Research Article

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

Comparison of the methods for tumor response assessment in patients with hepatocellular carcinoma undergoing transarterial chemoembolization

Eun Suk Jung[1], Ji Hoon Kim[1,*], Eileen L. Yoon[1], Hyun Jung Lee[1], Soon Jae Lee[1], Sang Jun Suh[1], Beom Jae Lee[1], Yeon Seok Seo[1], Hyung Joon Yim[1], Tae-Seok Seo[2], Chang Hee Lee[2], Jong Eun Yeon[1], Jong-Jae Park[1], Jae Seon Kim[1], Young Tae Bak[1], Kwan Soo Byun[1]

[1]*Department of Internal Medicine, Korea University College of Medicine, Seoul, Republic of Korea;* [2]*Department of Radiology, Korea University College of Medicine, Seoul, Republic of Korea*

Background & Aims: Recently, new methods, including the concept of viable enhancing tumor such as EASL and mRECIST, have been proposed for substitution of the conventional WHO and RECIST criteria in hepatocellular carcinoma (HCC) undergoing transarterial chemoembolization (TACE). Herein, we evaluated the differences of four methods and compared the association of these methods with the prognosis of HCC patients undergoing TACE.

Methods: We retrospectively reviewed 114 consecutive newly diagnosed HCC patients who underwent TACE as initial treatment. We evaluated the intermethod agreement (κ values) between the methods and compared their association with the prognosis of HCC patients.

Results: The κ values for EASL *vs.* WHO, EASL *vs.* RECIST, mRECIST *vs.* WHO, and mRECIST *vs.* RECIST were low, of 0.102, 0.088, 0.112, and 0.122, respectively. However, good correlations were observed for WHO *vs.* RECIST and EASL *vs.* mRECIST ($\kappa = 0.883$, $\kappa = 0.759$, respectively $p < 0.001$). The median OS was 32.3 months. Hazard ratios (HR) for survival in responders compared with non-responders were 0.21 (95% CI; 0.12–0.37, $p < 0.001$) for EASL and 0.27 (95% CI; 0.15–0.48, $p < 0.001$) for mRECIST. The mean survival of responders was significantly longer than that of non-responders in both EASL (40.8 *vs.* 16.9 months, $p < 0.001$) and mRECIST (41.1 *vs.* 20.7 months, $p < 0.001$). In multivariate analysis, EASL response (HR 0.21, 95% CI 0.11–0.40, $p < 0.001$) and mRECIST response (HR; 0.31, 95% CI, 0.17–0.59, $p < 0.001$) were independently associated with survival.

Conclusions: The response assessment by EASL and mRECIST could reliably predict the survival of HCC patients undergoing TACE and could be applicable in practice in preference to the conventional WHO and RECIST criteria.

Keywords: Hepatocellular carcinoma; Transarterial chemoembolization; Survival; Response.

Received 11 September 2012; received in revised form 26 December 2012; accepted 29 January 2013; available online 8 February 2013

* Corresponding author. Address: Department of Internal Medicine, Guro Hospital, Korea University College of Medicine, 97, Guro-Dong Gil, Guro-Dong, Guro-Gu, Seoul 152-703, Republic of Korea. Tel.: +82 2 2626 3011; fax: +82 2 2626 1038.

E-mail address: kjhhepar@naver.com (J.H. Kim).

Abbreviations: EASL, European Association for the Study of the Liver; RECIST, Response Evaluation Criteria in Solid Tumors; mRECIST, modified RECIST; WHO, World Health Organization; HCC, hepatocellular carcinoma; TACE, transarterial chemoembolization; BCLC, Barcelona Clinic Liver Cancer; OS, overall survival; TTP, time to progression; CT, computed tomography; MRI, magnetic resonance imaging; TARE, transarterial radioembolization; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; PVT, portal vein thrombosis; ECOG, Eastern Cooperative Oncology Group; HR, hazard ratio; afp, alfa feto-protein; DEB, drug eluting beads.

Introduction

Primary liver cancers, most of which consist of hepatocellular carcinomas (HCC), are increasing globally [1,2]. However, fewer than 20% of HCC patients could be candidates for curative therapy at the time of diagnosis, due to asymptomatic progression, underlying chronic liver disease, and specific tumor biologic characters, making this disease one of the poorest prognostic cancers.

Locoregional therapy, such as radiofrequency ablation or transarterial chemoembolization (TACE), is widely used to treat HCC as curative or palliative treatment. Although TACE has been the optimal therapy for patients with intermediate Barcelona Clinic Liver Cancer (BCLC) stage [3], TACE can be applied to patients in the early stage, who are ineligible for surgery due to poor residual liver function and/or co-morbidities, and for ablation due to tumor location. TACE is also a palliative treatment option for patients with advanced stage. In this respect, TACE is a widely applicable therapeutic option in the treatment of HCC and thus produces various results according to patient conditions. Nevertheless, TACE showed a survival benefit and became the standard treatments in HCC patients with BCLC intermediate stage [4,5].

Objective response assessment is important in the evaluation of the effect of anticancer treatment. The most important end point for approved anticancer therapy is overall survival (OS), but radiologic responses have been widely used as surrogate

Cancer

end points in phase II trials and as short-term decision guides to continue or change the ongoing therapy [6]. However, it has not been well evaluated whether object radiologic response could properly reflect prolonged survival of HCC patients undergoing TACE.

For the purpose of radiologic response evaluation, the World Health Organization (WHO) response criteria were introduced in 1979. The WHO criteria are based on the sum of bidimensional perpendicular products [7]. However, because of some limitations of the WHO criteria, the Response Evaluation Criteria in Solid Tumors (RECIST) were introduced in 2000 to unify and standardize the response assessment criteria [8]. The RECIST criteria, which were revised to version 1.1 in 2009 [9], are based on the sum of the unidimensional longest diameters. However, the WHO and RECIST criteria were designed for the evaluation of cytotoxic agents. In the case of molecular targeted therapy or locoregional therapy such as TACE, clinical benefit is not always correlated with shrinkage of tumor size, but could be correlated with necrosis of a viable tumor. TACE induces tumor necrosis with or without change in tumor size. Since the WHO and RECIST criteria are based on tumor size measurement, they have been considered as suboptimal methods for tumor response assessment in HCC patients, especially for those undergoing TACE. Therefore, recently, the European Association for the Study of the Liver (EASL) and the American Association for the Study of Liver Disease (AASLD) have proposed new methods, including the concept of viable enhancing lesion modifying WHO (EASL) [10] and RECIST (mRECIST) [11,12] criteria, respectively. However, EASL and mRECIST methods should be extensively validated by investigating their correlation with the survival in HCC patients undergoing TACE.

Herein, we investigated the differences among WHO, RECIST, EASL, and mRECIST and evaluated the optimal method for predicting radiologic end point of time to progression (TTP) and clinical benefits of OS.

Materials and methods

We retrospectively reviewed the medical records of 141 consecutive newly diagnosed HCC patients who underwent TACE as initial treatment between the period from August 2005 to November 2006. Diagnosis of HCC was confirmed by biopsy or radiologic imaging studies according to the guidelines [13]. Among the 141 patients, we excluded 43 patients due to early follow-up loss after 1st TACE (n = 6), no measurable enhancing lesions >1 cm (n = 8), other co-existing cancers (n = 4), main portal vein thrombosis (n = 5), extrahepatic metastasis (n = 13), and other treatment modalities before response assessment (n = 7). A total of 98 enrolled patients underwent TACE as initial therapy and then repeated TACE on demand at 4–8 weeks after the first cycle. Contrast-enhanced dynamic computed tomography (CT) was performed at baseline and 3–4 weeks after TACE, and was used for response assessment. When indicated, 15 (15.3%) patients underwent primovist-enhanced dynamic magnetic resonance imaging (MRI) to further clarify tumor viability. TACE was performed with the superselective method using 10–50 mg of adriamycin mixed with lipiodol (2–20 ml), through feeding arteries until arterial flow stasis was obtained or iodized lipiodol appeared in portal branches. Subsequently, we embolized the feeding arteries by absorbable gelform sponge particles. Tumor measurements were performed according to the WHO, RECIST, EASL, and mRECIST criteria (Supplementary Fig. 1 and Table 1) by two qualified radiologists (T.S.S, C.H.L). When there was any ambiguity in tumor measurement or response assessment, final determination was made with consensus. TACE is usually performed in more than one session and could be performed repeatedly, in case of suspicious viable lesions in follow-up examination after 4–8 weeks of each TACE as current guidelines. Although response assessment associating long-term prognosis at early time point during TACE sessions is needed to decide to continue TACE or change the treatment modality, the optimal time point to assess TACE response reflecting long-term clinical prognosis has been controversial. To guide the therapeutic decision at early time point during TACE and exclude the compounding factors through potential marginal or *de novo* recurrence during treatment, we compared treatment responses between baseline imaging at diagnosis and follow-up imaging at early time point after 1–2 sessions of TACE. The WHO criteria are based on the sum of bidimensional products and do not mention the minimum size or number of lesions [7]. The RECIST criteria are based on the sum of unidimensional diameters. The longest diameter needed to be more than 1 cm on spiral CT. Target lesions were a maximum of 5 lesions per organ and 10 lesions total [8]. The EASL criteria in 2001, based on the measurements of the WHO criteria, emphasized viable lesions consistent with enhancing lesions on dynamic CT or MRI [10]. The mRECIST criteria are based on the RECIST criteria with the addition of the concept of viable enhancing lesions [11]. Recently, the RECIST criteria were upgraded to version 1.1 in 2009 [9]. The revised criteria simplify and optimize the assessment of tumor burden that includes the overall tumor burden calculated by the sum of 2 lesions per organ (5 lesions total). Moreover, Riaz *et al.* reported that even a single index lesion was enough to measure the tumor response and it well correlated with the survival of HCC patients undergoing TACE and transarterial radioembolization (TARE) [14]. Therefore, to unify and simplify the heterogeneity of the four response assessment methods, we measured up to two lesions as target lesions with a minimum size of 1 cm as the measurable target lesion. Other lesions were regarded as non-target lesions in the response assessment. The final response assessment was decided by incorporating the target and non-target lesion response according to the respective guidelines. Objective response was defined as complete (CR) and partial response (PR). The patients with objective response were classified as responders and the others were classified as non-responders.

Baseline demographic and tumor characteristics were recorded for each patient, including the BCLC system. Intermethod agreement between the four methods was assessed by κ values.

Moreover, we investigated and compared the efficacy predicting TTP and OS through the presence of objective response of each method.

For statistical analysis, SPSS software version 11.0 was used (Chicago, IL).

Intermethod agreement was assessed using κ statistics. A κ coefficient over 0.75 represented excellent intermethod agreement and a κ coefficient of less than 0.21 represented poor intermethod agreement [15,16].

The Kaplan–Meier method with the log-rank test was used to calculate and compare the differences of TTP and OS between responders and non-responders of each method. Univariate and multivariate analyses were conducted to identify factors associated with survival.

The Cox proportional hazards model with forward conditional selection was used to make a multivariate model to determine the independent predictable factor for survival.

Results

The baseline characteristics of 98 enrolled patients are summarized in Table 1. 86.7% of the patients were male and the mean age was 59.6 ± 9.4 years. The etiology of HCC was mostly hepatitis B in 68 patients (69.4%). 77 patients (78.6%) were Child-Pugh class A and 21 (21.4%) Child-Pugh class B. Liver cirrhosis was documented in 72 patients (73.5%). ECOG performance status was 0 in 24 patients (24.5%), 1 in 66 (67.3%) and 2 in 8 (8.2%). Concerning the tumor characteristics, 24 patients (24.5%) showed portal vein thrombosis (PVT, 10 patients presented with portal vein thrombosis at first branch and the others showed it at less than second branch of portal vein).

Concerning tumor staging, 37, 38, and 23 patients were classified as early, intermediate, and advanced BCLC stage, respectively. The mean follow-up duration was 25.0 months.

Tumor response assessment according to each method is shown in Table 2. In the response assessment by the WHO criteria, CR was seen in 1 patient (1.0%), PR in 14 (14.3%), stable disease (SD) in 74 (75.5%), and progressive disease (PD) in 9 (9.2%). In the response assessment by the RECIST criteria, CR was seen in 1 patient (1.0%), PR in 12 (12.2%), SD in 77 (78.6%), and PD in 8 (8.2%). In the response assessment by the EASL criteria, CR was seen in 34 patients (34.7%), PR in 34 (34.7%), SD in 25 (25.5%), and PD in 5 (5.1%). In the response assessment by the