Glypican-3 versus alpha-fetoprotein as a biomarker for hepatocellular carcinoma: a diagnostic meta-analysis

Weiyu FU¹, Hongyu LU², Li LI³, Kefeng WU³, Yanping LI², Yi LIU⁴*, Jianli HU⁴*

¹ Institute of Biochemistry and Molecular Biology, Guangdong Medical University, Zhanjiang, Guangdong, China.
² College of Food Science and Technology, Guangdong Ocean University, Zhanjiang, Guangdong, China.
³ Guangdong Key Laboratory for Research and Development of Natural Drugs, Guangdong Medical University, Zhanjiang, Guangdong, China.
⁴ Cancer Center, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

ABSTRACT: Objective: To assess the diagnostic accuracy of serum GPC3 versus alpha-fetoprotein (AFP) for HCC by using the method of system review.

Methods: PubMed and EMBASE were searched from its inception to 20, April 2014 for studies that compared diagnostic accuracy of serum GPC3 with AFP for HCC. Sensitivity, specificity and other measures were pooled using random-effects models. Summary receiver operating characteristic curve analysis was used to summarize the overall test performance.

Results: Fourteen studies were included in this meta-analysis. Summary estimates for serum GPC3 and AFP in diagnosing HCC were as follows: sensitivity, 69% (95% confidence interval (CI), 56-80%) vs. 60% (95% CI, 50-69%); specificity, 91% (95% CI, 76-97%) vs. 92% (95% CI, 84-98%); diagnostic odds ratio (DOR), 22 (95% CI, 6-83) vs. 18 (95% CI, 6-83); and area under sROC, 0.85 (95% CI, 0.81-0.88) vs. 0.80 (95% CI, 0.76-0.83). The pooled sensitivity and specificity for (GPC3+AFP) and AFP were: sensitivity 80% (95% CI, 75-85%) vs. 64% (95% CI, 53-73%) and specificity 86% (95% CI, 74-93%) vs. 96% (95% CI, 86-99%). A significant heterogeneity was found among the ten studies, and meta-regression and subgroup meta-analysis suggested that race and assay type were probably responsible for the heterogeneity.

Conclusions: Serum GPC3 may be a promising diagnostic marker of HCC and it was helpful for early detection of primary hepatocellular carcinoma when combined with AFP. More studies for specific race of patients, and using certain methods for detecting GPC3 are required to further confirm the diagnostic value of GPC3 for HCC.

Key words: Glypican-3, AFP, Hepatocellular carcinoma, diagnostic, meta-analysis

Introduction

Primary hepatocellular carcinoma (HCC) is one of the most malignant tumors and its 5-year survival rate is less than 10% (Ferlay et al., 2010; Blechacz et al., 2013). Lots of patients are diagnosed at an advanced stage and lost the opportunity of effective treatment. Monitoring serum α-fetoprotein (AFP) and liver ultrasound (US) every 6 months are recommended by the European Association for the Study of the Liver (EASL) (European Association for the Study of the Liver, 2009). However, US depend on operator’s skill (European Association for the Study of the Liver, 2009), and AFP detection is also not satisfactory with the sensitivity ranging from 40% to 65% and specificity ranging from 76% to 96% (Abu El Makarem, 2012). A new marker with better accuracy is needed for HCC diagnosis.

Glypican-3 (GPC3) is a member of the glypican family of heparan-sulfate proteoglycans (Jia et al., 2007). Recently, it was frequently reported that GPC3 was over expressed in HCC patients (Grozdanov et al., 2006; Kandil et al., 2007; Abd El Moety et al., 2011; Honsová et al., 2011), and could be detectable in serum. In addition, serum GPC3 has been suggested as a potential serum marker for detection of HCC, but the results were inconsistent (Capurro et al., 2003; Hippo et al., 2004; Beale et al., 2008; Liu et al., 2010; Youssef et al., 2010; Tangkijvanich et al., 2010; Zhang et al., 2010; Abd El
Moety et al., 2011; Ozkan et al., 2011; Qiao et al., 2011; Wang et al., 2011; Gomma A et al., 2012; Abdelgawad et al., 2013; Lee et al., 2014). The objective of the present meta-analysis was to synthesize and analyze the results from studies that evaluated the diagnostic accuracy of serum GPC3 by directly comparing it with AFP.

Materials and Methods

Inclusion and exclusion criteria
Eligible studies were those original research articles that compared the diagnostic accuracy of GPC3 test with AFP for HCC in the same patients, or randomly assigned patients to one of the tests using blood as the only sample type. Studies that evaluated serum GPC3 or AFP levels by messenger RNA, DNA or DNA polymorphisms, or those without providing the sensitivity or specificity of GPC or AFP were excluded.

Only studies published in English were included. Abstracts, letters, editorials and expert opinions, reviews without original data, case reports and studies lacking control groups were excluded. No restriction was imposed on the year of publication.

Identification of studies
Diagnostic studies were identified through searches of electronic databases PubMed and EMBASE from its inception through April 20, 2014. Subject headings and keywords used in the search process were: (1) GPC3: GPC3, glypican-3; and (2) HCC: HCC, hepatocellular carcinoma, liver cell carcinoma, hepatic cell carcinoma, liver cancer. No restriction was set on study design, year of publication and publishes status. To avoid missing relevant studies, we did not use keywords or indexing terms for diagnostic test accuracy. We also manually searched the reference lists of the selected articles to identify additional studies.

Study selection and data extraction
The first selection was carried out by one of the authors (W. Fu), on the basis of the title and abstract. The full paper of each potentially eligible study was then obtained. Two authors (W. Fu, H. Lu) independently assessed eligible studies for inclusion. Discrepancies were resolved by discussion and consensus. The following characteristics were extracted from each selected study: authors, year of publication, journal, study design, number of patients, reference test, assay type of the markers, cutoff values and raw data for the calculation of sensitivity and specificity (the number of true positive, false negative, true negative and false positive results). Any disagreements were resolved through consultation with a third author (L. Li).

Assessment of methodological quality
According to the QUADAS (Quality Assessment of studies of Diagnostic Accuracy included in Systematic reviews) criteria recommended by Cochrane Collaboration, we chose five items from the checklist to assess the studies quality (Capurro et al., 2003; Ling et al., 2008): (1) study design (i.e. cross-sectional versus case–control design); (2) comparison of the index test with an appropriate reference standard; (3) recruitment of patients either consecutively or randomly; vice versa; (4) complete verification of test results with the reference standard; and (5) blind interpretation of the test.

Data analysis
Sensitivity and specificity were calculated, and the diagnostic accuracy was summarized for each study. We present the data as forest plots and receiver operating characteristic curves (ROC). The forest plots display the sensitivity and specificity of individual studies with the corresponding 95% confidence intervals (CIs). The receiver operating characteristic curves show individual study data as circles, the 95% confidence and 95% prediction regions around the pooled estimate, and the hierarchical summary curve resulting from the hierarchical summary receiver operating characteristic model. The Chi-square test and Fisher’s exact test were used to detect heterogeneity among studies and to evaluate the degree of variability. Univariate meta-regression analysis was performed to assess the effects of race, assay type and cutoff value on the diagnostic accuracy of HCC. The potential presence of publication bias was measured by Deeks’ funnel plot asymmetry test (Deeks et al., 2005). Statistical hypotheses (two-tailed) were tested at the level of 5% significance. STATA (version 12.0) and Meta-Disc (version 1.4) were used for statistical analysis.
The included studies were shown in Table 1.

**TABLE 1**

Characteristics and methodology assessment of the included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>HCC/controls</th>
<th>Assay type</th>
<th>GPC3 cutoff value</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Assay type</th>
<th>AFP cutoff value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee et al., 2014</td>
<td>Korea</td>
<td>120/40</td>
<td>ELISA</td>
<td>73 ng/ml</td>
<td>65</td>
<td>1</td>
<td>26</td>
<td>1</td>
<td>Not known</td>
<td>6 ng/ml</td>
</tr>
<tr>
<td>Abdelgawad et al., 2013</td>
<td>Egypt</td>
<td>40/20</td>
<td>ELISA</td>
<td>4.9 ng/ml</td>
<td>38</td>
<td>2</td>
<td>19</td>
<td>1</td>
<td>Not known</td>
<td>40.5 ng/ml</td>
</tr>
<tr>
<td>Gomaa et al., 2012</td>
<td>Egypt</td>
<td>31/30</td>
<td>ELISA</td>
<td>5.41 ng/ml</td>
<td>28</td>
<td>1</td>
<td>3</td>
<td>29</td>
<td>Electrochemiluminescence immunoassay</td>
<td>42.32 ng/ml</td>
</tr>
<tr>
<td>Wang et al., 2012</td>
<td>China</td>
<td>78/97</td>
<td>ELISA</td>
<td>Not known</td>
<td>28</td>
<td>44</td>
<td>50</td>
<td>53</td>
<td>ELISA</td>
<td>Not known</td>
</tr>
<tr>
<td>Abd El Moety et al., 2011</td>
<td>Egypt</td>
<td>10/50</td>
<td>ELISA</td>
<td>2.0 ng/ml</td>
<td>10</td>
<td>29</td>
<td>0</td>
<td>21</td>
<td>ELISA</td>
<td>20 ng/ml</td>
</tr>
<tr>
<td>Ozkan et al., 2011</td>
<td>Turkey</td>
<td>75/83</td>
<td>ELISA</td>
<td>3.9 pg/ml</td>
<td>46</td>
<td>48</td>
<td>29</td>
<td>35</td>
<td>Immunoassays utilizing enhanced chemiluminescence immunoassay</td>
<td>13 ng/ml</td>
</tr>
<tr>
<td>Qiao et al., 2011</td>
<td>China</td>
<td>101/88</td>
<td>ELISA</td>
<td>26.8 ng/ml</td>
<td>52</td>
<td>6</td>
<td>49</td>
<td>82</td>
<td>Electrochemiluminescence immunoassay</td>
<td>199.3 ng/ml</td>
</tr>
<tr>
<td>Liu et al., 2010</td>
<td>China</td>
<td>75/32</td>
<td>ELISA</td>
<td>300 ng/L</td>
<td>35</td>
<td>2</td>
<td>40</td>
<td>30</td>
<td>Electrochemiluminescence immunoassay</td>
<td>400ug/L</td>
</tr>
<tr>
<td>Youssef et al., 2010</td>
<td>Egypt</td>
<td>40/40</td>
<td>ELISA</td>
<td>4.6 ng/ml</td>
<td>33</td>
<td>2</td>
<td>7</td>
<td>38</td>
<td>Chemiluminescence immunoassay</td>
<td>66 U/L</td>
</tr>
<tr>
<td>Zhang et al., 2010</td>
<td>China</td>
<td>36/93</td>
<td>Immunoassay</td>
<td>3.10 ng/ml</td>
<td>33</td>
<td>0</td>
<td>3</td>
<td>93</td>
<td>Chemiluminescence immunoassay</td>
<td>25 ng/ml</td>
</tr>
<tr>
<td>Tangkijvanich et al., 2009</td>
<td>Thailand</td>
<td>100/190</td>
<td>ELISA</td>
<td>Not known</td>
<td>53</td>
<td>2</td>
<td>47</td>
<td>188</td>
<td>ELISA</td>
<td>20 ng/ml</td>
</tr>
<tr>
<td>Beale et al., 2008</td>
<td>UK</td>
<td>50/41</td>
<td>ELISA</td>
<td>Not known</td>
<td>34</td>
<td>22</td>
<td>16</td>
<td>19</td>
<td>ELISA</td>
<td>15 ng/ml</td>
</tr>
<tr>
<td>Hippo et al., 2004</td>
<td>Japan</td>
<td>69/38</td>
<td>ELISA</td>
<td>2.0 ng/ml</td>
<td>35</td>
<td>4</td>
<td>34</td>
<td>34</td>
<td>Not known</td>
<td>20 ng/ml</td>
</tr>
<tr>
<td>Capurro et al., 2003</td>
<td>Canada</td>
<td>34/91</td>
<td>ELISA</td>
<td>Not known</td>
<td>18</td>
<td>1</td>
<td>16</td>
<td>90</td>
<td>Not known</td>
<td>20 ng/ml</td>
</tr>
</tbody>
</table>

TP: true positive; FP: false positive; TN: true negative; FN: false negative

**Results**

**Study selection**

A total of 800 potentially relevant articles were identified by electronic databases searches. After reviewing the titles or abstracts, 502 articles, including overlapped studies, case reports, reviews and comments, were excluded. After referring to full texts, 258 studies were excluded due to not relevant study design, 24 studies were excluded due to having not enough data to estimate sensitivity or specificity and 2 studies were excluded for overlapped data. We also searched reference lists of the retrieved studies, and no additional citations met the inclusion criteria. Eventually, fourteen studies were included in this meta-analysis (Capurro et al., 2003; Hippo et al., 2004; Beale et al., 2008; Liu et al., 2010; Youssef et al., 2010; Tangkijvanich et al., 2010; Zhang et al., 2010; Abd El Moety et al., 2011; Ozkan et al., 2011; Qiao et al., 2011; Wang et al., 2011; Gomma A et al., 2012; Abdelgawad et al., 2013; Lee et al., 2014) (Fig. 1). The characteristics of the included studies were shown in Table 1.

**Quality of studies**

QUADAS quality assessment of the included studies is shown in Table 2. Nine studies used a prospective design, and in four studies the blood samples were collected from consecutive patients. All the studies reported the diagnostic standard of HCC, and ten completely verified the test results with reference standard. However, none of the fourteen studies clearly stated blinding interpretation of index results.

**Summary diagnostic accuracy of GPC3 vs. AFP for HCC**

Fig. 2A and Fig. 2B presents the forest plots of sensitivity (true positive rate) and specificity (false positive rate) for the 14 studies. Fig. 3 presents the diagnostic values of the studies in a hierarchical summary receiver operating characteristic graph (SROC) for GPC3 and AFP. The sensitivity of these studies ranged from 36% to 100%, 33% to 82% for GPC3 and AFP levels in the diagnosis of HCC, respectively, while the specificity ranged from 42% to 100%, 60% to 100%. The summary sensitivity and specificity for GPC3 were comparable to AFP: sensitivity 69% (95% CI, 56-80%) vs. 60%.
(95% CI, 50-69%) and specificity 91% (95% CI, 76-97%) vs. 92% (95% CI, 84-98%). The differences between sensitivities and specificities were significant. We also noted that the summary diagnostic odds ratio (DOR) and the area under SROC was 22 (95% CI, 6-83) and 0.85 (95% CI, 0.81-0.88) for GPC3, and 18 (95% CI, 8-41) and 0.80 (95% CI, 0.76-0.83) for AFP.

**Diagnostic accuracy of GPC3+AFP vs. AFP for HCC**

Ten studies reported the diagnostic values of GPC3 in combination with AFP (GPC3 + AFP) vs. AFP (Capurro et al., 2003; Hippo et al., 2004; Beale et al., 2008; Liu et al., 2010; Tangkijvanich et al., 2010; Zhang et al., 2010; Ozkan et al., 2011; Qiao et al., 2011; Wang et al., 2011; Abdelgawad et al., 2013; Lee et al., 2014). The pooled sensitivity and specificity for (GPC3+AFP) and AFP were: 80% (95% CI, 75-85%) vs. 64% (95% CI, 53-73%) and specificity 86% (95% CI, 74-93%) vs. 96% (95% CI, 86-99%). (Fig. 2C).

**Investigation for heterogeneity**

Heterogeneity among the included studies was found for analyzing of the above results and in the sROC as well. To explore the source of the heterogeneity we conducted a meta-regression and found that the differences of races of participants had a statistically significant effect on the diagnostic accuracy, and the differences of assay types was at margin of statistically significance (p=0.075). So we surmised that race and assay type might play major roles in the diagnostic accuracy. When we combined five studies (Hippo et al., 2004; Liu et al., 2010; Tangkijvanich et al., 2010; Qiao et al., 2011; Lee et al., 2014), all of which used Elisa for GPC3 test for Asians, together for analysis of heterogeneity, I2 of sensitivity and specificity for GPC3 was 0 and 92.72, respectively. Although it was still large, it became smaller when comparing to the former I2 of specificity (96.97). The pooled sensitivity and specificity for GPC3 were 52% (95% CI, 47-56%) and 93% (95% CI, 80-98%).

**Publication bias**

We used Deeks’ funnel plot asymmetry test to evaluate the potential publication bias among included studies. The slope coefficient was correlated to p values of 0.53 and 0.50, suggesting symmetry in the data and a lesser likelihood of publication bias (Fig. 4).

**Discussion**

GPC3, a 60kDa cell-surface protein, which is a member of the heparan sulfate proteoglycan family (GPC1 to GPC6), can be sliced by furin between Arg358 and Ser359 to make a 40-kd amino (N) terminal protein and a 30-kd, membrane-bound carboxyl (C) terminal protein. NH2-terminal portion [soluble GPC3(sGPC3)] can be specifically detected in the sera of patients with HCC. GPC3 has been reported to be increased in HCC at both mRNA and protein levels in comparison with cirrhotic tissues and pre-neoplastic lesions (Hippo et al., 2004). Interestingly, GPC3 mRNA levels are more frequently elevated than those of AFP, with the difference even greater in small HCC (Capurro et al., 2003). Meanwhile there are high expressions of GPC3 in liposarcoma (52%), grade 3 cervical intraepithelial neoplasia (41%), malignant melanoma (29%) (Baumhoer et al., 2008) and 13.5% (28/207) of lung cancer patients, 13.2% (9/68) of thyroid cancer patients and 40% of melanoma patients had positive results with sGPC3 (Nakatsura et al., 2004; Chen et al., 2013). Some studies tried to explore the relationship between
serologic concentrations of GPC3 and AFP, but no correlation was found. So the simultaneous use of both markers significantly increases the sensitivity of the test (Hippo et al., 2004; Tangkijvanich et al., 2010).

In this meta-analysis we identified fourteen studies that directly compared the diagnostic accuracy of serum GPC3 with AFP in same patient population. A significant heterogeneity had been found among the included studies, and Meta-regression result suggested that race and assay type were potentially responsible for the heterogeneity; therefore a subgroup was established to test the two factors, and we found that I2 of sensitivity for GPC3 in the five studies (Hippo et al., 2004; Liu et al., 2010; Tangkijvanich et al., 2010; Qiao et al., 2011; Lee et al., 2014), whose patients were Asians and using Elisa approach, was 0%; meanwhile, I2 of specificity for GPC3 was decreased. The pooled sensitivity and specificity for GPC3 was 0.52 and 0.93. Due to only five studies included in the subgroup, more studies would be needed to further confirm the impacts of race and assay type in the diagnostic accuracy of GPC3 for HCC.

HCC patients often have no obvious discomfort in the early period of HCC, but once the symptoms become obvious, they would already be in the advanced stage with expected short survival time of approximately six months. Therefore, early detection and diagnosis is critical to improve the treatment effects of HCC and therefore patients’ survival rate. Usually a single marker inevitably leads to a false negative result for early diagnosis of HCC, and combined detection by using multi-markers may reduce missed diagnosis (Bertino et al., 2012). So it has been advocated to simultaneously use GPC3 and AFP. Our meta-analysis found that when using AFP alone for diagnosing HCC, the sensitivity was 60%. While the combination of GPC3 and AFP yielded an improved sensitivity for detecting HCC to 80%. And the sensitivity was 86%, which remained a high level (Fig. 2C).

FIGURE 3. Summary receiver-operating characteristic curves for GPC3 and AFP from the hierarchical summary receiver operating characteristic model.
The results of this study should be interpreted with caution due to there being several limitations. First, the quality of the included studies was relatively poor. And second, we included only English language publications in this meta-analysis, which probably caused a potential bias related to the human populations studied.

Conclusion
Serum GPC3 is a potentially diagnostic marker for HCC, and combination of GPC3 and AFP significantly elevate the sensitivity for early diagnosis. Since the racial origin of the participants and the assay types may impact the diagnostic outcomes of HCC, more studies for specific racial groups (Asians, Caucasians or Africans), and using certain methods for detecting GPC3 (like Elisa) are required to further confirm the diagnostic value of GPC3 for HCC.

Acknowledgements
We thank Dr. Liu Liu (The Chinese University of Hong Kong) for English language support. This study was supported by a Financial Science and Technology Special Competitive Allocation Project of Zhanjiang (No. 2013A01008). The authors have declared no conflict of interest.

References
Abdelgawad IA, Mossallam GI, Radwan NH, Elzawahry HM, Elhfnawy NM (2013). Can Glypican3 be diagnostic for early hepatocellular carcinoma?


