Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

Usefulness of alpha-fetoprotein response in patients treated with sorafenib for advanced hepatocellular carcinoma

Nicola Personeni[1,*], Silvia Bozzarelli[1], Tiziana Pressiani[1], Lorenza Rimassa[1], Maria Chiara Tronconi[1], Francesco Sclafani[1], Carlo Carnaghi[1], Vittorio Pedicini[2], Laura Giordano[3], Armando Santoro[1]

[1]*Department of Oncology-Hematology, Humanitas Cancer Center, IRCCS, Rozzano, Italy;* [2]*Department of Diagnostic Imaging, Humanitas Cancer Center, IRCCS, Rozzano, Italy;* [3]*Biostatistician, Humanitas Cancer Center, IRCCS, Rozzano, Italy*

Background & Aims: Tumor shrinkage has been considered a fundamental surrogate efficacy measure for new cancer treatments. However, in patients treated with sorafenib for advanced hepatocellular carcinoma (HCC), tumor shrinkage rarely accompanies increased survival, thereby questioning the prognostic value of imaging-based Response Evaluation Criteria in Solid Tumors (RECIST). We investigated the prognostic usefulness of a decrease in serum alpha-fetoprotein (AFP) and compared it to RECIST.

Methods: In HCC patients treated with sorafenib with baseline AFP >20 ng/ml, AFP response was defined as a >20% decrease in AFP during 8 weeks of treatment. Patients were also assessed by RECIST and were categorized as having radiologically proven progressive disease or disease control (consisting of complete or partial responses and stable disease). Comparisons of survival by RECIST and AFP response were corrected for guarantee-time bias by the landmark method.

Results: We evaluated 85 patients for AFP response, among them, 82 were also evaluated by RECIST. In the analysis of AFP response, 32 out of 85 patients (37.6%) were responders, whereas 58 out of 82 patients (70.7%) achieved disease control. In landmark analysis, the hazard ratios (HR) for survival according to AFP response and disease control were 0.59 ($p = 0.040$) and 1.03 ($p = 0.913$), respectively. In multivariate analysis, only AFP response (HR = 0.52; $p = 0.009$) and Cancer of the Liver Italian Program dichotomized stage (HR = 0.42; $p = 0.002$) were prognostic factors of survival.

Conclusions: Assessment of AFP response may be considered as an alternative to RECIST to capture sorafenib activity in HCC.

Keywords: Alpha-fetoprotein; Hepatocellular carcinoma; Sorafenib; AFP response; Survival; RECIST; Landmark analysis.

Received 9 August 2011; received in revised form 6 January 2012; accepted 26 February 2012; available online 10 March 2012

* Corresponding author. Address: Department of Oncology-Hematology, Humanitas Cancer Center, IRCCS, Via Manzoni 56, 20089 Rozzano, Italy. Tel.: +39 02 8224 4516; fax: +39 02 8224 4592.

E-mail address: nic_personeni@hotmail.com (N. Personeni).

Abbreviations: HCC, hepatocellular carcinoma; RECIST, Response Evaluation Criteria in Solid Tumors; AFP, alpha-fetoprotein; HR, hazard ratio; CT, computed tomography; CLIP, Cancer of Liver Italian Program; OS, overall survival; TTP, time to disease progression; 95% CI, 95% confidence intervals.

Introduction

For patients with advanced hepatocellular carcinoma (HCC) not amenable to curative or locoregional therapies, few treatment options exist, as common cytotoxic chemotherapeutic regimes are ineffective. Recently, two phase III clinical trials found that sorafenib, an inhibitor of tyrosine kinases involved in tumor proliferation and angiogenesis, conferred significant survival advantages over placebo [1,2]. Radiologically documented tumor shrinkage (tumor response) has been generally used as an end point in research on new anticancer drugs, based on the assumption that a reduction in tumor size results in prolonged survival [3]. The increase in survival in sorafenib-treated patients compared to controls, however, was not accompanied by a corresponding decrease in tumor size [1,2].

Novel anticancer agents that are not directly cytotoxic may not reduce tumor volume, but rather stabilize it, as seems to be the case for sorafenib [1,2,4]. Given these premises, a dichotomization of radiological outcomes into disease progression and non-progression (also termed "disease control") may be more meaningful than the traditional classification suggested by the Response Evaluation Criteria in Solid Tumors (RECIST), which considers objective response, stable disease and disease progression [5]. Although in the SHARP trial [1] the sorafenib-treated group achieved a significantly higher disease control rate, the sizeable group of patients who do not have disease progression, which represents about 70% of patients enrolled in either arm of the trial, jeopardizes the possibility of identifying who really benefits from sorafenib. Furthermore, these figures question the prognostic value of imaging-based response criteria.

An alternative to the use of imaging-based criteria may be represented by the tumor marker alpha-fetoprotein (AFP). The serum concentration of AFP is increased in 60–70% of patients with HCC. Advanced or metastatic HCC and HBV-related disease identify those conditions most frequently associated with increased AFP levels [6,7].

This glycoprotein has long been measured in clinical practice for the surveillance of patients with cirrhosis at risk of HCC, although this use has been criticized due to its poor diagnostic performance [8].

Recent attention has turned to measuring changes in AFP during treatment for HCC. In 2005, Chen *et al.* [9] showed that,

Cancer

in patients with advanced HCC and elevated AFP (>200 ng/ml), a >50% decline in serum AFP was predictive of the efficacy of thalidomide treatment, and they therefore introduced the concept of "AFP response." Bearing in mind that the applicability of AFP response criteria is limited by definition to patients with increased AFP levels (>20 ng/ml), three subsequent studies confirmed the prognostic value of AFP response in HCC patients medically treated. Chan *et al.* [10] studied patients with baseline AFP levels >20 ng/ml and reported that a >20% decline in AFP after two or more cycles of palliative chemotherapy was prognostic of treatment efficacy. Vora *et al.* [11] analyzed data from five multicenter clinical trials of chemotherapy and targeted treatments and concluded that a $\geqslant$50% increase in AFP during the course of treatment was a useful surrogate marker of clinical outcome in patients whose baseline AFP levels were rated as not normal (i.e. exceeding 6.1–8.7 ng/ml depending on the institution). Most recently, Shao *et al.* [12] analyzed data from 72 patients enrolled onto clinical trials which investigated the efficacy of metronomic chemotherapy in combination with various antiangiogenic agents, including sorafenib; a >20% decrease in AFP within 2–4 weeks of treatment predicted response to treatment and prognosis. Therefore, the current study was specifically designed to assess whether AFP response is associated with lengthened survival after treatment with sorafenib alone. A second goal of this study was to compare AFP response to imaging-based response criteria as surrogate end point for survival.

Patients and methods

Study population

We retrospectively evaluated clinical records of patients with advanced HCC treated with sorafenib in one out of three clinical trials conducted at Humanitas Cancer Center. Two of these studies have already been published [1,4], while the third has been recently presented in abstract form [13]. Patients included in these trials had metastatic or locally advanced HCC that was not amenable to surgery or locoregional therapies, including transcatheter arterial (chemo)embolization, and local ablation. The patients received sorafenib as first-line treatment at a dose of 400 mg twice daily. Treatment was continued until disease progression or toxicity occurred. Tumor response to sorafenib was evaluated every 8 weeks by computed tomography (CT) according to RECIST [5]. Concomitant with the radiological evaluation, serum samples were taken for AFP determination using a chemiluminescent immunoassay; assays for individual patients were always done by the same laboratory. Patients were included in this retrospective study if AFP values at baseline and 8 weeks of treatment were available and if the baseline AFP was >20 ng/ml. The three clinical studies were approved by the Institutional review board of Humanitas Cancer Center. Written informed consent for the analysis of clinical records had been given by the patients at the time of enrollment in the trials. For each patient, we obtained pretreatment data regarding age, gender, cause of disease, and serum values of AFP, alanine transaminase, and bilirubin, Child–Pugh class, Eastern Cooperative Oncology Group performance status, Cancer of Liver Italian Program (CLIP) stage [14], presence of extrahepatic disease and esophageal varices.

Definitions

Overall survival (OS) time was measured from the first day of treatment to the date of death or last contact. Time to disease progression (TTP) was measured from the first day of treatment until the day when disease progression, according to RECIST [5], was recorded. AFP variation was defined as the percent change in AFP concentration according to the formula $(AFP_{8weeks} - AFP_{Baseline})/AFP_{Baseline} \times 100$. Disease control was defined as the absence of disease progression at 8 weeks, and included complete or partial response and stable disease according to RECIST [5].

Survival analyses

To examine the relationship between changes in AFP values and survival, we first tried a time-dependent receiver operating characteristics curve analysis in the attempt to generate a cut-off point for AFP variation that predicted survival. However, the area under the curve was 0.60, indicating limited prognostic reliability of AFP variation. Therefore, we categorized patients as either responders or non-responders on the basis of a decrease in AFP concentration by more than 20% suggested by Chan *et al.* [10], and we used the Kaplan–Meier product-limit method to estimate survival probabilities, compared with the log-rank test. This was done for both OS and TTP. A subgroup analysis was performed for patients with baseline AFP >200 ng/ml, considering the concern on the prognostic role of AFP response in patients with slightly or moderately elevated AFP levels [15].

Results of 8-week follow-up CT examinations were used to classify tumor response to treatment according to RECIST and to categorize patients into disease control and disease progression subgroups. The Kaplan–Meier product-limit method was used to determine survival probabilities in the two subgroups, compared by the log-rank test. This was done for both OS and TTP, for all patients and for those with baseline AFP >200 ng/ml.

Hazards ratios (HR), 95% confidence intervals (95% CI), and p values were calculated for OS according to both AFP response and disease control using a univariate Cox proportional hazards model. This analysis was performed for all patients, for patients with baseline AFP >200 ng/ml and for patients who survived to a prespecified, arbitrary landmark time of 90 days after the date treatment was started. This latter analysis was specifically undertaken to eliminate guarantee-time bias [16] and was proposed by others in similar settings [17].

A multivariate Cox proportional hazards model was built to assess the prognostic value of AFP response taking into account the possible influence of confounding clinical variables. Only those variables that were significant in univariate analysis were entered into the multivariate model. This analysis was performed for all patients and for those who survived to the landmark time of 90 days.

Statistical analysis

Pearson's χ^2, Fisher's exact and Cochran–Mantel–Haenszel tests were used to investigate the association between categorical variables. Wilcoxon Mann–Whitney and two-tailed t tests were used to investigate differences between groups for continuous variables.

Statistical analyses were performed using the R package (version 2.0.1) for all tests except the preliminary ROC analysis, done with SAS software (version, 9.1.3, SAS Institute, Cary, USA). A p <0.05 was considered to indicate statistical significance.

Results

Review of clinical records from HCC patients treated at Humanitas Cancer Center in trials of sorafenib led to the identification of 85 patients with elevated baseline serum levels of AFP. In this group, baseline AFP ranged over 4000-fold from just above the inclusion criterion of 20 ng/ml to almost 90,000 ng/ml (Table 1). Mean AFP variation during the 8 weeks of treatment with sorafenib was +28.3% and median OS was 9.6 months.

AFP variation during sorafenib treatment and survival

According to the AFP response criteria, 32 patients with a 20% AFP decrease were classified as responders while the remaining 53 patients were defined as non-responders.

Median OS was 13.3 months among responders but only 8.2 months in non-responders (Fig. 1); this difference was significant (p = 0.022). Median TTP was 7.9 months among responders but only 2.4 months for non-responders (p = 0.004). No differences were observed between responders and non-responders in disease severity as indicated by clinical parameters such as Child–Pugh status and CLIP stage (Table 2). While median base-