Liver Cancer

DOI: 10.1159/000489505
Published online: June 14, 2018
© 2018 S. Karger AG, Basel
www.karger.com/lic

Original Paper

Surrogacy of Time to Progression for Overall Survival in Advanced Hepatocellular Carcinoma Treated with Systemic Therapy: A Systematic Review and Meta-Analysis of Randomized Controlled Trials

Takeshi Terashima\textsuperscript{a} Tatsuya Yamashita\textsuperscript{a} Tadashi Toyama\textsuperscript{b} Kuniaki Arai\textsuperscript{a} Kazunori Kawauchi\textsuperscript{a} Kazuya Kitamura\textsuperscript{a} Taro Yamashita\textsuperscript{a} Yoshio Sakai\textsuperscript{a} Eishiro Mizukoshi\textsuperscript{a} Masao Honda\textsuperscript{a} Shuichi Kaneko\textsuperscript{a}

\textsuperscript{a}Department of Gastroenterology, Kanazawa University Hospital, Kanazawa, Japan; \textsuperscript{b}Department of Nephrology, Kanazawa University Hospital, Kanazawa, Japan

Keywords
Time to progression · Overall survival · Surrogate endpoint · Hepatocellular carcinoma · Clinical trials · Molecular targeted therapy

Abstract
Time to progression (TTP) is widely used as the endpoint in early-phase trials of advanced hepatocellular carcinoma (HCC). However, the relevance of using TTP as a surrogate marker for overall survival (OS) in pivotal trials remains uncertain. The PubMed database and ASCO Meeting Library were searched for reports of randomized controlled trials that investigated patients with advanced HCC, included data for both OS and TTP, and were launched between 2009 and 2016. The correlation between hazard ratios (HRs) for TTP and OS was determined using weighted linear regression. Correlations between median OS and TTP, and between median OS and postprogression survival (PPS), defined as the period obtained by subtracting the median TTP from the median OS, were also evaluated. The database search yielded 24 trials with 50 arms. Overall, TTP HR correlated with OS HR ($R = 0.73$); however, the coefficient in the regression equation was 0.48. The correlation between median OS and median TTP was not so strong ($R = 0.50$), whereas the correlation between median OS and median PPS was strong ($R = 0.78$). In advanced HCC, the OS HR can be predicted from the TTP HR, which is useful when considering whether to proceed to a pivotal trial based on the results of early-phase trials. TTP may be a better endpoint than OS for evaluating a novel agent in a pivotal trial, because an improvement in antitumor effect cannot fully reflect an improvement in OS due to the strong impact of PPS on OS.

© 2018 S. Karger AG, Basel

Tatsuya Yamashita, MD, PhD
Department of Gastroenterology, Kanazawa University Hospital
13-1 Takaramachi, Kanazawashi
Ishikawa 920-8641 (Japan)
E-Mail ytatsuya@m-kanazawa.jp

www.karger.com/lic

DOI: 10.1159/000489505
Introduction

Overall survival (OS) has been considered the optimal and firm endpoint in clinical trials for cancer patients who cannot receive curative treatment, because the aim for such patients is prolongation of survival. However, it takes a long time to complete a clinical trial when OS is set as the primary endpoint, and an alternative endpoint, such as time to progression (TTP), is commonly used for early-phase trials to evaluate the treatment efficacy of a novel agent over a short time [1, 2]. Although the objective response rate was also suggested as a representative indicator across various cancer patients, its suitability as a surrogate endpoint for an agent that does not act through shrinkage of the tumor is still debated [3, 4].

Hepatocellular carcinoma (HCC) is the sixth most common cancer and the third leading cause of cancer-related death worldwide [5]. Since sorafenib showed a survival benefit for patients with advanced HCC [6, 7], many novel agents have been tested in randomized phase III trials in first-line [8–14] and second-line [15–19] settings. Some agents have shown a significant antitumor effect in terms of prolongation of the TTP [9, 12, 15, 17, 18]; however, with the exception of regorafenib [20], none have been proven to prolong survival.

All trials to date have set OS as the primary endpoint, and the relevance of using TTP as a surrogate marker for OS remains unclear. We evaluated the correlation between TTP hazard ratio (HR) and OS HR based on a systematic review and meta-analysis of randomized controlled trials to clarify the value of improvement of TTP as a surrogate for improvement of OS. We also evaluated the correlation between median OS (mOS) and median TTP (mTTP) and between mOS and median postprogression survival (mPPS), which was defined as the period obtained by subtracting mTTP from mOS. This approach will clarify the applicability of TTP as a surrogate marker for OS and provide information that is useful for planning and interpreting clinical trials on patients with advanced HCC.

Materials and Methods

Search and Selection of Literature

We reviewed PubMed citations and the ASCO Meeting Library up to December 31, 2016. Keywords included in the search were “advanced hepatocellular carcinoma” and “randomized trial.” The search was limited to reports of randomized controlled trials written in English. Furthermore, we also checked the Cochrane Library and ESMO Conference Platform to confirm the randomized controlled trials which should be included in this study. We reviewed each report including publications on the website and selected studies that compared two or more arms for treatment of advanced HCC with a medical agent.

Included in this study were randomized phase II or phase III trials that provided an analysis of intent-to-treat data for OS HRs and either TTP HRs or progression-free survival (PFS) HRs of an experimental arm against a control arm. We excluded review articles, letters or commentaries, interim analyses, subgroup analyses of previously reported trials, or duplicate reports. Reports in which all or some patients were treated in an adjuvant setting or neoadjuvant setting – or with concomitant treatment procedures including resection, transplantation, transarterial chemoembolization, transarterial infusion chemotherapy, or radiation therapy in addition to systemic therapy – were also excluded. To avoid bias, two investigators (T. Terashima and K.A.) independently reviewed and abstracted the data from the reports, and another investigator (T. Yamashita) made the final decision if agreement was not reached between the first two investigators.

Data Extraction

For simplicity, PFS data were collectively referred to as TTP in the present analysis if TTP was not addressed, which is similar to an approach adopted in a recent report [21]. TTP and OS HRs, mTTP, and mOS were extracted from all reports that provided data. mPPS was defined as the period obtained by subtracting mTTP from mOS for each report. The survival data were converted into months; 1 month was considered to be 30.45 days in reports that described survival data in days. We also obtained the following information from each report: the number of arms, trial phase, publication status, treatment line, the type of control arm, and the number of patients in each arm.
Data Analysis

We summarized the survival data (TTP HR, OS HR, mOS, mTTP, and mPPS) as the median value for all treatment arms. The relationship between TTP HR and OS HR was estimated using weighted linear regression, with weights equal to the sample size of the arms from which the data were derived [22]. All reported \( p \) values correspond to two-sided tests, with \( p < 0.05 \) considered to be statistically significant. The strength of the association was assessed by using \( R \), and we considered the correlation to be strong when \( R \geq 0.7 \). We also evaluated whether the relationship changed according to treatment line (first vs. second line), assessment of tumor progression (TTP vs. PFS), trial phase (randomized phase II vs. phase III), and publication status (published vs. unpublished trials). The relationship between mOS and either mTTP or mPPS was analyzed in the same way. The data analyses were performed using Stata 12.1 (College Station, TX, USA).

Results

Characteristics of the Trials

We identified 30 trials in the initial search of PubMed and ASCO Meeting Library citations, but excluded 6 trials because of lack of information about TTP HR and/or OS HR (see online suppl. Fig. 1; see www.karger.com/doi/10.1159/000489505 for all online suppl. material). The remaining 24 trials were included in the present study. The characteristics of the 24 trials, which included 50 arms and 9,556 patients with advanced HCC, are shown in Table 1 and the online supplementary Table. Sixteen trials were in the first-line setting and 8 trials were in the second-line setting. The number of patients in each arm ranged from 19 to 578, with a median of 134.5.

Relationships between HRs for TTP and OS

The TTP HRs of experimental arms against the control arm ranged from 0.5 to 1.4, with a median TTP HR of 0.915, whereas the corresponding OS HR ranged from 0.44 to 1.3, with a median OS HR of 0.918.

The OS HR of the 24 trials was plotted against the TTP HR (Fig. 1). Overall, TTP HR strongly correlated with OS HR (\( R = 0.73, p < 0.001 \)) based on weighted linear regression.

Table 1. Characteristics of the trials used in this study

<table>
<thead>
<tr>
<th>Trials</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arms in each trial</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Phase of trial</td>
<td>14</td>
</tr>
<tr>
<td>Phase III</td>
<td>10</td>
</tr>
<tr>
<td>Randomized phase II</td>
<td></td>
</tr>
<tr>
<td>Publication status</td>
<td>19</td>
</tr>
<tr>
<td>Published</td>
<td>5</td>
</tr>
<tr>
<td>Unpublished</td>
<td></td>
</tr>
<tr>
<td>Treatment line</td>
<td>16</td>
</tr>
<tr>
<td>First</td>
<td>8</td>
</tr>
<tr>
<td>Second</td>
<td></td>
</tr>
<tr>
<td>Control arm</td>
<td>12</td>
</tr>
<tr>
<td>Placebo</td>
<td>2</td>
</tr>
<tr>
<td>Doxorubicin</td>
<td>10</td>
</tr>
<tr>
<td>Sorafenib</td>
<td></td>
</tr>
<tr>
<td>Arms</td>
<td>50</td>
</tr>
<tr>
<td>Patients in each arm</td>
<td>134.5</td>
</tr>
<tr>
<td>Median</td>
<td>19–578</td>
</tr>
<tr>
<td>Range</td>
<td></td>
</tr>
</tbody>
</table>
**Fig. 1.** Relationship between overall survival hazard ratio (OS HR) and time to progression/progression-free survival HR (TTP/PFS HR) among all trials. Each trial is represented by a circle, with its size proportional to the number of patients. TTP/PFS HR was strongly correlated with OS HR ($R = 0.73$, $p < 0.001$) on the basis of weighted linear regression (*).

**Fig. 2.** Relationship between overall survival hazard ratio (OS HR) and time to progression/progression-free survival HR (TTP/PFS HR) among second-line trials (a) and first-line trials (b). Each trial is represented by a circle, with its size proportional to the number of patients. TTP/PFS HR was more strongly correlated with OS HR in a second-line setting ($R = 0.91$, $p = 0.002$) than in a first-line ($R = 0.77$, $p < 0.001$) setting on the basis of weighted linear regression (*).
However, the coefficient in the regression equation was not much greater than 0.48, and the regression equation was \( \log(\text{OS HR}) = 0.48 \log(\text{TTP HR}) + 0.050 \). When trials were stratified by treatment line, TTP HR was more strongly correlated with OS HR in second-line \((R = 0.91, p = 0.002)\) (Fig. 2a) than in first-line \((R = 0.77, p < 0.001)\) (Fig. 2b) settings. Corresponding tendencies were observed in the sensitivity analysis conducted to confirm the robustness of the above results among the reports in which tumor progression was assessed by TTP \((R = 0.70, p < 0.001)\) (Fig. 3a) and PFS \((R = 0.71, p = 0.003)\) (Fig. 3b), the phase III trials (see online suppl. Fig. 2A), and the published trials (see online suppl. Fig. 2B).

**Fig. 3.** Relationship between overall survival hazard ratio (OS HR) and time to progression (TTP) HR among the trials in which tumor progression was assessed by TTP (a) and progression-free survival (PFS) (b). Each trial is represented by a circle, with its size proportional to the number of patients. TTP HR and PFS HR were consistently correlated with OS HR \((R = 0.72 \text{ and } R = 0.70, \text{ respectively})\) on the basis of weighted linear regression (*).

### Relationships between Median Times of OS and TTP or PPS

mOS, mTTP, and mPPS for all arms was 4.2–13.7 months (median 8.6), 0.95–6.4 months (median 3.0), and 2.8–9.8 months (median 5.2), respectively. The proportion of mPPS to mOS for all arms was 40.7–86.6%, with a median of 63.9%.

mOS in the 50 arms was plotted against mPPS (Fig. 4a) and mTTP (Fig. 4b). The correlation between mOS and mTTP was not so strong \((R = 0.50, p < 0.001)\), whereas the correlation between mOS and mPPS was strong \((R = 0.78, p < 0.001)\), based on weighted linear regression.
Discussion

This study investigated the correlation between TTP HR and OS HR and the correlation between mOS and either mTTP or mPPS by a systematic review and meta-analysis of randomized controlled trials of advanced HCC. Our analysis yielded three main findings: (1) TTP HR was strongly correlated with OS HR; (2) the coefficient in the regression equation between TTP HR and OS HR was not much greater than 0.48; and (3) mPPS was strongly correlated with mOS, whereas the correlation between mTTP and mOS was not so strong. These results have implications for planning and interpreting clinical trials on advanced HCC patients.

First, an important finding of this study was that the correlation between mTTP and mOS was not so strong, whereas the correlation between TTP HR and OS HR was strong. TTP is widely used as the primary endpoint in phase II trials of novel agents for unresectable solid tumors, and the judgement as to whether the agent should be tested in phase III trials and planning of such trials is generally formed based on the results of the phase II trial. However, the obtained information about OS may not have adequate power to predict a benefit from OS because the number of patients in the phase II trial is calculated based on the hypothesis
regarding the expected TTP. Our data, showing a not so strong correlation between mTTP and mOS, suggest that the wrong decision may be made if planning is based only on information about mTTP. In contrast, in our data the correlation between TTP HR and OS HR was strong, indicating that it is possible to predict the OS HR from data on the TTP HR. Based on these findings, it is essential to have a control arm in order to obtain information about the HR when designing a phase II trial, since sufficient information for successful planning of a consequent pivotal study cannot be obtained from a single-arm trial that provides only median survival data [23]. Moreover, the decision as to whether a novel agent should proceed to be tested in a pivotal trial should be made based on previous results showing that the agent does not merely provide a statistical improvement in antitumor effect but has a clinically significant antitumor effect contributing to an improvement in patient outcome for efficient and successful development, all of which can be evaluated via HRs provided from an early-phase trial of a randomized design.

Second, with regard to the rationale of clinical trials in advanced HCC it is very important to note that the line of the regression equation in this study did not correspond to the line representing TTP HR equal to OS HR (shown as a dotted line in Fig. 1), and that the coefficient in the regression equation between TTP HR and OS HR was only 0.48; the risk reduction in TTP HR directly reflects the risk reduction in OS HR if the TTP HR corresponds to the OS HR and the coefficient is 1.0. OS is an indicator for a patient’s prognosis, whereas TTP represents a direct clinical benefit, and whether TTP can be a surrogate for clinical benefit depends on the magnitude of the effect and the risk benefit of the new treatment compared to available therapies [24]. Our results suggest that the antitumor effect obtained from a treatment contributes only partially to the improvement in patient prognosis in advanced HCC. This tendency can be explained by our findings that mPPS strongly correlated with mOS [25]. In patients with advanced-stage HCC, locoregional therapies such as hepatic arterial infusion chemotherapy or transarterial chemoembolization have the potential to be efficacious even after sorafenib therapy [26–28], although such therapies have not been verified in proper prospective trials.

Lee et al. [29] concluded that TTP can be a surrogate endpoint for OS based only on the finding of a good correlation between TTP (or PFS) HR and OS HR; however, it is important to consider the objective of a clinical trial in addition to the correlation between TTP HR and OS HR when we evaluate the applicability of using TTP as a surrogate marker for OS. On one hand, we can certainly predict the OS HR from data on the TTP HR, as described above for an early-stage trial with the aim of judging as to whether there is value in proceeding to a pivotal trial. On the other hand, it is not appropriate to evaluate TTP as a surrogate endpoint of OS in a pivotal trial if the trial aims to verify the survival benefit from novel agents, because the contribution of the antitumor effect to the survival benefit is weakened by the strong impact of PPS on OS.

Although the US Food and Drug Administration or European Medicines Agency have usually requested proof of improvement in survival in pivotal trials for the regulatory approval of new agents [24], an increasing number of agents have been approved based on the results of a trial setting in which an indicator such as TTP directly represents the antitumor effect. This trend has often been seen for cancers such as breast cancer, colorectal cancer, or non-small cell lung cancer, for which several effective agents are available and it seems to be difficult to evaluate the survival benefit in a clinical trial due to the noise of PPS [30–32]. As it is well known that TTP has several disadvantages as an endpoint in this setting compared with OS [24], it is particularly important to verify this relationship and reconsider whether we can evaluate the improvement of survival prolongation in pivotal trials of advanced HCC in which the antitumor effect does not directly reflect the patients’ prognosis [33]. It should originally be the antitumor effect that is evaluated in any development study of a novel agent.
Now, we may have to consider setting an indicator such as TTP which directly reflects the antitumor effect as the primary endpoint in pivotal studies of HCC as well as of breast cancer, colorectal cancer, and non-small cell lung cancer.

After our analysis, the results of the RESORCE trial were published, revealing that regorafenib improved the outcome of patients with HCC who progressed on sorafenib treatment compared with placebo in a second-line setting [20]. The point plotting OS HR (0.63) of this trial against TTP HR (0.44) was located under the line representing the regression equation (see online suppl. Fig. 3), indicating that the risk reduction in TTP HR contributed more to the risk reduction in OS HR compared with previous failed trials. The reasons for this successful result were widely discussed, and they include the second-line setting, which had fewer negative effects from effective posttrial treatment. The last-line setting must be one of the elements of a successful clinical trial design that enables the antitumor effect to directly improve patients’ outcomes [25].

The present study has some limitations, including the lack of individual patient data and potential confounders because of the inclusion of heterogeneous trials. Further investigations including details such as subsequent therapy, liver function, and progression patterns affecting the relationship between OS and either TTP or PPS [1] will contribute to a more efficient development of novel agents.

Conclusions

In conclusion, by a systematic review and meta-analysis of randomized controlled trials of advanced HCC, we revealed that OS HR strongly correlated with TTP HR, especially in second-line trials, whereas the correlation between mOS and mTTP was not so strong, even though mPPS strongly correlated with mOS. Predicting OS HRs from data on TTP HRs is useful when considering whether to proceed to a pivotal trial based on the results of early-phase trials. OS may be unsuitable as a primary endpoint in a pivotal trial because the improvement in antitumor effect cannot fully reflect the improvement in OS due to a strong impact of PPS on OS; we should consider to set TTP as the primary endpoint in a pivotal trial where the antitumor effect should be originally evaluated.

Acknowledgments

We thank Mary Derry, PhD ELS, from Edanz Group Ltd (www.edanzediting.com/ac) for editing a draft of this manuscript, and Prof. Kenichi Yoshimura for statistical advice.

Disclosure Statement

The authors do not have any conflict of interest.
References

26 Kudo M: Regorafenib as second-line systemic therapy may change the treatment strategy and management paradigm for hepatocellular carcinoma. Liver Cancer 2016; 5:235–244.