Quantitative measurement of adiposity using CT images to predict the benefit of bevacizumab-based chemotherapy in epithelial ovarian cancer patients

YUNZHI WANG\textsuperscript{1}, THERESA THAI\textsuperscript{2}, KATHLEEN MOORE\textsuperscript{2}, KAI DING\textsuperscript{2}, SCOTT McMEEKIN\textsuperscript{2}, HONG LIU\textsuperscript{1} and BIN ZHENG\textsuperscript{1}

\textsuperscript{1}School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK 73019; \textsuperscript{2}Health Science Center of University of Oklahoma, Oklahoma, OK 73104, USA

Received July 8, 2015; Accepted May 12, 2016

DOI: 10.3892/ol.2016.4648

Abstract. The present study aims to quantitatively measure adiposity-related image features and to test the feasibility of applying multivariate statistical data analysis-based prediction models to generate a novel clinical marker and predict the benefit of epithelial ovarian cancer (EOC) patients with and without maintenance bevacizumab-based chemotherapy. A dataset involving computed tomography (CT) images acquired from 59 patients diagnosed with advanced EOC was retrospectively collected. Among them, 32 patients received maintenance bevacizumab following primary chemotherapy, while 27 did not. A computer-aided detection scheme was developed to automatically segment visceral and subcutaneous fat areas depicted on CT images of abdominal sections, and 7 adiposity-related image features were computed. Upon combining these features with the measured body mass index, multivariate data analyses were performed using three statistical models (multiple linear, logistic and Cox proportional hazards regressions) to analyze the association between the model-generated prediction results and the treatment outcome, including progression-free survival (PFS) and overall survival (OS) of the patients. The results demonstrated that applying all three prediction models yielded a significant association between the adiposity-related image features and patients’ PFS or OS in the group of the patients who received maintenance bevacizumab (P<0.010), while there was no significant difference when these prediction models were applied to predict both PFS and OS in the group of patients that did not receive maintenance bevacizumab. Therefore, the present study demonstrated that the use of a quantitative adiposity-related image feature-based statistical model may generate a novel clinical marker to predict who will benefit among EOC patients receiving maintenance bevacizumab-based chemotherapy.

Introduction

Ovarian cancer has the highest mortality rate of all gynecologic malignancies in the USA (1). The majority of ovarian cancer cases (>85%) are epithelial ovarian cancer (EOC) (2). Previous studies have shown that angiogenesis plays a fundamental role in the pathogenesis of EOC, with high vascular endothelial growth factor (VEGF) expression, which promotes tumor growth, ascites and metastases (3). Thus, novel therapies that target the angiogenesis-specific pathways have been developed and tested in clinical trials (4–6). Among them, bevacizumab, a 149-kDa recombinant humanized monoclonal immunoglobulin G1 anti-VEGF antibody, has been extensively tested in several randomized clinical phase III trials (4,5). The results of those studies revealed that certain groups of EOC patients, such as those with high risk of disease progression, may benefit from bevacizumab therapy by increasing their progression-free survival (PFS) and/or overall survival (OS) (4), while other groups of EOC patients may not be suitable for receiving bevacizumab-based chemotherapy due to its high toxicity levels and other serious adverse events (6). For example, a fraction of EOC patients with obesity usually have poor response to bevacizumab-based chemotherapy, including short PFS and OS (7). As a result, how to rationally identify EOC patients who are most likely to benefit from receiving bevacizumab and/or other antiangiogenic therapies remains one of the unsolved critical issues in treating EOC patients (8).

In order to identify effective clinical biomarkers to accurately predict the benefit of EOC patients receiving concurrence and/or maintenance bevacizumab-based chemotherapy, the present authors previously investigated an adiposity-based clinical biomarker based on the fact that obesity is not only common in a large fraction of patients with advanced EOC, but it is also highly associated with increased circulating levels of VEGF, a key target of bevacizumab (9). In that study, the
visceral fat area (VFA) and subcutaneous fat area (SFA) of an EOC patient were measured by a radiologist from one computed tomography (CT) image slice acquired in the umbilical section of the patient (9). Applying that technique to a group of 25 EOC patients who received bevacizumab demonstrated that EOC patients with lower levels of VFA and/or SFA were likely to yield a significantly longer PFS (P=0.030) and probably a longer OS (P=0.060) than those with high levels of VFA and/or SFA (9).

In order to effectively and reliably extract adiposity-related image features and to build a multivariate statistical prediction model to assess the clinical benefit of treating individual EOC patients with bevacizumab-based chemotherapy, the present study was performed based on the hypothesis that the perfusion CT images of the EOC patients contain various useful image features that could be quantitatively extracted and computed to describe the extent and nature of adiposity in detail. Using an automated quantitative image feature analysis method may also provide more accurate and reliable results to predict clinical benefit by avoiding the inter-reader variability observed in the previous study (9). Therefore, the objective of the present study is to test the aforementioned hypothesis by first developing a computer-aided detection (CAD) scheme to automatically segment VFA and SFA from the multiple CT image slices acquired in the abdominal section and compute both fat area and density distribution-related image features. Then, using these quantitative image features and combining them with the measured body mass index (BMI), three multivariate statistical data analysis models were built and tested to investigate and cross-validate the association between the computed adiposity-related image features and the clinical outcome (PFS and OS) of EOC patients receiving bevacizumab-based chemotherapy.

Materials and methods

Based on the present data collection and study protocol approved by the Institutional Review Board of the University of Oklahoma (Norman, OK, USA; approval no. #4168), a patient dataset was retrospectively assembled for the present preliminary study, which included 59 patients diagnosed and treated with advanced EOC (stages III and IV) in the Health Science Center of the University of Oklahoma (Oklahoma, OK, USA) from January 2005 to January 2010. For each patient, a set of post-contrast-enhanced perfusion CT images, which were acquired following primary cytoreductive surgery but prior to chemotherapy initiation, were collected. All these CT images were scanned using a LightSpeed VCT 64-detector (GE Healthcare Life Sciences, Chalfont, UK) or a Discovery 600 16-detector (GE Healthcare Life Sciences) CT machine. The X-ray power output was set at 120 kVp and a variable range from 100 to 600 mA depending on patient body size. Prior to CT image scanning, 100 cc contrast agent Isovue-370 (Bracco Imaging, Milan, Italy) was intravenously injected using a standard power injector at a rate of 2-3 cc/sec.

In the above dataset, the 59 patients were divided into two groups. The first group included 32 patients who received chemotherapy consisting of bevacizumab (175 mg/kg) plus paclitaxel (175 mg/m²) and carboplatin [target area under the concentration versus time curve=6] with follow-up maintenance bevacizumab. The PFS and OS data for all of these patients were also available and collected in the dataset. The patient characteristics, including demographic information, BMI and clinical outcome of these two groups of patients, are presented in Table I. With the exception of PFS, there is no significant difference between these two groups of patients regarding all other characteristics.

From each patient, a perfusion CT examination was performed to scan from the lung to the pelvis, crossing the entire abdominal region. A CAD scheme was developed and applied to automatically process multiple CT image slices across the entire abdominal region from the CT image slice just below the lung area to the slice just above the umbilicus level. The CAD scheme segments VFA and SFA on the selected abdominal CT image slices using the following steps: i) CAD is applied to segment a body trunk region from the air and CT bed background in each CT image slice using a previously developed and tested automated CT image segmentation method (10); ii) since a previous study has defined that a pixel in the fat area depicted on perfusion CT images should have a CT number between -140 HU and -40 HU (11), CAD applies these two threshold values to the segmented body region within the body trunk mask in the CT image and thus generates two novel masks to cover the fat and non-fat regions; iii) in order to differentiate between the visceral and subcutaneous regions, CAD generates a visceral region mask by applying several image processing processes to the non-fat region mask (12,13). A four-connect pixel labelling method is then performed on the non-fat region mask, and the connected regions with sizes smaller than a predefined threshold (200) are removed. This process removes all the small and isolated pixels located in the subcutaneous fat region. Then, a morphological dilation operation with a spherical kernel (8 pixels) is applied to the image (non-fat region), which aims to break the possible connection between SFA and VFA in certain CT image slices. Subsequently, a visceral region mask is created by performing a hole-filling process to cover all non-fat organs or structures, followed by a morphological erosion operation; and iv) CAD finally defines a VFA mask, by performing an ‘AND’ operator between the fat region mask and the visceral region mask, and a SFA mask, by performing an ‘AND’ operator between the fat region mask and the body trunk mask. In this way, all the pixels in the fat region mask are classified as either subcutaneous fat pixels or visceral fat pixels.

These four CAD processing steps are iteratively applied to all the CT image slices in the selected abdominal section of a patient. Fig. 1 represents several examples of segmentation results of VFA and SFA depicted on several CT image slices. Upon segmentation of VFA and SFA, the CAD scheme computes 7 image features, as indicated in Table II, which include i) the percentage or ratio of two fat areas or volumes (VFA and SFA) compared with the whole body volume computed from all CAD-processed CT image slices; ii) the mean and standard deviation (SD) of the CT number (pixel value) of the VFA and SFA; and iii) the ratio between the segmented volume between SFA and VFA.
Upon extracting 8 features, the next step was to investigate whether these features and their combinations are significantly associated with the clinical outcome (PFS and OS) of the EOC patients in the two groups (with or without maintenance bevacizumab). For this purpose, three multivariate statistical data analysis methods commonly used in medical data analysis literature (14) were applied, including multiple linear regression, multiple logistic regression and Cox proportional hazards regression, in order to evaluate the association between the BMI and CT image features and the clinical outcome of EOC patients. The purpose of using all these three models is to cross-validate the present study results and to improve their robustness.

A multiple linear regression model is a generalization of a simple linear regression model, which is used to assess the linear association between the outcome variable and ≥2 predictor variables (features) based on certain mutually independent observations (15), which can be expressed as:

\[ y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon \]

where \( y \) is the outcome variable, \( x_1, x_n \) are \( n \) distinct predictor variables, \( \varepsilon \) is the random error variable with zero-mean Gaussian distribution and \( \beta_0, \beta_n \) are estimated regression coefficients obtained by minimizing the variation of \( \varepsilon \). \( \beta_0 \) is the intercept of the regression model, while \( \beta_n \) (\( 1 \leq k \leq n \)) represents the changes in \( y \) relative to one unit change in \( x_k \).

Since the present study involves two groups of patients and two clinical outcomes (PFS and OS), four multiple linear regression models were built. Each model was used to assess the association between one clinical outcome and the extracted features in one group of EOC patients. To identify the optimal feature sets selected from 8 features, an exhaustive-search method was applied to select a specific feature set with the highest predicting power in one evaluation model. Specifically, all the possible combinations of 8 features were used as the predictor variables to establish the regression model, and the regression model with the highest \( R^2 \) value was considered as the best prediction of the outcomes. An F-test (also known as Wald test) was performed to evaluate the overall significance of the model.

A logistic regression is a regression-based statistical classification model (16) that is able to predict binary outcomes from ≥2 continuous predictor variables, where the binary outcomes indicate which class the observations belong to. The formula of basic logistic regression can be written as:

\[ F(x) = \frac{1}{1 + e^{-\sum_{i=0}^{n} \beta_i x_i}} \]

where the output variable \( F(x) \) is a continuous value, indicating the probability of the outcome belonging to class 1. Logistic regression is a ‘linear’ classifier because the output of the model is determined by a linear combination of the predictor variables, as indicated by the above formula.

In the present study, similarly to the multiple linear regression process, four multiple logistic regression classifiers were built. In each patient group, PFS and OS were binarized using the median PFS and OS value as the threshold to create two classes of ‘long’ and ‘short’ survival (for both PFS and OS). A multiple logistic regression model was performed to distinguish the two classes based on the predictor variables. An exhaustive-search feature subset selection process was also performed to select the specific feature set with the highest discriminatory power. To minimize the model training bias, performance of a classifier (model) was trained and tested using a leave-one-case-out validation method (17). The classification accuracy was compared to a binomial null distribution \( B(\eta, q) \) in order to assess the statistical significance of classifier over chance level, where \( \eta \) is the total number of observations and \( q \) is the prediction accuracy of chance level (0.5). The classification scores were processed by a maximum likelihood data analysis-based receiver operating characteristic (ROC) curve fitting program (ROCKT; http://metz-roc.uchicago.edu/MetzROC/software). The area under the ROC curve was used as performance assessment index to assess model performance.
The Cox proportional hazards regression is a statistical model commonly used in survival analysis, which investigates the association between the time that passes prior to the occurrence of a particular event (such as mortality) and several predictor variables (18). The basic formula of a Cox regression model can be written as:

\[ h(t) = h_0(t) e^{\sum_{i=1}^{n} \beta_i x_i} \]

where \( h(t) \) is a hazard function assessing the instantaneous probability of the event occurring at time \( t \), \( h_0(t) \) is the baseline hazard function (which is unspecified in the model and is estimated by the method of partial likelihood from the observations) and \( \beta_1, \beta_n \) are the optimal coefficients obtained by fitting the model using the observations.

In the present study, four Cox regression models were built to assess the association between PFS/OS and the extracted features in the two patient groups. In each model, all 8 features were used as predictor variables. The overall significance of the model was evaluated using a likelihood-ratio test. The null hypothesis was that all the coefficients \( \beta_1, \beta_n \) were equal to zero and that the hazard function was mainly determined by the baseline hazard function. The significance of each individual feature was evaluated by a Wald test, with the null hypothesis that the coefficient of that feature was equal to zero.

**Results**

Table III summarizes the association between the model-generated prediction results obtained using the multiple linear regression models and the EOC patient outcomes (PFS or OS) in the two patient groups. The results indicate that neither PFS nor OS are significantly associated with the extracted features or models in the group of patients not receiving maintenance bevacizumab (\( P>0.117 \)). However, in the group of patients who received maintenance bevacizumab, a significant association was detected between the model-generated prediction results and both clinical outcomes (PFS and OS) with \( P<0.003 \).

Based on the selected features used in the models, it was also observed that, in the group of patients receiving maintenance bevacizumab, PFS was best predicted using the feature set including 6 image features, while OS was best predicted using BMI and 2 other image features.

Table IV summarizes the performance of applying the multiple logistic regression models for predicting the EOC patients into two ‘long’ and ‘short’ survival (PFS or OS) classes. The results indicate that, in the group without maintenance bevacizumab, the model had no statistically significant predictive power compared with chance (\( P>0.060 \)), while in the group with maintenance bevacizumab, the model prediction performance was significantly different from chance in a binomial null
distribution (P<0.01). Fig. 2 represents a scatter diagram of two features (BMI and SD of the CT number of SFA) used in a multiple logistic regression model to predict or classify each case in the group of patients who received maintenance bevacizumab into ‘long’ and ‘short’ PFS. Based on this scatter diagram, two confusion matrices were generated, as indicated in Table V. The overall classification accuracy was 56.3% (18/32) when using BMI only. Upon adding an image feature, the classification accuracy of the prediction model increased to 75.0% (24/32).

Table IV. Performance summary of using the multiple logistic regression models to classify patients as ‘long’ or ‘short’ survival.

<table>
<thead>
<tr>
<th>Patient group</th>
<th>Clinical outcomes</th>
<th>Selected feature set</th>
<th>LOO classification accuracy</th>
<th>P-value over null hypothesis</th>
<th>Area under ROC curve</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without maintenance bevacizumab</td>
<td>PFS</td>
<td>f₁</td>
<td>16/27</td>
<td>0.124</td>
<td>0.487</td>
<td>(0.273-0.705)</td>
</tr>
<tr>
<td></td>
<td>OS</td>
<td>f₂, f₃</td>
<td>17/27</td>
<td>0.061</td>
<td>0.543</td>
<td>(0.323-0.750)</td>
</tr>
<tr>
<td>With maintenance bevacizumab</td>
<td>PFS</td>
<td>f₁, f₃, f₄, f₅, f₆, f₇</td>
<td>24/32</td>
<td>0.001</td>
<td>0.653</td>
<td>(0.449-0.820)</td>
</tr>
<tr>
<td></td>
<td>OS</td>
<td>f₁, f₃, f₅, f₇</td>
<td>22/32</td>
<td>0.010</td>
<td>0.681</td>
<td>(0.468-0.846)</td>
</tr>
</tbody>
</table>

PFS, progression-free survival; OS, overall survival; ROC, receiver operating characteristic; LOO, leave-one-out.

Fig. 3 includes three box-plots representing the distribution of PFS in three subgroups. The results demonstrate that, although in the present dataset the PFS of the group of patients not receiving maintenance bevacizumab is significantly shorter than that of the group of patients receiving maintenance bevacizumab (Table I), the subgroup of patients who received maintenance bevacizumab but was classified as short PFS class by the aforementioned multivariate logistic regression model (middle box) has no significant difference in PFS compared with the group not receiving maintenance bevacizumab (P=0.114).

When applying the Cox proportional hazards regression models, the results indicate that the model prediction results were significantly associated with both PFS (P=0.009) and OS (P=0.007) in the group of patients who received maintenance bevacizumab, according to the likelihood-ratio test. However, in the patient group not receiving maintenance bevacizumab, no significant association was detected in PFS (P=0.706) or OS (P=0.753). Using the Wald test, the Cox model also analyzed the prediction power of each individual feature. The results indicate that none of the features analyzed had significant power to predict PFS or OS in the group without maintenance bevacizumab, while in the group of patients receiving bevacizumab, one image feature (SD of the CT number of VFA) had a significant association with PFS (P=0.020), while three features had a significant association with OS, including BMI (P=0.001), SD of the CT number of SFA (P=0.002) and SD of the CT number of VFA (P=0.037).
Table V. Comparison of two confusion matrices between using BMI only and a combination of BMI and the standard deviation of the computed tomography number of subcutaneous fat area ($f_s$) in classifying patients who received maintenance bevacizumab into ‘long’ and ‘short’ PFS classes.

<table>
<thead>
<tr>
<th>PFS class</th>
<th>BMI</th>
<th>BMI and $f_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>Long</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Short</td>
<td>7</td>
<td>9</td>
</tr>
</tbody>
</table>

PFS, progression-free survival; BMI, body mass index.

Figure 3. Box plots showing the PFS of three groups of patients not receiving maintenance bevacizumab and patients receiving maintenance bevacizumab with predicted short or long PFS. PFS, progression-free survival.

Discussion

Applying antiangiogenic agents such as bevacizumab to block the VEGF signaling pathway is changing the landscape of treating numerous advanced cancers, including EOC (4,5). However, the lack of effective prognostic prediction markers to stratify the patients remains a major challenge to establish an optimal personalized treatment paradigm of using these antiangiogenic agents to date. In a continuous effort to identify more effective prognostic markers, 'adiposity' has recently attracted research interest as a potential clinical biomarker to evaluate the efficacy of applying VEGF-targeted antiangiogenic therapy to treat a variety of cancers, including colon cancer (7,19), renal cell carcinoma (20,21) and ovarian cancer (9). However, previous studies generated somehow conflicting results (7,19,20,21). In the present study, a CAD approach was developed for the first time to quantitatively extract adiposity-related image features and to demonstrate the feasibility of applying multivariate data analysis models to stratify EOC patients with and without potential clinical benefit by receiving maintenance bevacizumab-based chemotherapy.

The present study has a number of unique characteristics regarding the computation of the quantitative image features and fusion of these features to improve prediction performance. First, since VFA and SFA ratios were automatically computed without any human intervention from the multiple CT images acquired from the abdominal section, this process should provide more reliable and consistent feature values than using one subjectively selected CT image slice, as conducted in previous studies (9,11). Second, besides the size ratio of VFA and SFA, the present study also investigated and added novel image features to compute the mean CT number and SD of VFA and SFA. These novel features may correlate with the density (or probably stiffness) and heterogeneity of VFA and SFA, which have never been used and tested in previous studies. Third, using these quantitative image features and combining them with BMI data, three multivariate statistical data analysis models were built and tested to assess the potential association between these adiposity-based features and the clinical outcome of EOC patients with or without receiving maintenance bevacizumab. By cross-validation and comparison of the prediction results generated from the three models, the present study results are more robust than using individual data analysis models.

The present findings indicate that the computed adiposity image features were significantly associated with both PFS and OS in the group of EOC patients who received maintenance bevacizumab. Therefore, using a statistical data analysis model to fuse these features may provide a novel prognostic prediction marker or tool for improved prediction of the potential clinical benefit of EOC patients receiving maintenance bevacizumab-based chemotherapy. For example, using the prediction results of a multivariate logistic regression model, the patients were classified into ‘long’ and ‘short’ PFS class categories. The patients classified as ‘long’ PFS had significantly improved PFS, while the patients classified as ‘short’ PFS did not exhibit significant differences in PFS compared with patients not receiving maintenance bevacizumab (Fig. 3).

The definition of obesity itself and whether BMI is the most appropriate measure of obesity are controversial (22). In the present study, the performance in predicting PFS or OS was compared between using BMI and other quantitative image features. It was observed that adding the image features could substantially improve prediction performance. For example, by adding the image feature of SD of the CT number of SFA, the correctly classified number of cases increased >33% (from 18 to 24), compared with using BMI only (Table V). Therefore, quantitative adiposity-related image features may provide supplementary information to BMI to better measure obesity and aid to improve prediction performance when using obesity as a clinical biomarker.

Previous biological studies indicated that not all fat was equal in terms of differential gene expression, which could affect the efficacy of therapeutic targets differently (23,24). As a result, the image features based only on the size (or ratio) of VFA and/or SFA measured from one CT image slice may not be optimal. In the present study, novel quantitative image features were computed to distinguish the difference between VFA and SFA based on their density (or stiffness) and heterogeneity. The results demonstrated that these novel features were more frequently selected to build the prediction models than the VFA and SVA size ratio features. In particular, the highest discriminatory power comes from two fat density heterogeneity features (SD of the CT number of VFA and SFA) in all three prediction models. This demonstrates the potential link between the
genotype features discovered in biology or genomic studies and the quantitative CT image-based phenotype features.

In summary, the present study provides the first evidence and demonstrates that i) EOC patients with high levels of adiposity and/or high heterogeneity of fat tissues may not benefit from receiving maintenance bevacizumab-based chemotherapy; and ii) using quantitatively computed adiposity-related CT image features provides a useful and supplementary information that could yield higher discriminatory power than BMI in predicting the association between adiposity and clinical outcome of EOC patients (including PFS and OS). Despite these promising results, the present is a preliminary study with a number of limitations, including i) a small dataset size and ii) the use of CT images collected from a single medical institution. As a result, due to the small dataset size, the 95% confidence intervals of the prediction performance (the area under the ROC curve values in Table IV) are considerably wide. Therefore, further studies and cross-validation using larger and more diverse datasets are required prior to the acceptance of the adiposity prediction model and the related novel quantitative image feature-based prediction models in the clinical practice as a useful prognostic biomarker for EOC patients.

Acknowledgements

The present study was approved by the Institutional Review Board of the University of Oklahoma (Norman, OK, USA; protocol no. IRB# 4168) on April 9, 2014, and was supported in part by the Oklahoma Tobacco Settlement Endowment Trust Cancer Center Program (Oklahoma, OK, USA) and the Peggy and Charles Stephenson Cancer Center of the University of Oklahoma (Oklahoma, OK, USA; grant no. A00-3995).

References