronic hepatitis C patient. *Virology* 2008;370:237–245
36. Soriano V, Puoti M, Sulkowski M, et al. Care of patients with hepatitis C and HIV co-infection. *Aids* 2004;18:1–12
37. Barbotte L, Ahmed-Belkacem A, Chevalier S, et al. Characterization of V36c, a novel amino acid substitution conferring hepatitis C virus (HCV) resistance to telaprevir, a potent peptidomimetic inhibitor of HCV protease. *Antimicrob Agents Chemother* 2010;54:2681–2683