The Added Value of Circulating Tumor Cell Enumeration to Standard Markers in Assessing Prognosis in a Metastatic Castration-Resistant Prostate Cancer Population

Glenn Heller1, Karim Fizazi2-3, Robert McCormack4, Arturo Molina5, David MacLean6, Iain J. Webb6, Fred Saad7, Johann S. de Bono8, and Howard I. Scher9,10

Abstract

Purpose: Metastatic castration-resistant prostate cancer (mCRPC) is a heterogeneous disease for which better prognostic models for survival are needed. We examined the added value of circulating tumor cell (CTC) enumeration relative to common prognostic laboratory measures from patients with CRPC.

Methods: Utility of CTC enumeration as a baseline and post-baseline prognostic biomarker was examined using data from two prospective randomized registration-directed trials (COU-AA-301 and ELM-PC4) within statistical models used to estimate risk for survival. Discrimination and calibration were used to measure model predictive accuracy and the added value for CTC enumeration in the context of a Cox model containing albumin, lactate dehydrogenase (LDH), PSA, hemoglobin, and alkaline phosphatase (ALK). Discrimination quantifies how accurately a risk model predicts short-term versus long-term survivors. Calibration measures the closeness of actual survival time to the predicted survival time.

Results: Adding CTC enumeration to a model containing albumin, LDH, PSA, hemoglobin, and ALK (“ALPHA”) improved its discriminatory power. The weighted c-index for ALPHA without CTCs was 0.72 (SE, 0.02) versus 0.75 (SE, 0.02) for ALPHA + CTCs. The increase in discrimination was restricted to the lower-risk cohort. In terms of calibration, adding CTCs produced a more accurate model-based prediction of patient survival. The absolute prediction error for ALPHA was 3.95 months (SE, 0.28) versus 3.75 months (SE, 0.22) for ALPHA + CTCs.

Conclusions: Addition of CTC enumeration to standard measures provides more accurate assessment of patient risk in terms of baseline and postbaseline prognosis in the mCRPC population. Clin Cancer Res. 23(8); 1967–73. ©2016 AACR.
Translational Relevance

Previous studies in multiple tumor types have shown that pre- and posttreatment circulating tumor cell (CTC) number measured with the CellSearch assay is prognostic for survival. Use of CTC enumeration in practice has been limited, in part due to costs, variations in reimbursement by third-party payers, and questions about the prognostic utility. In this study, we explored in greater depth the magnitude of the added value of CTCs in terms of its predictive accuracy for survival time. We found that CTC enumeration measured at baseline and early in the treatment phase provided incremental value in predictive accuracy relative to known biomarkers acquired in the course of routine clinical practice. These findings and those in previous studies suggest a role for posttreatment CTC-containing biomarkers as an indicator of patient risk.

Materials and Methods

Study design and participants

A proportional hazards model for survival was developed using baseline and posttreatment data from patients enrolled in the completed phase III registration trials of abiraterone acetate plus prednisone (COU-AA-301) and orteronel plus prednisone (ELM-PC4). The results of the trials have been reported previously (7–9). The biomarkers used for clinical prognostication were albumin, lactate dehydrogenase (LDH), PSA, hemoglobin, and alkaline phosphatase (ALK), which have previously been shown to be prognostic for survival in multivariate analysis and are components of several prediction models (nomograms) that estimate survival times in men with mCRPC (10–14). The baseline values of albumin, LDH, PSA, hemoglobin, and ALK, along with the increase or decrease in PSA at week 13 relative to baseline, represent our submodel (called "ALPHA" in this article). For the ELM-PC4 analysis, a weighted proportional hazards model was used to account for the nonproportionality in the model. The weights enable the interpretation of the model coefficients as the average hazard ratio over time. All analyses were based on a landmark time of 12 weeks.

The objective of this study was to determine the incremental information provided by early posttreatment CTC measures in predicting patient survival. For each patient, a risk score was computed from both the ALPHA model and the model developed by adding baseline CTC enumeration and the increase or decrease in CTCs at week 13 relative to baseline. This model is referred to as ALPHA + CTC. Adding CTCs to the submodel indicates that both the baseline CTC values and the relative change in CTCs were included. For the CTC and PSA markers, the relative change was defined as:

\[
\text{marker value at week 13} - \text{marker value at baseline} \\
\text{marker value at baseline}
\]

Because patients may not have detectable CTCs at baseline, CTC baseline values equal to 0 were recorded as 1 for the CTC relative change variable. The risk score for each patient is a weighted sum of his biomarkers in the proportional hazards model, where the weights are the regression coefficients derived from the model.

Statistical analysis methods

Discrimination represents the model's strength in differentiating long-term survivors from short-term survivors. It is illustrated graphically using negative and positive predictive value statistics (15). The negative predictive value is depicted with a Kaplan–Meier curve estimating the probability of survival using the cohort of patients with low-risk scores derived from the proportional hazards models. The positive predictive value is defined as the probability of dying (one minus Kaplan–Meier) for patients with model-based high-risk scores. For the current analysis, the negative predictive value is computed using patients with the lowest 25% of the risk scores (presumed best prognosis), and the positive predictive value is calculated using patients with the highest 25% of the risk scores (presumed worst prognosis). To assess the added value of CTC enumeration, the negative and positive predictive values were computed for both models. An enhanced negative predictive value with CTCs would be indicated by a Kaplan–Meier survival curve that is above the survival curve derived from the model developed without the CTC marker, whereas an improvement in the positive predictive value is represented by a higher one minus Kaplan–Meier curve for the model that included CTCs.

In addition to a graphical analysis, the weighted concordance index (c-index) was computed to summarize the discriminatory power of each model to evaluate the added value of CTCs (16). The concordance index is the proportion of all pairs of patients, where the patient with a longer survival time also has a smaller risk score. The weighted c-index ranges between 0.5 and 1.0, with the value 0.5 indicating that the model cannot discriminate between long-term survivors and short-term survivors and a value of 1.0 indicates perfect discrimination. A 10-fold cross-validation procedure was used to compute the weighted c-indices and bootstrap 95% confidence intervals were computed for the associated parameter.

A calibration metric is used to compare the model-based predicted survival time with the actual survival time of a patient (17). To gauge calibration for each model, the relationship between the observed time to death (for patients who died) and the median predicted survival time is illustrated graphically. Perfect calibration is represented by a 45° line through the origin showing that the observed times and the predicted median survival times were equal for all patients.

A summary measure of calibration, termed the absolute prediction error (APE), is computed (18). The APE computes a weighted difference between the actual survival time and a
Figure 1. Consort diagrams for the COU-AA-301 (A) and ELM-PC4 (B) studies.

predicted median survival time (19). The patient’s predicted survival time is computed from his model-based risk score. An APE equal to zero indicates that the model-based predicted survival time is exactly equal to the true survival for all patients, which equates to the data lining up in a 45° line in the calibration plot described above. As the disparity between the observed and predicted survival increases, the APE increases.

Graphs and summary statistics of discrimination and calibration were computed. We evaluated the added value of CTC enumeration as a clinical predictor of survival. The results were validated using a patient cohort from the independent randomized clinical ELM-PC4 trial, which was designed to test the efficacy of orteronel plus prednisone in men with mCRPC.

Results

The randomized clinical trial comparing abiraterone acetate plus prednisone versus prednisone alone for patients with mCRPC enrolled 1,195 patients. Of these, 949 had all baseline plus prednisone versus prednisone alone for patients with mCRPC enrolled 1,195 patients. Of these, 949 had all baseline markers recorded and 648 patients were alive with CTC and PSA

Table 1. Summary of association analysis in the COU-AA-301 and ELM-PC4a studies.

<table>
<thead>
<tr>
<th>Factor</th>
<th>COU-AA-301</th>
<th>ELM-PC4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>SE (coefficient)</td>
</tr>
<tr>
<td>Baseline albumin</td>
<td>-0.678</td>
<td>0.171</td>
</tr>
<tr>
<td>Baseline log (ALK)</td>
<td>-0.031</td>
<td>0.076</td>
</tr>
<tr>
<td>Baseline log (CTC)</td>
<td>0.253</td>
<td>0.043</td>
</tr>
<tr>
<td>Baseline hemoglobin</td>
<td>-0.008</td>
<td>0.041</td>
</tr>
<tr>
<td>Baseline log (LDH)</td>
<td>1.161</td>
<td>0.155</td>
</tr>
<tr>
<td>Baseline log (PSA)</td>
<td>0.050</td>
<td>0.037</td>
</tr>
<tr>
<td>Relative change in CTC</td>
<td>0.098</td>
<td>0.014</td>
</tr>
<tr>
<td>Relative change in PSA</td>
<td>0.222</td>
<td>0.029</td>
</tr>
</tbody>
</table>

*3.95* 3.83

Table 2. Predictive accuracy summary measures from the COU-AA-301 and ELM-PC4 studies.

<table>
<thead>
<tr>
<th>Weighted c-index</th>
<th>APE</th>
</tr>
</thead>
<tbody>
<tr>
<td>COU-AA-301</td>
<td></td>
</tr>
<tr>
<td>ELM-PC4</td>
<td></td>
</tr>
</tbody>
</table>

A deeper analysis, presented visually, demonstrated where CTC enumeration provides added value as a prognostic marker of survival. A visual assessment of the improvement in markers recorded at week 13 and were eligible for the landmark analysis. In addition, four patients had data values with extreme leverage points that had an overly influential effect on the analysis. One patient had a PSA relative change (from baseline to week 13) approximately equal to 100, two patients had CTC relative changes equal to 99, and the final patient had a CTC relative change approximately equal to 23. With these additional four patients excluded, data from 644 patients were used in the analysis (CONSORT diagram; Fig. 1A). The median survival time for the 644 patients under study was 16.8 months (95% CI, 15.6–18.0). A summary of the marker values for these 644 patients is provided in Table 1.

Table 2 provides a summary of the log-relative risk coefficients from the proportional hazards model and their attendant standard errors and P-values. In addition to the significance of many of the factors, the baseline and relative change in CTC enumeration demonstrated a strong association with survival time. However, association alone is not a sufficient measure of the prognostic utility of individual markers.

We examined the weighted c-index to numerically evaluate the value of adding CTCs to the model in terms of discrimination. The discriminatory power of ALPHA + CTC was 0.75 (SE, 0.02). The weighted c-index for the model containing ALPHA (without CTCs) was 0.72 (SE, 0.02; Table 3). A 95% bootstrap confidence interval for the difference in these two measures, indicating the magnitude of the improvement in discrimination due to the addition of CTCs, is (0.02, 0.05).

A deeper analysis, presented visually, demonstrated where CTC enumeration provides added value as a prognostic marker of survival. A visual assessment of the improvement in
discrimination, based on the addition of CTC, was obtained by graphing the negative and positive predictive values. For the negative predictive value (Fig. 2A), patients with low risk scores are expected to have prolonged survival times relative to the entire cohort. This improvement in survival is magnified in the model that includes CTCs. For the positive predictive value, which reflects patients with the poorest prognosis, there is no benefit to adding CTCs to the prognostic model (Fig. 2B). Thus, in terms of discrimination, the prognostic utility of CTC enumeration is manifested in the lower-risk cohort; the existing markers are sufficient for prognosis with the high-risk cohort.

Taken together, the graphical and numerical discrimination analyses show that the addition of CTC enumeration improves the discriminatory power of the risk model relative to standard prognostic factors in this patient population and that the discrimination benefit is found in the lower-risk cohort.

Calibration measures the closeness of the actual survival time to the model-based predicted survival. The APE is recorded to summarize the distance between the observed survival times and predicted median survival times. The APE for ALPHA + CTC is 3.75 months (SE, 0.22). Removing CTCs from this model (ALPHA) increases the APE to 3.95 (SE, 0.28). Therefore, using APE, the inclusion of CTCs to the model provides a modest improvement in the accuracy of the predicted survival time (Table 3). The reduction in the APE due to CTC is 0.20 and the 95% confidence for this difference is (−0.02, 0.51).

The graphical calibration analysis compared the estimated median survival times to the observed survival times using smoothed curves for an uncluttered visual of this relationship (Fig. 2C). The frequencies for the observed survival times are illustrated by the vertical lines at the bottom of the plot. As shown, the calibration curve for ALPHA + CTC is closer to the 45° line early in the follow-up period, but there remains significant distance between this curve and exact calibration in the early time period.

Data from 908 patients enrolled on the ELM-PC4 study were used to validate the discrimination and calibration analyses (Fig. 1B and Table 2). In this patient population, the discrimination metric for ALPHA + CTC produced a weighted c-index of 0.75 (SE, 0.01). When removing CTCs from the model, the weighted c-index decreased to 0.71 (SE, 0.01; Table 3). A 95% confidence for the difference in the two measures is (0.02, 0.64). A similar pattern occurred with the calibration metric. The APE for ALPHA + CTC was 3.56 months (SE, 0.24) and increased to 3.83 months (SE, 0.22) when CTCs were omitted from the model. The 95% confidence for this difference is (0.02, 0.66). The negative (Fig. 3A) and positive (Fig. 3B) predictive value curves and the calibration curves (Fig. 3C) produce parallel information to these summary measures. These results validate the discrimination and calibration analyses obtained from the COU-AA-301 set.

patients would be shown by an increase in the Kaplan–Meier estimate. An improvement in the positive predictive value among these high-risk patients would be demonstrated by an increase in the one minus Kaplan–Meier estimate. Calibration represents the relationship between observed survival time and the median predicted survival time. A model with good calibration would closely approximate the 45° line through the origin.
Discussion

Understanding the clinical importance of a posttherapy measure to assess long-term trial endpoints is an essential step in establishing outcome measures as indicators of clinical benefit that can be used to support drug approval. To do so requires a determination of the prognosis of the patient cohort using risk factors measured at baseline and postbaseline. This study shows that the incorporation of the CTC number at baseline and the relative change in CTC number from baseline to week 13 provides an improvement in the predictive accuracy of a prognostic model for low-risk patients with respect to survival time. The results were validated using patient data from a comparable mCRPC trial, but with chemotherapy-naïve patients, in contrast to the initial study based on patients with prior docetaxel treatment.

The discrimination analysis in the COU-AA-301 marker data demonstrated that, for defining a low-risk cohort, the addition of CTC number to the risk model produced higher survival rates relative to a risk model developed without CTC enumeration. Thus, the addition of CTC enumeration to the submodel improved the negative predictive value of the risk classification model. The finding was validated using an independent cohort of patients treated in the ELM-PC4 trial, where the addition to CTCs to the submodel showed even greater separation in the survival curves among the low-risk cohort developed with and without CTCs.

The calibration analysis established that the CTC-based model provided only a modest improvement in predicting survival time. Neither curve approaches the 45° line of equivalence between the model and actual survival time. There are two components that impact the accuracy of a calibration analysis. First, point prediction of survival time is complicated by many factors not related to disease, such as age or comorbidities. Second, there are disease-related factors, such as number of bone metastases, performance status, and Brief Pain Inventory-Short Form score, to name a few, that are not included in the model. In our analysis, this is evident by the calibration curves in Fig. 3: the prediction of early deaths is poor, as shown by the distance or separation from the 45° line.

Complete biomarker data were not available for all patients. For patients who are alive at 13 weeks and missing either the week 13 CTC or PSA recording, the data are not missing at random. For the COU-AA-301 study, the median survival time for these patients is 13.4 months compared to a median survival time of 16.8 for the analysis cohort. However, the potential biases in the discrimination or calibration comparisons should approximately cancel because the same patient cohort was used for both risk models (ALPHA and ALPHA + CTC).

The use of biomarkers during follow-up to accurately determine prognosis is essential for disease management. To this end, serial biologic profiling of the disease before treatment, during treatment, and at progression, and determining the

Figure 3.

ELM-PC4 negative (A) and positive (B) predictive value and calibration (C) curves. Note: Negative predictive value includes patients in the lowest quartile of risk scores within each model; positive predictive value includes patients in the highest quartile of risk scores within each model; calibration represents the relationship between observed survival time and the median predicted survival time.
association of the profiles with later events such as radiographic progression-free survival and/or overall survival, was included in the updated recommendations from the Prostate Cancer Clinical Trials Working Group 3 (20). If applied early in the posttreatment period, the biomarker results may be useful for informing the decision to continue treatment if the patient-specific risk score is favorable or to discontinue treatment if the computed risk score is not.

In this study, we found that CTC enumeration measured at baseline and early in the treatment phase, regardless of treatment received, provided incremental value to the clinical factors and laboratory test results acquired in the course of routine clinical practice. A limitation of the CTC CellSearch assay is that it only defines one circulating tumor cell type. Whether non-selection based assays that enable the identification of multiple cell types is more informative is unknown. At this point, however, the results support a role for postbaseline CTC-containing biomarkers as an indicator of prognosis. More research is needed to go beyond its prognostic utility and examine its clinical utility as an intermediate response variable.

Disclosure of Potential Conflicts of Interest
K. Fizazi is a consultant/advisory board member for Amgen, Astellas, AstraZeneca, Bayer, Essa, Genentech, Janssen, Orion, and Sanofi. D. MacLean is an employee of Takeda Pharmaceuticals. F. Saad is a consultant/advisory board member for Astellas, Janssen, and Takeda. J.S. de Bono is a consultant/advisory board member for AstraZeneca, Genentech, Genmab, GlaxoSmithKline, Janssen, Pfizer, Sanoﬁ, and Vertex. H.I. Scher reports receiving commercial research grants from and is a consultant/advisory board member for Janssen. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions
Conception and design: G. Heller, R. McCormack, A. Molina, J.S. de Bono, H.I. Scher
Development of methodology: G. Heller, R. McCormack, J.S. de Bono
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): K. Fizazi, A. Molina, D. MacLean, F. Saad, J.S. de Bono, H.I. Scher
Analysis and interpretation of data (e.g., statistical analysis, bio-statistics, computational analysis): G. Heller, K. Fizazi, J.I. Webb, F. Saad, J.S. de Bono
Writing, review, and/or revision of the manuscript: G. Heller, K. Fizazi, R. McCormack, A. Molina, D. MacLean, J.I. Webb, F. Saad, J.S. de Bono, H.I. Scher
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A. Molina
Study supervision: D. MacLean, F. Saad, J.S. de Bono

Acknowledgments
Editorial assistance was provided by Lashon Pringle, PhD, of PAREXEL, and was funded by Janssen Global Services, LLC.

Grant Support
These analyses were funded by the following sources: NIH/NCI P50-CA92629SPOR in Prostate Cancer, NIH/NCI Cancer Center Support Grant P30 CA08748, Department of Defense Prostate Cancer Research Program (PC071610 and PC121111), Prostate Cancer Foundation Challenge Award, David H. Koch Fund for Prostate Cancer Research.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 13, 2016; revised August 29, 2016; accepted September 8, 2016.

References
The Added Value of Circulating Tumor Cell Enumeration to Standard Markers in Assessing Prognosis in a Metastatic Castration-Resistant Prostate Cancer Population

Glenn Heller, Karim Fizazi, Robert McCormack, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-16-1224

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2016/09/27/1078-0432.CCR-16-1224.DC1

Cited articles
This article cites 16 articles, 7 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/23/8/1967.full#ref-list-1

Citing articles
This article has been cited by 3 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/23/8/1967.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://clincancerres.aacrjournals.org/content/23/8/1967.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.