Toward the Establishment of a Prediction System for the Personalized Treatment of Chronic Hepatitis C

Hidenori Ochi, C. Nelson Hayes, Hiromi Abe, Yasufumi Hayashida, Tomotaka Uchiyama, Naoyuki Kamatani, Yusuke Nakamura, and Kazuaki Chayama

1Laboratory for Digestive Diseases, Center for Genomic Medicine, RIKEN, Hiroshima, 2Department of Medicine and Molecular Science, Division of Frontier Medical Science, Programs for Biomedical Research, Graduate School of Biomedical Sciences and 3Liver Research Project Center, Hiroshima University, 4Center for Genomic Medicine, Riken, Yokohama, and 5Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Japan

Background. Although several direct-acting antivirals (DAAs) are now available, the therapy regimen for chronic hepatitis C will continue to include pegylated interferon and ribavirin for the foreseeable future. Despite their improved rate of sustained virological response (SVR), DAAs pose increased risks of side effects and selection for antiviral resistance. Not all patients require DAA to achieve SVR, whereas others are unlikely to respond even to triple therapy. Therefore, a personalized approach to candidate selection is necessary.

Methods. In this retrospective study, data from 640 Japanese patients who were treated for chronic hepatitis C genotype 1, 2, or 3 with pegylated interferon plus ribavirin combination therapy was compiled to identify robust pretreatment predictive factors for SVR.

Results. A logistic regression model for personalized therapy was developed based on age, viral genotype, initial viral load, aspartate aminotransferase/alanine aminotransferase ratio, α-fetoprotein levels, and IL28B single-nucleotide polymorphism genotype. The area under the receiver-operating characteristic curve (AUC) was 0.85. The mean AUC following 10 rounds of 10-fold cross validation was 0.82, with a true positive rate of 78.2%.

Conclusions. A personalized approach to therapy may better inform treatment decisions and reduce incidence of side effects and antiviral resistance.

The hepatitis C virus (HCV) affects >100 million people worldwide and is a major global cause of chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma [1–5]. The current standard of care, pegylated interferon (PEG-IFN) plus ribavirin combination therapy, is both expensive and poorly tolerated, and treatment efficacy is <50% for genotype 1b [6]. Telaprevir and boceprevir, 2 direct-acting antiviral (DAA) protease inhibitors, have recently been approved for clinical use in the United States [7] and are expected to improve the rate of sustained virological response (SVR) to 65%–75% [8]. However, the addition of a DAA to the current standard of care increases the risk of side effects, including anemia and rash, and failure to achieve SVR may pose an increased risk of accumulating protease inhibitor–resistant viral strains that may be recalcitrant to future treatment [8]. Consequently, it may be advantageous to identify patients who are unlikely to respond to therapy, as well patients who are likely to achieve SVR under the current standard of care without requiring a DAA. Patients who are able to achieve at least a transient response (relapsers) under combination therapy are more likely to achieve a SVR under triple therapy, whereas patients who fail to respond to combination therapy are also less likely to respond to triple therapy [9]. Therefore, it may be possible for patients who are highly likely to respond to combination therapy to be spared the additional risks.
and costs of triple therapy, but to determine the optimal treatment for each patient, reliable and inexpensive pretreatment predictors are needed for SVR.

A number of pretreatment predictors associated with SVR or nonresponse have been reported. Older female patients have been shown to respond poorly to therapy in Japan [10–12], and metabolic factors such as obesity [13], insulin resistance [13], hepatic steatosis [14], low-density lipoprotein (LDL) cholesterol levels [15, 16], and γ-glutamyl transpeptidase (γ-GTP) levels [17] have been shown to influence treatment outcome. Baseline virus titer is also an important predictor of treatment outcome [14, 18]. HCV genotypes differ in the response to interferon therapy, with genotypes 1 and 4 considered more difficult to treat than genotypes 2 and 3 [18, 19]. The importance of these and other factors in triple therapy remains unclear, although they may influence the effectiveness of interferon and ribavirin in suppressing emergence of resistant strains.

Genetic differences among patients also influence response to treatment and incidence of side effects. Genomewide association studies have reported common single-nucleotide polymorphisms (SNPs) predictive of response to interferon therapy. A set of linked SNPs within the IL28B locus on chromosome 19 has recently been shown to be the strongest predictor of sustained virological response as well as spontaneous viral clearance [20–26]. So far, the SNP also appears to be the strongest predictor for triple therapy [9, 27]. Other SNPs are associated with the occurrence of side effects. In particular, SNPs in the ITTPA locus have been found to be associated with anemia in patients treated with PEG-IFN plus ribavirin combination therapy [28–30] and appear to be predictive of anemia in triple therapy as well [31]. Although there are currently few options for treating HCV, SNP genotyping may nonetheless help gauge expectations and help identify patients at risk for severe side effects that may disrupt the course of therapy.

Even though telaprevir and boceprevir are now available for use in clinical practice, DAAs must be coadministered with PEG-IFN and ribavirin, to prevent rapid selection for resistance mutations [32]. As a result, patients who respond poorly to PEG-IFN and ribavirin may not only fail to achieve SVR under triple therapy but may be more likely to encounter viral breakthrough, with confounding effects on future treatment efforts. Consequently, there remains a need to identify robust predictors for response to PEG-IFN and ribavirin to establish a personalized approach for treatment of chronic hepatitis C.

 METHODS

 Patients

Data from 640 patients who were treated with PEG-IFN plus ribavirin combination therapy for chronic HCV infection were compiled from hospitals belonging to the Hiroshima Liver Study Group (http://home.hiroshima-u.ac.jp/naika1/e/) in Hiroshima, Japan. All patients were interferon treatment–naive and were infected with HCV genotype 1, 2, or 3. Study participants tested positive for HCV RNA over a span of >6 months, tested negative for hepatitis B and human immunodeficiency virus (HIV), and showed no evidence for other liver diseases. All patients gave written informed consent to participate in the study in accordance with the ethical guidelines of the 1975 Declaration of Helsinki and according to the process approved by the ethical committees of Hiroshima University and the SNP Research Center at the Institute of Physical and Chemical Research in Yokohama. Patient profiles are shown in Table 1 and Figure 1.

 PEG-IFN Plus Ribavirin Combination Therapy

Patients received weekly injections of PEG-IFN-α-2b at 1.5 μg/kg body weight for 48 weeks. Ribavirin was administered orally, and the dosage was determined based on the patient’s body weight (600 mg for <60 kg, 800 mg for 60–80 kg, 1000 mg for >80 kg), based on guidelines by the Ministry of Health, Labor, and Welfare of Japan [33]. The ribavirin dose was reduced when hemoglobin levels fell below 10 g/dL, and both PEG-IFN and ribavirin were discontinued when hemoglobin levels dropped to <8.5 g/dL.

 Outcome of Therapy

We evaluated treatment success based on SVR, defined as undetectable HCV RNA levels 24 weeks after cessation of treatment. Some patients showed a transient response (relapsers), in which HCV RNA dropped to undetectable levels during treatment but then rebounded during the follow-up period. In nonresponders, HCV-RNA levels failed to decline by 2 log10 IU/mL by week 12 of treatment and never dropped below detectable levels. Histopathological diagnosis was made as described previously [34].

 HCV RNA Levels

We monitored HCV RNA levels throughout the course of therapy using polymerase chain reaction (PCR)–based methods (the original Amplicor method, the high-range method, or the TaqMan real-time PCR test). The measurement ranges of these assays were 0.5–850 KIU/mL, 5–5000 KIU/mL, and 1.2–7.8 log10 IU, respectively. Samples exceeding the measurement range were diluted with phosphate-buffered saline and reanalyzed. All values were reported as log10 international units per milliliter.

 SNP Genotyping

We genotyped each patient for 2 SNPs: rs8099917 in the IL28B locus, which is associated with therapy outcome, and rs127354 in the ITTPA locus, which is associated with ribavirin-induced anemia. Although there are 2 SNPs associated with ITTPA enzyme activity in white patients [28], 1 of these SNPs appears to
be fixed in the Japanese population, and so only rs1127354 was genotyped [29]. Samples were genotyped using the Illumina HumanHap610-Quad Genotyping BeadChip or the Invader or TaqMan assay, as described previously [35].

Statistical Analysis

All analysis was performed using the R statistical package (http://www.r-project.org). Nonparametric tests (χ² and Mann–Whitney U tests) were used to detect significant associations. All statistical analyses were 2 sided, and P < .05 was considered significant. Multiple logistic regression analysis with forward/backward stepwise selection of variables was used to identify independent factors associated with SVR. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated for each factor. Receiver-operating characteristic (ROC) curves and areas under the curve (AUC) were calculated for each model using the ROCR software package. CIs for predicted SVR probabilities were calculated over a range of ages and viral loads, and results were stratified by IL28B SNP genotype and viral genotype using the rms software package. Models were validated based on 10 rounds of 10-fold cross-validation using the WEKA data-mining package [36].

RESULTS

Patient Characteristics

Patient profiles are shown in Table 1. In total, 388 (61%) patients achieved SVR, 119 (19%) were transient responders, and 119 (19%) were nonresponders. The frequency of the deleterious allele for the IL28B SNP rs8099917 (G) was 0.14. A total of 476 patients had the favorable TT genotype, and 148 and 13 patients had the unfavorable GT and GG genotypes, respectively. Genotype data for rs12979860, another commonly reported IL28B SNP, were not available for all patients, but the 2 SNPs are in high linkage disequilibrium and genotypes are highly correlated (0.99). The frequency of the favorable allele for the ITPA SNP rs1127354 (A) was 0.15. A total of 468 patients had the anemia-susceptible CC genotype, and 152 and 19 patients had the protective AC and AA genotypes, respectively.
Predictive Factors for SVR

The following predictors were significantly associated with SVR using univariate analysis after Bonferroni correction for 26 tests: male sex, age, genotype 1b, initial viral load, aspartate aminotransferase/alanine aminotransferase (AST/ALT) ratio, rs8099917 TT genotype, diabetes mellitus, α-fetoprotein level, white blood cell count, platelet count, and hemoglobin level (Table 2). When all factors were included in multivariate analysis, genotype, age, viral genotype, and initial viral load. The probability of achieving SVR, whereas some patients with an unfavorable genotype were able to achieve SVR. Inclusion of other viral and host factors is therefore expected to improve the accuracy of treatment-outcome predictions. Although a number of predictors have been reported, this study achieves relatively high accuracy using only a simple subset of pretreatment predictors, the most important of which are IL28B SNP genotype, age, viral genotype, and initial viral load.

A prediction equation based on the coefficients in Table 3 was used to generate the predicted response over a range of ages and viral loads (Figure 3). For example, a 60-year-old patient with the favorable IL28B SNP genotype and HCV genotype 1b has a probability of SVR of 0.61, whereas the probability is only 0.23 for a patient with an unfavorable IL28B genotype. On the other hand, the probability increases to 0.80 for a 40-year-old patient or 0.88 for a patient with genotype 1a, 2, or 3. Based on this model, it appears that older patients who have high viral load for genotype

DISCUSSION

In this study, we present a simple predictive model for outcome of PEG-IFN plus ribavirin combination therapy for patients infected with HCV. Although the IL28B SNP is the best single predictor of SVR, not all patients with the favorable genotype achieved SVR, whereas some patients with an unfavorable genotype were able to achieve SVR. Inclusion of other viral and host factors is therefore expected to improve the accuracy of treatment-outcome predictions. Although a number of predictors have been reported, this study achieves relatively high accuracy using only a simple subset of pretreatment predictors, the most important of which are IL28B SNP genotype, age, viral genotype, and initial viral load.

A prediction equation based on the coefficients in Table 3 was used to generate the predicted response over a range of ages and viral loads (Figure 3). For example, a 60-year-old patient with the favorable IL28B SNP genotype and HCV genotype 1b has a probability of SVR of 0.61, whereas the probability is only 0.23 for a patient with an unfavorable IL28B genotype. On the other hand, the probability increases to 0.80 for a 40-year-old patient or 0.88 for a patient with genotype 1a, 2, or 3. Based on this model, it appears that older patients who have high viral load for genotype

Table 2. Univariate Predictors for Sustained Virological Response

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR</th>
<th>95% CI</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>1.93</td>
<td>(1.39–2.7)</td>
<td>.00101</td>
</tr>
<tr>
<td>Age</td>
<td>0.948</td>
<td>(0.933–0.963)</td>
<td>1.66 × 10^{-12}</td>
</tr>
<tr>
<td>BMI</td>
<td>0.961</td>
<td>(0.91–1.01)</td>
<td>.0135</td>
</tr>
<tr>
<td>Genotype 1b vs others</td>
<td>0.202</td>
<td>(0.127–0.311)</td>
<td>1.09 × 10^{-13}</td>
</tr>
<tr>
<td>Virus titer, log IU/mL</td>
<td>0.39</td>
<td>(0.285–0.524)</td>
<td>7.12 × 10^{-9}</td>
</tr>
<tr>
<td>ALT</td>
<td>1.28</td>
<td>(1.01–1.63)</td>
<td>.07166</td>
</tr>
<tr>
<td>AST/ALT ratio</td>
<td>0.311</td>
<td>(0.177–0.528)</td>
<td>5.97 × 10^{-6}</td>
</tr>
<tr>
<td>rs8099917 17 TT genotype</td>
<td>4.24</td>
<td>(2.9–6.26)</td>
<td>2.75 × 10^{-6}</td>
</tr>
<tr>
<td>rs1127354 CC genotype</td>
<td>0.64</td>
<td>(0.432–0.939)</td>
<td>.0237</td>
</tr>
<tr>
<td>γ-GTP</td>
<td>0.999</td>
<td>(0.966–1)</td>
<td>.04023</td>
</tr>
<tr>
<td>Fibrosis</td>
<td>0.744</td>
<td>(0.599–0.929)</td>
<td>.02826</td>
</tr>
<tr>
<td>Activity</td>
<td>1.12</td>
<td>(0.909–1.39)</td>
<td>.2868</td>
</tr>
<tr>
<td>α-Fetoprotein</td>
<td>0.69</td>
<td>(0.559–0.844)</td>
<td>1.30 × 10^{-5}</td>
</tr>
<tr>
<td>LDL cholesterol</td>
<td>0.997</td>
<td>(0.992–1)</td>
<td>.3184</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>0.998</td>
<td>(0.995–1)</td>
<td>.04564</td>
</tr>
<tr>
<td>HDL cholesterol</td>
<td>0.993</td>
<td>(0.976–1.01)</td>
<td>.558</td>
</tr>
<tr>
<td>Iron</td>
<td>0.995</td>
<td>(0.99–1)</td>
<td>.01145</td>
</tr>
<tr>
<td>Fasting blood sugar</td>
<td>0.989</td>
<td>(0.981–0.997)</td>
<td>.01578</td>
</tr>
<tr>
<td>White blood cells</td>
<td>1</td>
<td>(1–1)</td>
<td>.000859</td>
</tr>
<tr>
<td>Platelets</td>
<td>2.32</td>
<td>(1.46–3.82)</td>
<td>.000342</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>1.25</td>
<td>(1.11–1.41)</td>
<td>.000224</td>
</tr>
<tr>
<td>Core aa 70 substitution (0–1 vs &gt;1)</td>
<td>1.35</td>
<td>(1.57–3.69)</td>
<td>.548</td>
</tr>
<tr>
<td>Core aa 91 substitution</td>
<td>1.24</td>
<td>(0.491–3.19)</td>
<td>.6604</td>
</tr>
<tr>
<td>ISDR substitutions</td>
<td>3.87</td>
<td>(1.4–12)</td>
<td>.01034</td>
</tr>
<tr>
<td>Hypertension</td>
<td>0.497</td>
<td>(0.278–0.878)</td>
<td>.01599</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>0.206</td>
<td>(0.0841–0.454)</td>
<td>6.90 × 10^{-5}</td>
</tr>
</tbody>
</table>

Abbreviations: ALT, aspartate aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; CI, confidence interval; γ-GTP, γ-glutamyl transpeptidase; HDL, high-density lipoprotein; OR, odds ratio.

P < .001; **P < .01; *P < .05; (*) P < .05 but not significant following Bonferroni correction for multiple testing (α = 0.05/26 = .0019)
Several predictive models for outcome of combination therapy for HCV have been reported and have used a variety of different approaches. Several studies have used artificial neural networks [37–39] and support vector machines [40] to predict SVR, although these types of models are more difficult to interpret than regression-based methods and are less amenable to adoption in clinical use. Other studies have used decision trees [41] and classification and regression tree analysis [42], both of which provide an intuitive, flowchart-based approach to prediction. However, small changes can dramatically alter the topology of the tree, and individual paths through the tree may make use of only a fraction of the available data. Medrano et al proposed a logistic-regression SVR prediction model for patients in Spain coinfected with HCV and HIV [43]. The model achieved high accuracy (AUC, 0.85–0.89) using 4 predictors: IL28B SNP genotype, liver stiffness, viral genotype, and viral load. For HCV-monoinfected patients, O’Brien et al [44] proposed a prediction model for SVR in European–American patients with genotype 1 based on IL28B SNP genotype, viral load, AST/ALT ratio, fibrosis score, and prior ribavirin treatment. The model proposed here is similar to the model proposed by O’Brien et al, differing mainly in patient ethnicity (Japanese vs European–American) and treatment history (prior ribavirin treatment vs treatment-naive), although patients in the model of O’Brien et al had more severe fibrosis (Ishak fibrosis score ≥3 vs 0–4), and were younger (median age, 49 vs 59 years) and more likely to be male (73% vs 51%). Both models had similar factors and AUC scores (0.79 vs 0.82), and the inclusion of various host and viral factors in both models underscores the variability in response to therapy and the limitations of IL28B SNP genotype alone in predicting the outcome of therapy. Presumably, future studies will introduce models geared specifically for response to triple therapy, but until additional data become available, predictions based on response to combination therapy may help guide patient selection.

**CONCLUSIONS**

Pretreatment predictors based on clinical and viral factors may be used to predict the outcome of therapy. Regardless of the approach or the specific predictors analyzed, most prediction studies report a consistent set of important predictive factors, including viral genotype, IL28B SNP genotype, age, viral load, and ≥1 clinical factors reflecting liver function (eg, γ-GTP, LDL cholesterol, blood sugar, α-fetoprotein, and platelet count). By adopting a personalized approach to treatment, clinicians may be better able to determine the most appropriate course of therapy for individual patients while minimizing the risk of side effects.

**Notes**

**Acknowledgments.** We thank Mika Tsuzuno, Sakura Akamatsu, Sanae Furuya, and other members of the clerical staff and Yoshikiku Kawakami and other physicians at Hiroshima University Hospital for their help.

**Financial support.** This study was supported by the Ministry of Health, Labor, and Welfare of Japan (21390229).
Figure 3. Predicted probabilities of sustained virological response (SVR) for patients with genotype 1b. Confidence intervals for predicted probability of SVR based on logistic regression by age, grouped by rs8099917 genotype and initial viral load for patients with genotype 1b, are shown. Solid lines represent the favorable rs8099917 TT single-nucleotide polymorphism genotype, and dashed lines represent the unfavorable GT or GG genotype.

Potential conflicts of interest. All authors: No reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References


