CrossMark

A novel representation of inter-site tumour heterogeneity from pre-treatment computed tomography textures classifies ovarian cancers by clinical outcome

Hebert Alberto Vargas[1] · Harini Veeraraghavan[2] · Maura Micco[1] ·
Stephanie Nougaret[1,3,4] · Yulia Lakhman[1] · Andreas A. Meier[1] · Ramon Sosa[1] ·
Robert A. Soslow[5] · Douglas A. Levine[6] · Britta Weigelt[5] · Carol Aghajanian[7] ·
Hedvig Hricak[1] · Joseph Deasy[2] · Alexandra Snyder[7] · Evis Sala[1]

Abstract

Purpose To evaluate the associations between clinical outcomes and radiomics-derived inter-site spatial heterogeneity metrics across multiple metastatic lesions on CT in patients with high-grade serous ovarian cancer (HGSOC).

Methods IRB-approved retrospective study of 38 HGSOC patients. All sites of suspected HGSOC involvement on preoperative CT were manually segmented. Gray-level correlation matrix-based textures were computed from each tumour site, and grouped into five clusters using a Gaussian Mixture Model. Pairwise inter-site similarities were computed, generating an inter-site similarity matrix (ISM). Inter-site texture heterogeneity metrics were computed from the ISM and compared to clinical outcomes.

Results Of the 12 inter-site texture heterogeneity metrics evaluated, those capturing the differences in texture similarities across sites were associated with shorter overall survival (inter-site similarity entropy, similarity level cluster shade, and inter-site similarity level cluster prominence; $p \leq 0.05$) and incomplete surgical resection (similarity level cluster shade, inter-site similarity level cluster prominence and inter-site cluster variance; $p \leq 0.05$). Neither the total number of disease sites per patient nor the overall tumour volume per patient was associated with overall survival. Amplification of 19q12 involving *cyclin E1 gene* (*CCNE1*) predominantly occurred in patients with more heterogeneous inter-site textures.

Conclusion Quantitative metrics non-invasively capturing spatial inter-site heterogeneity may predict outcomes in patients with HGSOC.

Key Points
- *Calculating inter-site texture-based heterogeneity metrics was feasible*
- *Metrics capturing texture similarities across HGSOC sites were associated with overall survival*
- *Heterogeneity metrics were also associated with incomplete surgical resection of HGSOC.*

✉ Hebert Alberto Vargas
  vargasah@mskcc.org

1 Department of Radiology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

2 Department of Medical Physics, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

3 Service de Radiologie, Institut Régional du Cancer de Montpellier, Montpellier, France

4 INSERM, U1194, Institut de Recherche en Cancérologie de Montpellier (IRCM), Montpellier, France

5 Department of Pathology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

6 Department of Surgery, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

7 Department of Medicine, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

Abbreviations

CCNE1	Cyclin E1 gene
CLOVAR	Classification of Ovarian Cancer transcriptomic profiles
CSR	Complete surgical resection
CT	Computed tomography
EM	Expectation-Maximization algorithm
FIGO	The International Federation of Gynecology and Obstetrics
GLCM	Gray-level correlation matrix
GMM	Gaussian Mixture Model
HGSOC	High-grade serous ovarian cancer
IQR	Interquartile range
ISM	Inter-site similarity matrix
KM	Kaplan-Meier
ROI	Region of interest
SCP	Inter-site similarity level cluster prominence
SCS	Similarity level cluster shade
SE	Inter-site similarity entropy
SLV	Inter-site similarity level variance
TCGA	The Cancer Genome Atlas
VOI	Volume of interest

Introduction

Ovarian cancer is the deadliest malignancy of the female genital tract, with high-grade serous ovarian cancer (HGSOC) the most common histologic subtype. Clinical outcomes are widely variable [1]. This discrepancy cannot be explained solely by differences in clinico-pathological features or treatment regimens. Genomics play an important role in the natural history of ovarian cancer: improved survival has been reported in patients with germline BRCA-associated ovarian cancers [2] and worse outcomes in tumours with mesenchymal transcriptomic profiles [3, 4].

Tumour heterogeneity within a single tumour site (intra-tumour heterogeneity) or between different metastatic sites in the same patient (inter-site heterogeneity) is a feature of many solid malignancies, and preliminary evidence linking tumour heterogeneity and outcomes in ovarian cancer is emerging [5, 6]. A robust imaging-based method to quantify non-invasively heterogeneity within and among separate tumour sites in the same patient could improve our ability to select effective therapies and curtail treatment resistance. Texture analysis is a "radiomics" [7] method that involves extraction of metrics, termed texture parameters, that provide detailed quantitative characterization of local variations in intensity levels throughout an image [8]. To date, radiomics methods have been applied to evaluate intra-tumour heterogeneity based on a single site of disease per patient [9, 10]. The purpose of this study was to evaluate the associations between clinical outcomes and radiomics-derived inter-site spatial heterogeneity metrics across multiple metastatic lesions on CT in HGSOC patients evaluated as per The Cancer Genome Atlas (TCGA) Research Network ovarian cancer pilot project [11].

Methods

This retrospective study was compliant with the Health Insurance Portability and Accountability Act and received approval by the institutional review board with a waiver of informed consent.

Patient characteristics, image acquisition, and transcriptomic classification of ovarian cancer (CLOVAR) transcriptomic subtyping

Through our institution's electronic records we identified patients fulfilling the following inclusion criteria: (i) FIGO stage IIIC-IV high-grade serous ovarian cancer, (ii) attempted primary cytoreductive surgery (total abdominal hysterectomy, bilateral salpingo-oophorectomy, omentectomy, and extensive tumour debulking, with or without pelvic and retroperitoneal lymph node dissection) performed at our institution, (iii) standard of care CT of abdomen and pelvis performed prior to surgery and demonstrating at least three tumour sites, and (iv) molecular analysis performed as per the TCGA Research Network ovarian cancer pilot project. The only exclusion criterion was neoadjuvant chemotherapy given prior to surgery. The final study population consisted of 38 patients; their clinical characteristics are shown in Table 1. Gene expression profiles were obtained as per the publication by Verhaak et al. [4]. All patients were included in a prior study that investigated the associations between qualitative CT imaging features, CLOVAR gene signatures and survival in women with HGSOC [11]. CT examinations were performed within a median of 12 days (interquartile range [IQR] 5-21 days) before surgery using multidetector helical scanners as per the previously reported technique and acquisition parameters [11].

Image analysis and texture feature extraction

An oncologic imaging research fellow with >6 years experience in radiology evaluated all CT exams and identified all sites of suspected involvement by ovarian cancer, including the primary ovarian tumour(s) and metastatic peritoneal implants in the abdomen and pelvis (Table 2). The sites of disease were numerically coded according to a pre-defined template based on anatomic location of the metastatic deposits (for example site #2 is pelvic cul-de-sac, site #4 is omentum, with location numbers ranging from 1 to 13). For each lesion, a region of interest (ROI) was manually segmented on all CT slices where the lesion

Table 1 Clinical characteristics of the 38 patients used in the analysis

Characteristic	
FIGO Stage III	23
FIGO Stage IV	15
Overall survival	56.4 (36-75)
Complete Surgical Resection	7
Incomplete Surgical Resection (<1 cm)	23
Incomplete Surgical Resection (>1 cm)	8
CCNE1 amplification	9
CLOVAR Differentiated	9
CLOVAR Immunoreactive	9
CLOVAR Mesenchymal	16
CLOVAR Proliferative	4

Overall survival is shown as median time in months (inter-quartile range), while the remaining variables are shown as number of patients

FIGO The International Federation of Gynecology and Obstetrics; *CCNE1* cyclin E1 gene; *CLOVAR* Classification of Ovarian Cancer transcriptomic profiles

Table 2 Number and anatomical location of tumour sites from which texture metrics were extracted

Subtype	Total
No. of patients	38
Ovarian mass	36
Omentum	35
Cul de sac	28
Left upper quadrant	25
Right paracolic	25
Left paracolic	19
Diaphragm	18
Capsular liver	17
Subcapsular liver	13
Gastro-hepatic ligament	11
Mesentery	9
Lesser sac	8
Subcapsular spleen	7
Capsular spleen	4

was visible and a volume of interest (VOI) was generated, using a multi-platform, free and open source software package for visualization and medical image computing (3D Slicer, version 4.2.2-1; available at: http://slicer.org/). The manually segmented lesions provided the VOI for further computer-based image analysis.

The methodology for texture feature extraction was as follows:

(i) From each lesion, Haralick texture features [8] were computed for each voxel inside the VOI from the Gray Level Correlation Matrix (GLCM) (calculation details given in Supplementary methods). The Haralick features consisted of energy, entropy, homogeneity, and contrast. Second, pairwise similarities for all sites were computed using squared Euclidean distance from the weighted mean texture and intensity values. The pairwise similarities generated an inter-site similarity matrix (ISM) wherein site pairs that were highly similar in appearance had lower values while those that were highly dissimilar had higher values in the corresponding cell of the similarity matrix (Fig. 1).

Inter-site tumour heterogeneity texture metrics

The following 12 inter-site tumour heterogeneity texture metrics were computed from the ISM:

- First order similarity features: computed from a histogram of the similarities of all the site pairs. The (1) mean, (2) s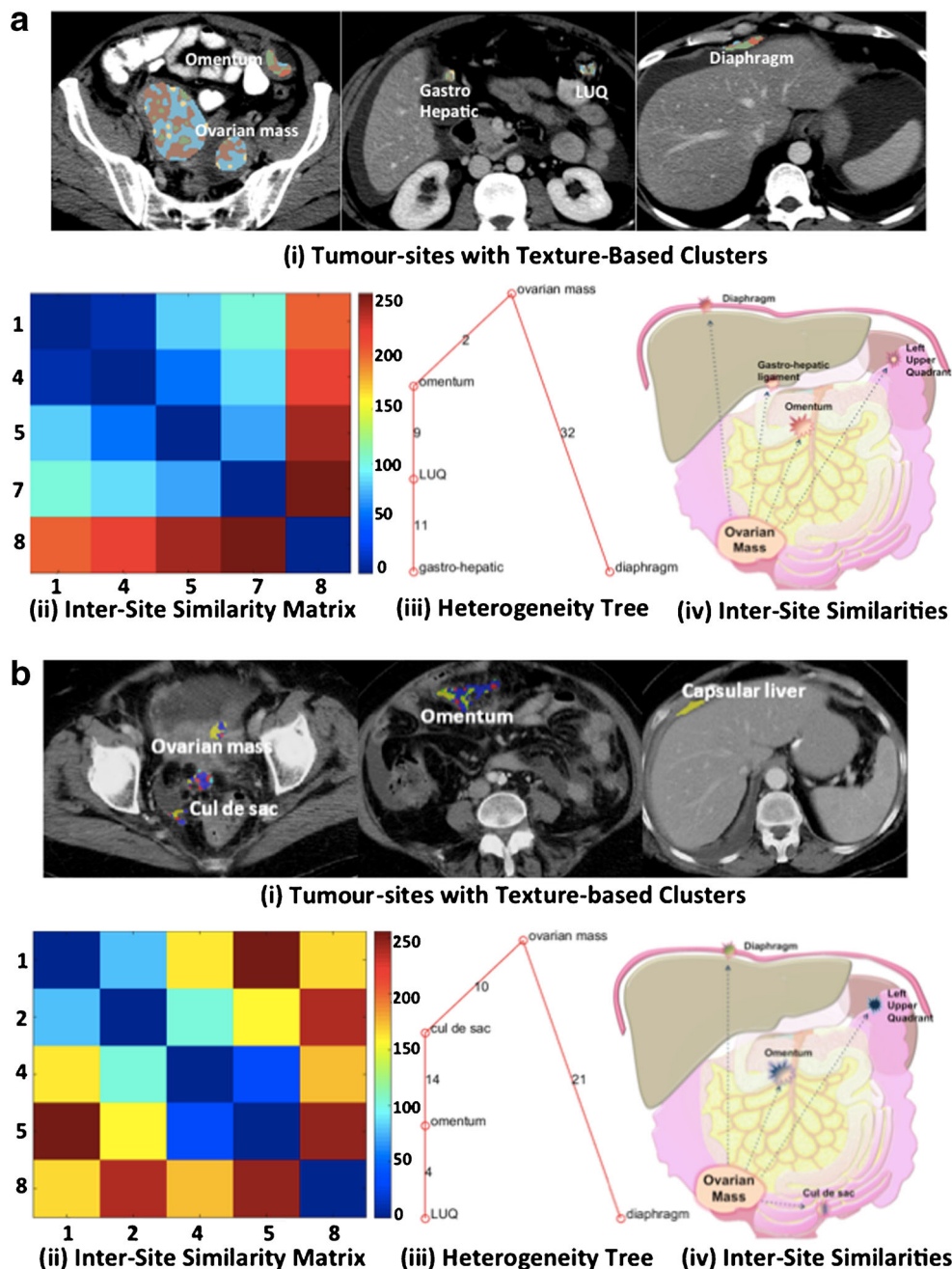tandard deviation (SD), (3) kurtosis, and (4) skewness capture the different moments of the histogram and reflect the differences in the similarity of all the site pairs.

- GLCM Haralick texture of similarities: The computed features comprised of (5) energy (6) correlation (7) contrast and (8) homogeneity. Multiple GLCM matrices were computed using offsets in all directions (n = 8) and space (m = 1,2,3), and the resulting texture feature values from all the GLCMs were averaged to generate the final texture value for the similarity matrix.

- Inter-site similarity entropy (SE) (9): captures the amount of variation in the similarities between the different sites. In other words, the more different each site is with respect to the other sites in appearance, the higher the inter-site entropy. In patients where all tumour sites appear very similar to one another, the inter-site entropy will be close to 0.

- Similarity Level-Area matrix metrics: these were designed specifically to extract the uniformity in the distribution of similarities for a group of sites. For instance, they would enable differentiation between two patients where in one all the tumour sites are *equally* similar or *equally* dissimilar from one another versus another patient where the tumour sites have *highly variable* similarities with respect to each other. The metrics derived from the similarity level-area matrix were (10) inter-site cluster variance (SCV), (11) cluster shade (SCS), and (12) cluster prominence (SCP). Both SCS and SCP are measures of skewness and asymmetry in the distribution of intensity levels in an image and are inversely proportional to the amount of asymmetry in the image. Given that the measures are computed from the similarity level-area matrix, a highly skewed matrix will result in small values for the SCS and SCP indicating that patients with homogeneous similarities will have low values of SCS and SCP.

Fig. 1 (**a**) Patient with Classification of Ovarian Cancer (CLOVAR) mesenchymal subtype and an overall survival of 69 months. Texture-based results within each tumour site (i), the inter-site similarity matrix (ISM) (ii), the heterogeneity tree (iii), and the schematic of the dissimilarity of the various sites compared to the ovarian mass (iv). For instance, the diaphragmatic tumour implant (#8 in ii) has the largest dissimilarity compared to the ovarian mass. (**b**) Patient with CLOVAR mesenchymal subtype and an overall survival of 10 months. Texture-based results within each tumour site (i), the ISM (ii), the heterogeneity tree (iii), and the schematic of the dissimilarity of the various sites compared to the ovarian mass (iv). For instance, the left upper quadrant (LUQ) (#5 in b) has the largest dissimilarity compared to the ovarian mass, followed by the diaphragmatic and the omentum tumour implants. The numbers listed in the x axis of both figures indicates the numerical codes for lesion location (1 = primary ovarian mass, 2 = cul de sac, 4 = omentum, 5 = left upper quadrant, 7 = gastro-hepatic ligament, 8 = diaphragm)

a

(i) Tumour-sites with Texture-Based Clusters

(ii) Inter-Site Similarity Matrix   **(iii) Heterogeneity Tree**   **(iv) Inter-Site Similarities**

b

(i) Tumour-sites with Texture-based Clusters

(ii) Inter-Site Similarity Matrix   **(iii) Heterogeneity Tree**   **(iv) Inter-Site Similarities**

Since all the aforementioned features are computed from the inter-site similarity matrix, meaningful computation requires a minimum of three metastatic sites. With less than three sites, there are too few distinct numbers to compute the inter-site similarity metrics introduced in this work. This is not a major issue in HGSOC, as most patients present with stage III-IV and multiple sites of disease. In our patient population, one patient had four sites, four patients had five sites, and the remaining patients had six or more sites of disease. As patients can present with metastatic disease in different locations, the measures were designed to be independent of the number of sites and their specific anatomical location.

Tumour heterogeneity tree generation

We developed a minimum spanning tree-based method to automatically generate a spatial tree-like representation of the heterogeneity (dissimilarity) among the various sites with respect to either ovarian mass or cul de sac disease (when ovarian masses were not present, two cases out of 38) from the

ISM. Tree generation using image-based features was previously introduced in [12]. Further details on Tree generation are given in the Supplementary methods.

Decision tree-based threshold and feature selection

The thresholds for splitting the patients into two groups using the various heterogeneity metrics were computed using decision trees. A recursive feature elimination technique was used to extract the cut-points and rank the features by their importance using the decision tree method available in Matlab (The MathWorks Inc, USA). A feature identified to be the root node of the decision tree was chosen and the value associated with that node used as the cut-point. Following its selection, the same feature was removed from the feature pool and the decision tree was recomputed with the remaining features until no more features remained. The decision tree method was trained to select features for distinguishing patients by their survival (>60 vs. ≤60 months).

Feature ranking through Least Absolute Shrinkage and Selection Operator (LASSO) regression

Feature ranking was assessed using LASSO regression [13], and subsequently the ranking was evaluated for predicting surgical resection status and overall survival (>60 vs. ≤60 months). The LASSO method implementation available in R software (R Foundation for Statistical Computing, Austria) was used, with a shrinkage penalty of $\alpha = 0.5$, thereby performing elastic net regression. Leave-one-out cross-validation was performed to extract the models. The model was chosen using the regularizer (λ) value that resulted in greatest amount of shrinkage and whose cross-validation mean square error was within one standard error from the minimum cross-validated mean square error.

Supervised clustering

The inter-site tumour heterogeneity metrics were used to divide the patients into two groups through supervised k-means clustering. For meaningful distance computation, the metrics were scaled using z-score normalization, wherein each feature was subtracted by the mean and divided by the standard deviation. The clustering method was evaluated through its performance for separating between low and high risk groups. The high-risk group was defined as patients who underwent incomplete surgical resection or had CCNE1 amplification and had an overall survival of ≤60 months. Conversely low risk group consisted of patients who underwent complete resection and had a survival of >60 months and no CCNE1 amplification. A leave-one-out cross-validation (LOOCV) strategy was then applied (additional information in Supplementary methods).

Statistical analysis

Complete surgical resection (CSR) was defined as no residual disease after surgery as assessed by the operating surgeon as part of the patients' standard of care. Overall survival was defined as the time interval between the preoperative CT exam and date of death or censure. Patients who survived were followed for at least 5 years. The curves for predicting CSR and overall survival were estimated using Kaplan-Meier (KM) method. The difference between the curves was computed using log-rank test. The cut points for splitting the patients into the different groups for each variable was computed using binary decision trees (implementation available in Matlab). Previously decision trees have been used with KM-based survival analysis for breast cancers [14]. Similarly, differences in inter-site heterogeneity metrics were evaluated according to Classification of Ovarian Cancer (CLOVAR) subtype (mesenchymal versus non-mesenchymal) subtypes using unpaired t-tests. R software (R Foundation for Statistical Computing, Austria) was used for analysis.

Results

Inter-site similarity matrices (ISM)

The similarity matrix provides a visual representation of the distribution of texture heterogeneity across all pairs of tumour sites from a patient, while the heterogeneity tree illustrates the relative similarity of each tumour site with respect to another tumour site (Fig. 1). The cases shown in Fig. 1 provide examples of two patients with mesenchymal subtype HGSOC and similar number of disease sites whose tumours nonetheless exhibit very distinct inter-site textural heterogeneity characteristics. Case 1 had complete surgical resection and experienced longer overall survival (69 months), whereas Case 2, which demonstrated greater texture heterogeneity on ISM, had incomplete surgical resection and shorter overall survival (10.4 months).

Associations between inter-site heterogeneity metrics, survival, and surgical resection status

Of the inter-site texture heterogeneity metrics evaluated, three measures that capture the differences in texture similarities across sites were associated with shorter OS [Table 3, log-rank test for inter-site similarity entropy (SE), similarity level cluster shade (SCS), and inter-site similarity level cluster prominence (SCP), $p = 0.02, 0.017$ and 0.028, respectively]. There was no significant association between OS and the remaining nine heterogeneity metrics that average the inter-site similarities. The following features were ranked highest for predicting OS using when using recursive feature elimination with decision trees: SE, SCP, correlation, SCS, and contrast. Independent evaluation of the data with leave one-out cross validation using

Table 3 Decision tree threshold values determined for texture features and their association (log-rank test) to overall survival

Measure	Threshold	*p*-value
SE	1.81	0.02*
SCP	1751	0.028*
Correlation	0.037	0.08
SCS	239	0.017*
Contrast	32	0.11
Inter-site skewness	0.97	0.36
SCV	0.69	0.066
Inter-site SD	52.3	0.42
Inter-site Mean	122.7	0.20
Inter-site kurtosis	3.30	0.74
Homogeneity	0.04	0.12
Energy	0.36	0.19

SE inter-site similarity entropy, *SCP* inter-site similarity level cluster prominence, *SCS* inter-site similarity level cluster shade, *SCV* inter-site similarity level variance, *SD* standard deviation

Table 4 Relationship between inter-site tumour texture heterogeneity metrics and surgical resection status

Measure	Complete	Incomplete	*p*-value
SCS	146.2	261.7	0.02*
SCP	1236.1	2272.7	0.033*
SCV	0.32	0.85	0.0097*
Homogeneity	0.37	0.34	0.78
SE	2.24	2.65	0.16
Inter-site skewness	0.74	0.51	0.94
Inter-site SD	76.7	65.6	0.53
Energy	0.07	0.04	0.88
Inter-site Mean	108.9	105.7	0.60
Inter-site kurtosis	3.2	3.4	0.81
Correlation	0.09	0.05	0.057
Contrast	21.3	20.5	0.49

SE inter-site similarity entropy, *SCP* inter-site similarity level cluster prominence, *SCS* inter-site similarity level cluster shade, *SCV* inter-site similarity level variance, *SD* standard deviation

Complete surgical resection was defined as no residual disease after surgery as assessed by the operating surgeon as part of the patients' standard of care

LASSO regression identified SE (w = -0.084SE + 0.0SCV + 0.0SCS + 0.0Mean + 0.0SD + 0.0Kurtosis + 0.0Skewness + 0.0Contrast + 0.0Correlation + 0.0Homogeneity + 0.0Energy) for predicting overall survival. The negative weight indicates that smaller values of SE (less heterogeneity) predict longer OS. For example, in our dataset, a patient with an SE of 2.21 had an OS of 71 months as opposed to a different patient with an SE of 2.3 who had an OS of 49 months. Measures of heterogeneity (SCV, SCS, and SCP) were significantly lower for patients who had complete compared to incomplete surgical resection (Table 4). In patients who had complete surgical resection, inter-site similarities were positively correlated and had higher energy and homogeneity compared to those with incomplete surgical resection. When trained using LOOCV strategy, LASSO regression identified the model (w = -0.49SE − 0.29SCV -0.0SCP + 0.0SCS + 0.0Mean -0.0002SD + 0.0Kurtosis + 0.0Skewness + 0.0Contrast + 0.0Correlation + 0.0Homogeneity + 0.0Energy) for predicting surgical resection status. The model for resection status indicates that lower heterogeneity is associated with complete resectability. The association between complete surgical resection and OS was not statistically significant ($p = 0.215$), possibly because optimal debulking was achieved in the majority (79%) of patients. In addition, neither the total number of disease sites per patient nor the overall tumour volume per patient were associated with OS ($p = 0.33$) (Fig. 2).

When splitting patients into two groups by thresholding according to respective inter-site tumour heterogeneity metrics, increased heterogeneity (SE < 3.0; SCV < 0.9; SCS < 250; SCP < 1500) was associated with poorer survival (median OS 50 months vs. 69 months) and decreased complete surgical resection rates (29% vs. 71%).

Associations between inter-site heterogeneity metrics, CCNE1 amplification and CLOVAR subtyping

Amplification of 19q12 involving cyclin E1 gene (*CCNE1*) predominantly occurred in patients with more heterogeneous inter-site textures (Fig. 3). Concretely, eight out of 13 patients with CCNE1 amplification occurred in the heterogeneous cluster. The median and inter-quartile range of the heterogeneity measures in the two clusters were as follows: low heterogeneity [SE 2.34(1.96-2.71); SCV 0.68(0.69-0.76); SCS 150.6(119.9-297.5); SCP 1282.5(981.8-2763.4), and high heterogeneity cluster (SE 2.56(2.14-2.87); SCV 0.56(0.46-0.94); SCS 164.8(117.2-239.7); SCP 1429.6(842.4-2041.5)]. *CCNE1* amplification has been associated with worse prognosis and primary treatment failure in patients with HGSOC [15, 16]. Ongoing studies by our group integrate tumour heterogeneity at the level of imaging, somatic alterations and expression using three-dimensional molds. We did not find significant associations between inter-site texture heterogeneity metrics and TCGA Classification of Ovarian Cancer (CLOVAR) subtype [4] when grouped by mesenchymal versus non-mesenchymal (better prognosis) subtypes (Table 5), with the exception of inter-site homogeneity.

Leave-one-out cross-validation

Supervised fuzzy c-means clustering using the inter-site tumour heterogeneity metrics outperformed the other texture measures for separating low and high risk groups as defined in the methods section [accuracy 71%, true positive rate (TPR) 60%, and true negative rate (TNR) 86%]. In comparison, the

Fig. 2 Relationship between (**a**) the total number of disease sites per patient, inter-site texture heterogeneity metrics and 60 month survival and (**b**) the overall tumour volume per patient (calculated from CT as the sum of the volume from each segmented individual tumour site per patient), inter-site texture heterogeneity metrics and 60-month survival
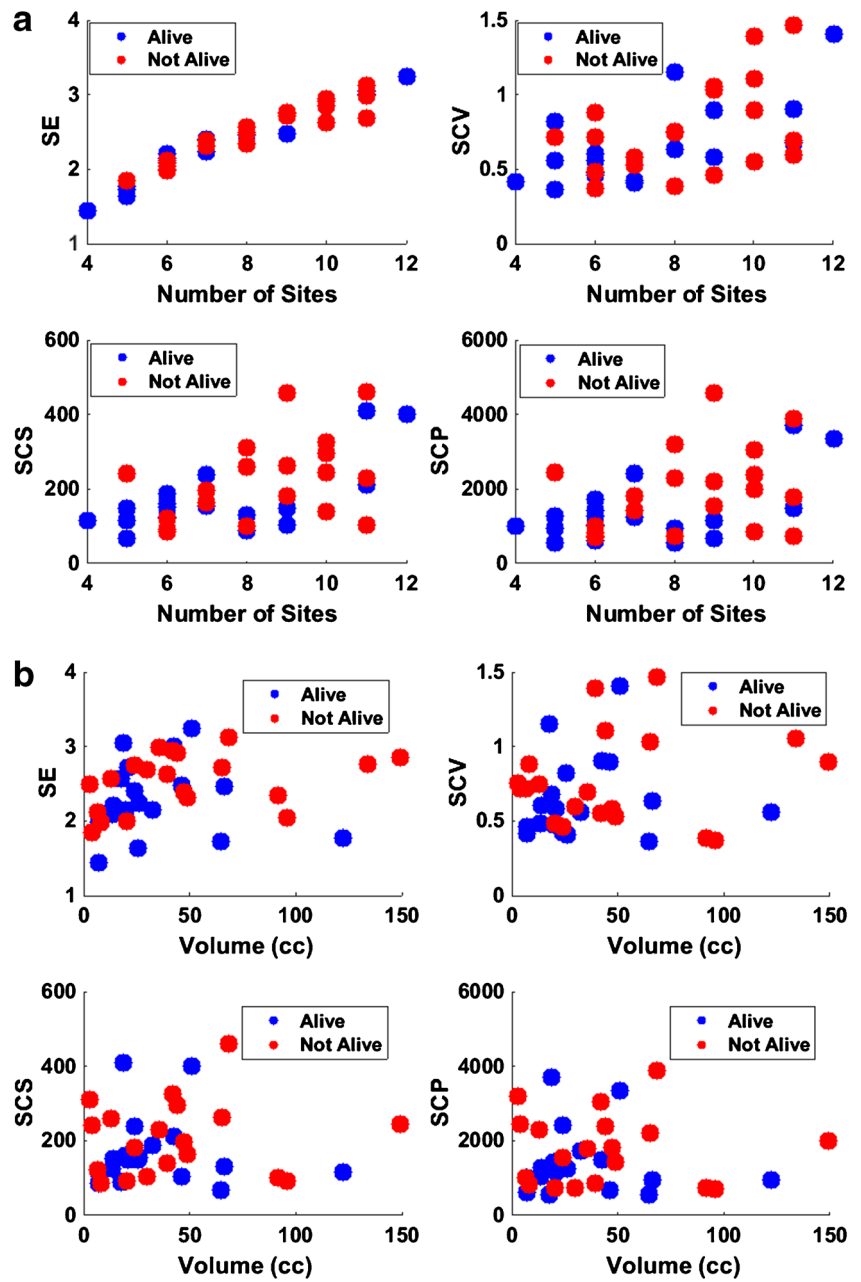

combination of first order and Haralick-based textures resulted in an accuracy of 56%, TPR of 40%, and TNR of 79% (Supplementary Table 1). Our results indicate that the algorithm was able to produce a separation between the two patient groups and the distribution of data produced using the algorithm generated clustering comparable to the ground truth (Fig. 3).

Discussion

In this study, we developed quantitative metrics that non-invasively capture spatial inter-site imaging heterogeneity from multiple metastatic tumour sites in patients with HGSOC. The computed measures were significantly associated with complete surgical resection and survival, even though there was no significant association between these clinical endpoints and simple radiographic findings such as the number of sites and total volume of disease identified on CT. Supervised clustering results with leave-one-out cross-validation show that clustering using the inter-site tumour heterogeneity measures achieved the best separation between different risk groups compared to the texture-based measures averaged across the tumour sites. The results suggest that textural analysis may provide added value in the evaluation of HGSOC patients, beyond the standard more traditional use for simple delineation of disease extent.

Fig. 3 Relation between inter-site tumour texture heterogeneity metrics and genomic alterations. (**a**) Heat map showing the relative similarities of the patients with respect to each other computed using the inter-site texture heterogeneity metrics. Similarity between two patients was computed using Euclidean distance measure after standardizing the features using z-score normalization. Two patients that are highly similar have a value close to 0 (red). (**b**) Patients were divided into two groups; the high risk group was defined as patients who underwent incomplete resection or had CCNE1 amplification and had an overall survival of ≤60 months. Conversely, low risk group consisted of patients who underwent complete resection and had a survival of >60 months and no CCNE1 amplification. The SCP, SE, and SCV values for each patient in the low and high risk categories according to the clinical variables and the model (classifier) are shown
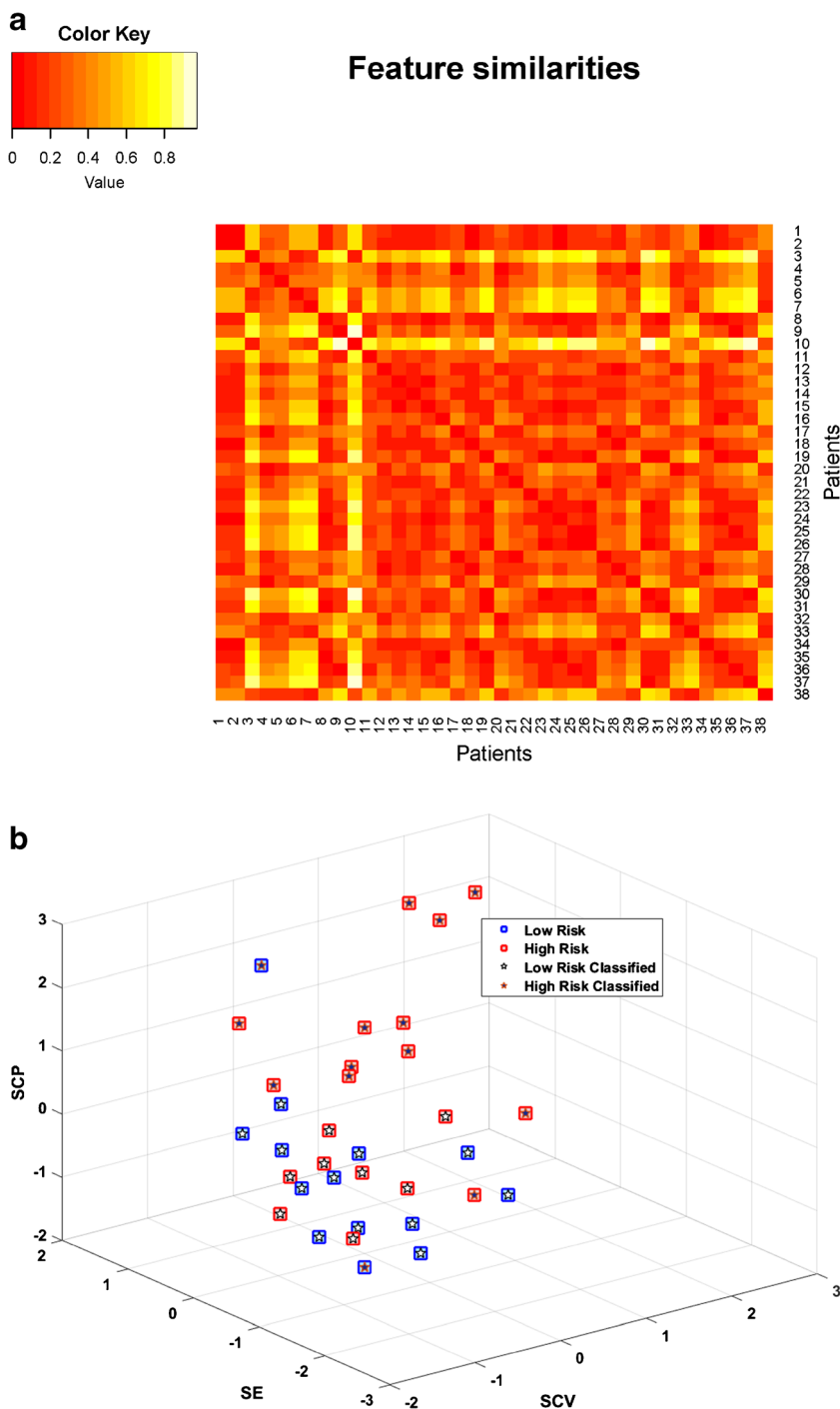
Our study contrasts to previous radiogenomic studies that evaluated a single tumour site per patient [9, 10]. Following extraction of >100 CT imaging features, Segal et al. [10] found a subset of 14 features predicted 80% of the gene expression patterns and outcome in patients with hepatocellular carcinoma. Recently, texture analysis of 440 CT-based features of 1,019 patients with lung or head-and-neck cancer found a number of prognostically relevant metrics associated with gene-expression patterns and predicted outcome [9]. The texture metrics in these studies were derived from a single tumour site, and no account was made for inter-site tumour heterogeneity, an important factor for tumour evolution and outcome which may affect the ability to design and select effective therapies. The clinical and therapeutic importance of inter-site tumour heterogeneity is becoming increasingly clear [17–19]. In HGSOC, high degrees of clonal expansion

Table 5 Relationship between inter-site tumour texture heterogeneity metrics and Classification of Ovarian Cancer (CLOVAR) subtypes

Measure	Mean value of the heterogeneity measure				Mesenchymal vs. Non-mesenchymal p-value
	Mesenchymal	Differentiated	Immunoreactive	Proliferative	
SE	2.27	2.68	2.43	1.62	0.11
SCV	0.50	0.67	0.41	0.34	0.94
SCS	160.4	248.6	214.14	114.59	0.08
SCP	1340.9	2136.2	1939.4	1355.6	0.08
Inter-site mean	105.7	95.5	111.9	122.8	0.64
Inter-site SD	71.9	62.7	67.9	91.4	0.056
Inter-site kurtosis	3.96	3.29	3.09	2.01	0.34
Inter-site skewness	0.88	0.67	0.27	0.57	0.18
Correlation	0.11	0.086	-0.04	0.54	0.41
Contrast	22.5	17.4	23.6	17.03	0.34
Energy	0.076	0.04	0.04	0.06	0.87
Homogeneity	0.36	0.36	0.33	0.39	0.014

SE inter-site similarity entropy, *SCV* inter-site similarity level variance, *SCS* inter-site similarity level cluster shade, *SCP* inter-site similarity level cluster prominence

are associated with shorter survival [20]. Approximately 50% of HGSOC contain genetic alterations related to homologous recombination (HR) deficiency. Therapies based on PARP inhibition, which lethally targets tumours with HR deficiency, could benefit many HGSOC patients [21]. However, establishing genomic profiles and defining potential targets in ovarian cancers is dependent on tumour heterogeneity. Primary ovarian lesions and metastases show different mutational landscapes with branched tumour evolution [20, 22]. Thus, although there is an increase in potential therapeutic targets being identified in HGSOC, their success will be dependent on the availability of biomarkers able to document the presence and extent of spatial and temporal heterogeneity [22]. An accurate understanding of inter-tumour heterogeneity may facilitate more appropriate triage of patients to effective therapies, as well as anticipation and treatment of resistant lesions.

The heterogeneity metrics developed in this work are independent of the number of sites or the specific anatomical tumour sites with the exception of the inter-site entropy. However, this modest increase in the level of dissimilarity between sites is expected as the number of sites increases, as corroborated by the near linear relation of the site entropy with the number of sites. This is because entropy is derived using Shannon entropy from information theory where entropy captured the number of bits necessary to carry information; thus by definition, more bits (tumour sites) would increase entropy. However, when comparing patients with the same number of sites, those with sites that depict large variations in the similarities had higher entropy compared to the patients whose sites appeared equally similar. This is illustrated in Fig. 1,

where two patients have the same number of sites involved by disease and yet display very different texture characteristics and outcomes. SCS, SCV, and SCP are independent of the number of tumour sites.

In our dataset, tumour heterogeneity metrics including SE, SCP, SCS, and SCV were most relevant metrics for predicting CSR and survival, whereas the first order and the GLCM-based features did not predict patient outcomes. First-order measures average the similarity thereby potentially losing information pertaining to a specific group of sites. Similarly, the GLCM measures examine only the immediately neighboring site and again do not capture the similarity in a specific group of sites.

The literature regarding the differential imaging features in relation to varying gene expression profiles in HGSOC is limited. Our group recently reported on associations between qualitatively assessed CT features, CLOVAR subtypes and survival in HGSOC [11]. The presence of mesenteric infiltration and diffuse peritoneal involvement by tumour on CT were significantly associated with mesenteric CLOVAR subtype of HGSOC. In the current work we have taken a further step and found that spatial heterogeneity metrics may predict CSR and survival regardless of CLOVAR subtype. We found that amplification of 19q12 involving *CCNE1* predominantly occurred in patients with more heterogeneous inter-site textures (Fig. 3). *CCNE1* amplification has been associated with worse prognosis and treatment failure in HGSOC [15, 16]. Ongoing studies by our group integrate tumour heterogeneity at the level of imaging, somatic alterations and expression using three-dimensional molds.

Our study has several limitations. The small sample size limits the power of the statistical analysis is, and thus the reported conclusions require further validation in a larger cohort. Sample size also limits correlation with additional factors such as germline and somatic BRCA status, homologous repair pathway and other somatic alterations, which may also correlate with texture features as they do with outcomes. Validation will need to take into account the potential impact of methodologic differences on the ability to apply thresholds from one dataset to another. This has been demonstrated in the field of genomics, where, for example, mutational load is consistently associated with response to checkpoint blockade immunotherapy, but the threshold at which this effect occurs is different in each study [23–25]. While the retrospective study design is considered a limitation, we do not expect that this design biased the results of the texture analysis methods. Finally, we could not directly correlate the texture features extracted from the images to the anatomical site of the tissue that was used for the genomic analysis. As previously reported, individual HGSOCs expressed multiple subtype signatures; however, for the purpose of the study, we used the best gene expression assignment as per Verhaak et al. [4].

In summary, this work introduced CT texture-based measures of inter-tumour heterogeneity. Our study provides preliminary data suggesting that texture measures can be used to capture the spatial heterogeneity of multiple tumour sites, as well as to predict non-invasively clinically relevant outcomes for women with HGSOC. As the clinical and therapeutic importance of inter-site tumour heterogeneity becomes increasingly clear [17–19], the potential for standard-of-care scans to noninvasively provide heterogeneity metrics may facilitate more appropriate triage of HGSOC patients to effective therapies, as well as anticipation and treatment of resistant lesions.

References

1. Engel J, Eckel R, Schubert-Fritschle G, Kerr J, Kuhn W, Diebold J et al (2002) Moderate progress for ovarian cancer in the last 20 years: prolongation of survival, but no improvement in the cure rate. Eur J Cancer 38:2435–2445

2. Hyman DM, Zhou Q, Iasonos A, Grisham RN, Arnold AG, Phillips MF et al (2012) Improved survival for BRCA2-associated serous ovarian cancer compared with both BRCA-negative and BRCA1-associated serous ovarian cancer. Cancer 118:3703–3709

3. Tothill RW, Tinker AV, George J, Brown R, Fox SB, Lade S et al (2008) Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. Clin Cancer Res 14:5198–5208

4. Verhaak RG, Tamayo P, Yang JY, Hubbard D, Zhang H, Creighton CJ et al (2013) Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. J Clin Invest 123:517–525

5. Soslow RA (2008) Histologic subtypes of ovarian carcinoma: an overview. Int J Gynecol Pathol 27:161–174

6. Khalique L, Ayhan A, Weale ME, Jacobs IJ, Ramus SJ, Gayther SA (2007) Genetic intra-tumour heterogeneity in epithelial ovarian cancer and its implications for molecular diagnosis of tumours. J Pathol 211:286–295

7. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB et al (2012) Radiomics: the process and the challenges. Magn Reson Imaging 30:1234–1248

8. Haralick RM, Shanmugam K, Dinstein IH (1973) Textural features for image classification. IEEE Trans Syst Man Cyber SMC-3:610–621

9. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006

10. Segal E, Sirlin CB, Ooi C, Adler AS, Gollub J, Chen X et al (2007) Decoding global gene expression programs in liver cancer by non-invasive imaging. Nat Biotechnol 25:675–680

11. Vargas HA, Micco M, Hong SI, Goldman DA, Dao F, Weigelt B et al (2015) Association between morphologic CT imaging traits and prognostically relevant gene signatures in women with high-grade serous ovarian cancer: a hypothesis-generating study. Radiology 274:742–751

12. Veeraraghavan H, Schrater P, Papanikolopoulos NP (2006) Adaptive templates for feature matching. Proc IEEE Int Conf Robot Autom 3393–3398

13. Tibshirani R (1996) Regression shrinkage and selection via the lasso. J R Stat Soc Ser B Methodol 58:267–288

14. Ture MTF, Kurt I (2009) Using Kaplan-Meier analysis together with decision tree methods in determining recurrence-free survival of breast cancer patients. Expert Syst Appl 36:2017–2026

15. Etemadmoghadam D, deFazio A, Beroukhim R, Mermel C, George J, Getz G et al (2009) Integrated genome-wide DNA copy number and expression analysis identifies distinct mechanisms of primary chemoresistance in ovarian carcinomas. Clin Cancer Res Off J Am Assoc Cancer Res 15:1417–1427

16. Patch AM, Christie EL, Etemadmoghadam D, Garsed DW, George J, Fereday S et al (2015) Whole-genome characterization of chemoresistant ovarian cancer. Nature 521:489–494

17. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E et al (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366:883–892

18. Juric D, Castel P, Griffith M, Griffith OL, Won HH, Ellis H et al (2015) Convergent loss of PTEN leads to clinical resistance to a PI(3)Kalpha inhibitor. Nature 518:240–244

19. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W et al (2013) Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. Clin Cancer Res 19:2240–2247

20. Schwarz RF, Ng CK, Cooke SL, Newman S, Temple J, Piskorz AM et al (2015) Spatial and temporal heterogeneity in high-grade serous ovarian cancer: a phylogenetic analysis. PLoS Med 12:e1001789

21. Mittempergher L (2016) Genomic characterization of high-grade serous ovarian cancer: dissecting its molecular heterogeneity as a road towards effective therapeutic strategies. Curr Oncol Rep 18:44

22. Paracchini L, Mannarino L, Craparotta I, Romualdi C, Fruscio R, Grassi T et al (2016) Regional and temporal heterogeneity of epithelial ovarian cancer tumor biopsies: implications for therapeutic strategies. Oncotarget. doi:10.18632/oncotarget.10505

23. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ et al (2015) Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science (New York, NY) 348:124–128

24. Rosenberg JE, Hoffman-Censits J, Powles T, van der Heijden MS, Balar AV, Necchi A et al (2016) Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet (London, England) 387:1909–1920

25. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S et al (2016) Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. Cell 165:35–44