




Integration of proteomics with CT-based qualitative and radiomic features in high-grade serous ovarian cancer patients: an exploratory analysis

Lucian Beer¹ · Hilal Sahin¹ · Nicholas W. Bateman^{2,3} · Ivana Blazic⁴ · Hebert Alberto Vargas⁵ · Harini Veeraraghavan⁶ · Justin Kirby⁷ · Brenda Fevrier-Sullivan⁷ · John B. Freymann⁷ · C. Carl Jaffe⁸ · James Brenton^{9,10} · Maura Miccò¹¹ · Stephanie Nougaret¹² · Kathleen M. Darcy^{2,3} · G. Larry Maxwell^{2,3,13} · Thomas P. Conrads^{2,3,13,14} · Erich Huang¹⁵ · Evis Sala^{1,16} 

Received: 5 September 2019 / Revised: 21 January 2020 / Accepted: 17 February 2020 / Published online: 6 April 2020
© The Author(s) 2020

Abstract

Objectives To investigate the association between CT imaging traits and texture metrics with proteomic data in patients with high-grade serous ovarian cancer (HGSOC).

Methods This retrospective, hypothesis-generating study included 20 patients with HGSOC prior to primary cytoreductive surgery. Two readers independently assessed the contrast-enhanced computed tomography (CT) images and extracted 33 imaging traits, with a third reader adjudicating in the event of a disagreement. In addition, all sites of suspected HGSOC were manually segmented texture features which were computed from each tumor site. Three texture features that represented intra- and inter-site tumor heterogeneity were used for analysis. An integrated analysis of transcriptomic and proteomic data identified proteins with conserved expression between primary tumor sites and metastasis. Correlations between protein abundance and various CT imaging traits and texture features were assessed using the Kendall tau rank correlation coefficient and the Mann-Whitney *U* test, whereas the area under the receiver operating characteristic curve (AUC) was reported as a metric of the strength and the direction of the association. *P* values < 0.05 were considered significant.

Results Four proteins were associated with CT-based imaging traits, with the strongest correlation observed between the CRIP2 protein and disease in the mesentery ($p < 0.001$, AUC = 0.05). The abundance of three proteins was associated with texture features that represented intra- and inter-site tumor heterogeneity, with the strongest negative correlation between the CKB protein and cluster dissimilarity ($p = 0.047$, $\tau = 0.326$).

Conclusion This study provides the first insights into the potential associations between standard-of-care CT imaging traits and texture measures of intra- and inter-site heterogeneity, and the abundance of several proteins.

Key Points

- CT-based texture features of intra- and inter-site tumor heterogeneity correlate with the abundance of several proteins in patients with HGSOC.
- CT imaging traits correlate with protein abundance in patients with HGSOC.

Keywords Ovarian neoplasms · Radiomics · Proteomics · Gene expression profiling · Prognosis

Lucian Beer, Hilal Sahin, Erich Huang and Evis Sala contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00330-020-06755-3>) contains supplementary material, which is available to authorized users.

✉ Evis Sala
es220@cam.ac.uk

Extended author information available on the last page of the article

Abbreviations

AUC	Area under the receiver operating characteristic curve
CLOVAR	Classification of ovarian cancer transcriptomic profiles
CT	Computed tomography
CV	Coefficients of variation
HGSOC	High-grade serous ovarian cancer
HIPAA	Health Insurance Portability and Accountability Act

IISTH	Intra- and inter-site heterogeneity measures
MRI	Magnetic resonance imaging
PROVAR	Protein-driven index of ovarian carcinoma
TCGA	The Cancer Genome Atlas
TCIA	The Cancer Imaging Archive
VOI	Volume of interest

Introduction

It is estimated that over 14,000 women will die from epithelial ovarian cancer in 2019 in the USA, the preponderance of whom will have high-grade serous ovarian cancer (HGSOC) [1]. Despite improvements in the treatment of ovarian cancer, the rate of recurrence remains high and both disease-free survival and overall survival are poor, due to resistance to standard platinum-based chemotherapy [2]. Although several clinical-pathological factors, such as age, stage, histologic grade, and surgical debulking status, have been shown to be important prognostic indicators of outcome in ovarian cancer patients, there are still no reliable biomarkers for predicting response to therapy and outcome [3].

Several studies have identified genetic markers that are associated with outcome in HGSOC patients [4–7], with TP53 the major driver mutation [8, 9]. Risk stratification using gene expression is reliable only in a subset of patients [6], as gene function may not necessarily correlate with the cognate gene product (e.g., protein) function. An analysis of 412 cases from The Cancer Genome Atlas (TCGA) study, which used reverse-phase protein arrays, identified a protein-driven index of ovarian carcinoma (PROVAR), which enabled classification of patients with different risks of recurrence and survival [10]. In addition, the Clinical Proteomic Tumour Analysis Consortium performed a mass spectrometry-based proteomic analysis of 174 ovarian tumors previously analyzed by the TCGA and integrated those data with genome-level data from a whole-exome sequencing [11]. This work showed that an abundance of selected proteins was associated with genomic changes, such as chromosomal structural abnormalities, copy number alterations, and the homologous recombination deficiency status. In addition, protein signatures were identified as a strong independent predictor for patient survival.

Imaging by computed tomography (CT) and, in selected cases, by magnetic resonance imaging (MRI), is used to evaluate the extent of disease and monitor treatment response in patients with HGSOC. To date, a few studies have applied radiomic feature analysis in patients with ovarian cancer. They showed that MRI-derived radiomic features can discriminate between benign and malignant ovarian masses, with a high accuracy of 87% [12]. Furthermore, CT radiomic features of patients with ovarian cancer correlate with response to therapy [13], progression-free survival [14, 15], and overall survival [15], and can identify patients at higher risk for

recurrence [16]. Recent work by our group focused on evaluating the possible associations between CT imaging traits and texture metrics with genomics data and patient outcome [17–19]. The integration of clinical, proteomic, and radiomic data may enable to stratify patients according to risk for progression thereby allowing for tailored therapy [20].

In this pilot, hypothesis-generating study, we investigated the association between CT imaging traits and texture metrics with proteomic data in a small cohort of patients with HGSOC.

Materials and methods

Study population

This was a multi-institutional, institutional review board-approved, and Health Insurance Portability and Accountability Act (HIPAA)-compliant retrospective study, with waiver of informed consent from all institutions that participated in The Cancer Genome Atlas-Ovarian Cancer (TCGA-OV) Imaging Research Group [21].

The eligibility criteria included the following: (1) confirmed diagnosis of HGSOC, (2) HGSOC tissue submitted and analyzed by TCGA, (3) no neoadjuvant chemotherapy administered, (4) intravenous contrast-enhanced CT of the abdomen and pelvis performed prior to primary cytoreductive surgery, and (5) protein relative abundance measurements available. The final patient cohort consisted of 20 patients. Figures 1 and 2 summarize the experimental design. All patients were included in two prior studies that evaluated the association between the Classification of Ovarian Cancer (CLOVAR) subtype signatures and CT features in a single- [17] and multi-institutional setting [21], respectively. All patients were also included in two other studies that evaluated the feasibility of CT-based texture measures in the quantification of inter-site tumor heterogeneity and combined intra- and inter-site tumor heterogeneity [17, 22] in patients with HGSOC. None of these prior studies involved any analysis of protein abundance data.

Image analysis

Thirty-six image features, including the 33 radiologist-scored CT imaging traits used in the study by Vargas et al, such as lesion size and laterality, locations of peritoneal disease, nodal stations involved, and locations of metastases, and three computer-extracted texture metrics were obtained [21]. All CT-derived variables are shown in Table 1 and Table 2.

The CT images were a part of the TCIA TCGA-OV [23] collection in The Cancer Imaging Archive (<https://cancerimagingarchive.net>) [24], an NCI-supported archive of anonymized medical images. The cases were reviewed on a

Fig. 1 Study flow chart. Patients analyzed in this study were part of the TCGA dataset ($n = 169$). Protein candidates (5504 total) that exhibited stable expression between primary and metastatic sites and those that highly correlated with transcript abundance ($n = 107$ HGSOE patients, TCGA, and CPTAC) were selected using the Student's t test (p value < 0.1) after boxCox normalization of transcript expression coefficients of variation and Spearman correlation distributions, respectively. Forty-seven genes exhibited both low correlation of variables between primary and metastatic sites (529 candidates) and were highly correlated at the protein and transcript levels (569 candidates). Among those, 16 proteins were associated with amino acid metabolism and selected for final analysis. CT data were available in 20 patients. Thirty-three radiologist-scored imaging traits and three CT-based texture measures of inter-site tumor heterogeneity were assessed and their correlation with the 16 cancer proteins was calculated

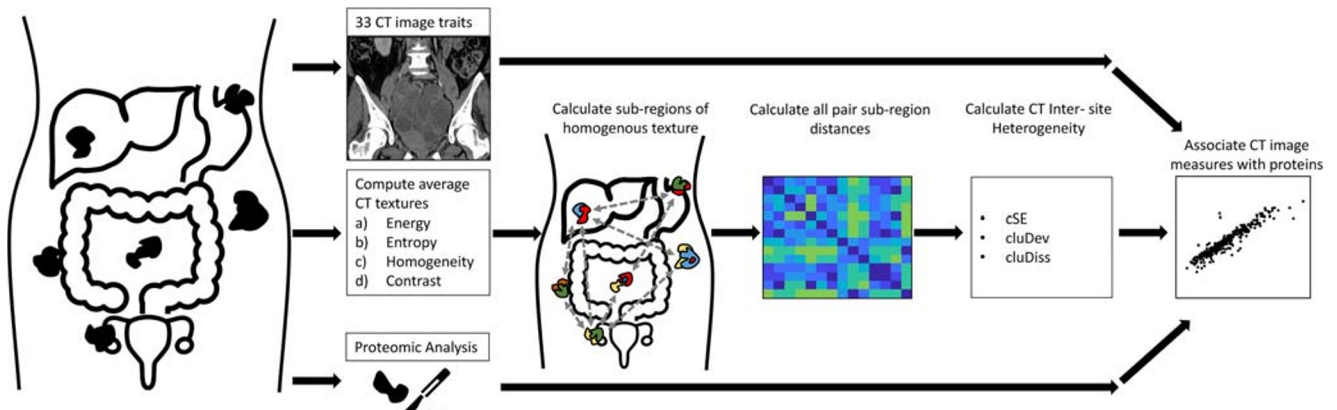
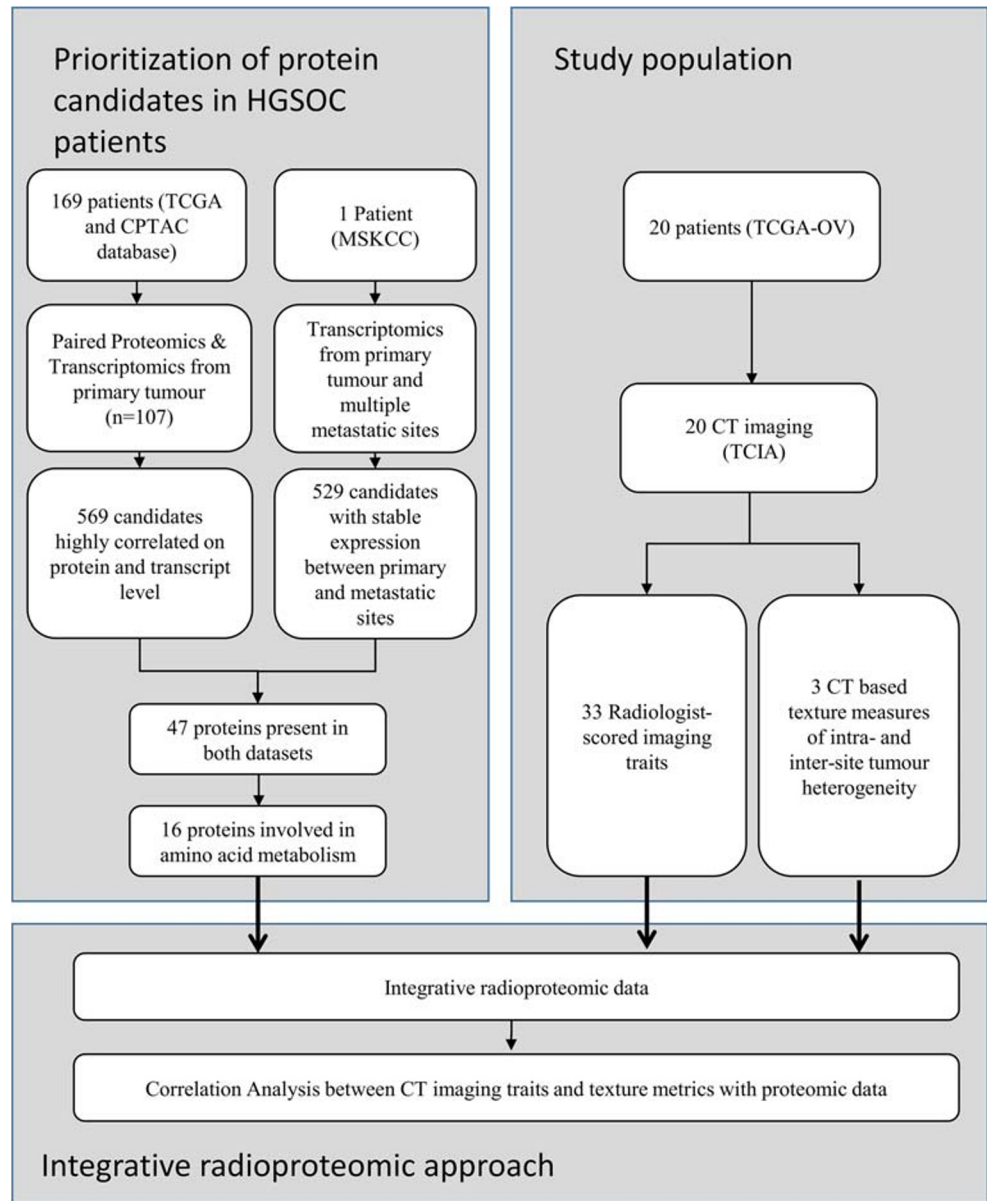


Fig. 2 Overview of the experimental workflow. For each patient, 33 imaging traits, proteomics from a single-site tumor biopsy, as well as intra- and inter-site tumor heterogeneity texture measures, were

obtained. The 33 imaging traits and measures of intra- and inter-site heterogeneity were correlated with proteomic data

Table 1 The CT-based image features used in this study, with possible values, summary statistics and inter-reader agreement

Feature	Possible values and summary statistics	Inter-reader agreement (%); Cohen's kappa
Disease in mesentery	Yes: 50% (10/20) No: 50% (10/20)	95%; 0.90
Peritoneal disease in paracolic gutters	Yes: 80% (16/20) No: 20% (4/20)	100%; 1.0
Peritoneal disease in Pouch of Douglas	Yes: 75% (15/20) No: 25% (5/20)	85%; 0.57
Peritoneal disease in spleen/left upper quadrant	Yes: 85% (17/20) No: 15% (3/20)	100%; 1.0
Peritoneal disease in lesser sac	Yes: 30% (6/20) No: 70% (14/20)	90%; 0.76
Peritoneal disease in liver/right upper quadrant	Yes: 70% (14/20) No: 30% (6/20)	100%; 1.0
Number of locations with peritoneal disease	0: 10% (2/20) 1: 5% (1/20) 2: 5% (1/20) 3: 5% (1/20) 4: 30% (6/20) 5: 20% (4/20) 6: 25% (5/20)	90%; 0.62
Peritoneal disease calcifications	Yes: 5% (1/20) No: 95% (19/20)	95%; N/A
Peritoneal disease omental implant	Yes: 95% (19/20) No: 5% (1/20)	90%; 0.5
Shape of peritoneal disease	Diffuse: 15% (3/20) Predominantly diffuse: 50% (10/20) Predominantly nodular: 30% (6/20) Nodular: 0% (0/20) Peritoneal enhancement only: 0% (0/20) No peritoneal disease: 5% (1/20)	90%; 0.71
Infrarenal retroperitoneal lymphadenopathy	Yes: 30% (6/20) No: 70% (14/20)	100%; 1.0
Pelvic lymphadenopathy	Yes: 15% (3/20) No: 85% (17/20)	95%; 0.77
Inguinal lymphadenopathy	Yes: 0% (0/20) No: 100% (20/20)	100%; N/A
Porta/cealic/gastro lymphadenopathy	Yes: 25% (5/20) No: 75% (15/20)	100%; 1.0
Retrocruial lymphadenopathy	Yes: 5% (1/20) No: 95% (19/20)	95%; N/A
Supradiaphragmatic lymphadenopathy	Yes: 45% (9/20) No: 55% (11/20)	100%; 1.0
Thoracic lymphadenopathy	Yes: 0% (0/20) No: 100% (20/20)	100%; N/A
Number of locations with lymphadenopathy	0: 45% (9/20) 1: 15% (3/20) 2: 20% (4/20) 3: 5% (1/20) 4: 15% (3/20)	100%; 1.0
Metastases in liver	Yes: 0% (0/20) No: 100% (20/20)	100%; N/A
Metastases in lung	Yes: 5% (1/20) No: 95% (19/20)	100%; 1.0
Metastases in pleura	Yes: 10% (2/20) No: 90% (18/20)	100%; 1.0
Metastases in spleen	Yes: 0% (0/20) No: 100% (20/20)	100%; N/A
Metastases in other locations	Yes: 0% (0/20) No: 100% (20/20)	100%; N/A
Number of locations with metastases	0: 85% (17/20) 1: 15% (3/20)	100%; 1.0

Table 1 (continued)

Feature	Possible values and summary statistics	Inter-reader agreement (%); Cohen's kappa
Metastases calcifications	Yes: 0% (0/20) No: 100% (20/20)	100%; N/A
Size of pleural effusions	No pleural effusions: 65% (13/20) Small: 20% (4/20) Moderate/Large: 15% (3/20)	95%; 0.91
Pleural metastases removed	Pleural effusion: 5% (1/20) [NA]: 95% (19/20)	100%; 1.0
Pleural effusion metastases removed	Small: 5% (1/20) [NA]: 95% (19/20)	100%; 1.0
Size of ascites	No ascites: 15% (3/20) Trace or small: 40% (8/20) Moderate or large: 35% (9/20)	95%; 0.92
Mass laterality	No mass: 10% (2/20) Left: 5% (1/20) Right: 5% (1/20) Bilateral: 85% (16/20)	90%; 0.73
Mass calcifications	Yes: 10% (2/20) No: 90% (18/20)	95%; 0.64
Mass septations	Yes: 10% (2/20) No: 90% (18/20)	90%; N/A
Mass internal architecture	Cystic: 0% (0/20) Predominantly cystic: 0% (0/20) Mixed: 35% (7/20) Predominantly solid: 55% (11/20) Solid: 10% (2/20)	75%; 0.38
Length of largest lesion	Median: 63 mm Range: 24 mm to 116 mm	N/A

N/A, not applicable. Inter-reader agreement is shown as absolute agreement and Cohen's kappa statistics for the two radiologists. In case of disagreement, a third radiologist acted as arbitrator. In case one or both readers scored all cases the same value, we were not able to calculate kappa values

cloud-based virtual machine using the ClearCanvas with Annotation and Imaging Markup viewing software (Northwestern University) [25]. Image data in DICOM format, and features and segmentations generated for this paper, are available at <https://doi.org/10.7937/TCIA.2019.9stoinfl> [26].

CT-based qualitative imaging features

The CT imaging traits were assessed by two radiologists (E.S. and H.A.V.; 17 and 12 years' experience in ovarian cancer imaging) independently, who were blinded to the patients' clinical, pathological, and proteomic data. The details of the CT imaging trait evaluation are provided in the study by

Vargas et al [17]. Briefly, each reader individually recorded the imaging interpretation on an electronic case report form, which was uploaded to a central server, compiled, and submitted for statistical analysis. Both ovarian masses and peritoneal disease were evaluated. If a definable ovarian mass was present, its features were recorded, including laterality, maximum size, internal architecture, and presence of calcifications and/or septations. The presence or absence of definable peritoneal implants and their specific location (small bowel mesentery, omentum, paracolic gutters, Pouch of Douglas, perisplenic/left upper quadrant region, lesser sac, and perihepatic/right upper quadrant region) was also recorded. Lymphadenopathy was defined as a short-axis dimension above a predefined size cutoff specific to each location or specific imaging appearance, as detailed in Vargas et al [21].

In the event, the two radiologists disagreed on a categorical feature (e.g., presence of peritoneal disease in the mesentery), and a third radiologist (L.B. 5th year of training) served as an arbitrator. For quantitative features, such as lesion size, the measurements of the two radiologist were averaged. The number of sites in which peritoneal disease was present was also recorded and was considered an additional human-read feature. This was also done for the presence of lymphadenopathy and intra-parenchymal metastases.

Table 2 The texture features used in this study, with summary statistics

Feature	Distribution
Cluster site entropy	Median: 3.724 Range: 2.170 to 4.969
Cluster standard deviation	Median: 2.271 Range: 0.241 to 13.32
Cluster dissimilarity	Median: 4394 Range: 483.8 to 15,857

CT-based texture analysis

A detailed description of image segmentation and texture feature extraction are provided in Veeraraghavan et al [22]. A slice thickness of 5 mm was used for analysis. kVp were variable, as CT machines from different vendors in different institutions were used. In brief, two oncologic imaging research fellows with 4 and 6 years of experience (M.M., S.N.), respectively, in consensus, manually segmented primary ovarian tumor(s) and all metastatic tumor implants in the abdomen and pelvis. Segmentation was performed using 3DSlicer [27] by tracing the contour of each lesion on each slice to produce the volumes of interest (VOI). In total, tumor burden in 14 distinct anatomical regions was manually segmented. Voxel-wise Haralick textures (energy, entropy, contrast, and homogeneity) were computed from within the manually delineated VOIs using in-house software implemented in C++ using the Insight ToolKit [28]. Site-specific sub-regions were computed by voxel-wise clustering of the Haralick textures using the kernel K-means method [29]. Following clustering, tumor sites were divided into distinct sub-regions of similar texture features. The intra- and inter-site heterogeneity (IISTH) measures (i.e., cluster site entropy, cluster standard deviation, and cluster dissimilarity) were computed. IISTH measures summarized the heterogeneity across all tumor sites in each patient. The more different each site of disease, the higher are the IISTH measures.

Protein abundance measurements

Given that the molecular profiles of TCGA patients were based on primary ovarian tumor specimens and our radiologic measurements were made on metastatic patterns of disease, we sought to identify proteins conserved between primary and metastatic tumors in an ovarian cancer patient and then cross-reference these with proteins associated with radiologic patterns of metastasis and proteomic profiles in TCGA. Protein relative abundance measurements from the CPTAC analysis of 107 TCGA HGSOC patients were available for 3377 proteins (normalized log-ratios, standardized across two institutions), which were obtained from Zhang et al [11]. When multiple measurements were available, their abundances were averaged.

As we included only 20 patients in this study, a simple correlation analysis between the protein abundance of 3377 proteins and the CT imaging traits and texture features would result in a high number of false-positive results. Therefore, the purpose of the workflow described below was to reduce the number of proteins used for the final analysis. First, the 3377 proteins were prioritized for inclusion in the analysis, based on evidence of conserved transcript expression between primary and metastatic sites and on agreement between transcript and protein abundance levels. Transcript expression (Affymetrix) data (Supplemental Table 2 in reference [30]) from the

primary tumor and multiple metastatic sites (spleen, right-upper-quadrant, liver, and vaginal cuff) from an HGSOC patient [30], along with transcript expression (mRNA-seq) data from primary HGSOC tumors ($n = 107$) [31] with matching protein expression data from the same case set ($n = 107$) [11], were merged by gene name, yielding a final matrix of 5504 co-measured proteins and transcripts (Supplemental Table 1, in reference [11]). Second, coefficients of variation (CV) values in transcript expression between the primary and metastatic sites, as reported by Jimenez-Sanchez et al [30], and Spearman correlations calculated for transcript and protein expression from the $n = 107$ patients, were normalized using the boxCox function in R (version 3.3.2). Transcripts that exhibited a low CV between the primary and metastatic sites (i.e., stable expression independent of metastatic loci), and with expression levels highly correlated to their proteins, were selected by comparison with normal distributions using T test analyses and a significance threshold of a p value < 0.1 . These analyses identified 529 candidates with conserved expression between primary and metastatic sites and 569 candidates highly correlated at the protein and transcript level; 47 of these intersected between these two analyses. To further reduce the number of proteins, we prioritized proteins based on their molecular function and their interaction with each other using MetScape [32]. Sixteen proteins were associated with the regulation of the amino acid metabolism ($p < 0.001$). The selection of these 16 proteins was done solely on the basis their summary statistics (coefficients of variation in transcript expression between primary and metastatic sites, correlations between transcript and protein expression) and molecular function while correlations of imaging features were not considered in any way in selecting these proteins. Finally, these 16 candidates were included in subsequent analyses to identify those that further correlated with clinical imaging features (Fig. 1). The 16 proteins are given in Supplementary Table 1.

Statistical methods

The percentage agreement and the Cohen's kappa coefficient used to assess the agreement between the two radiologists were calculated for all CT-based imaging features. Univariate associations were assessed between the expression levels of the proteins and CT-based image features in which previous results provided evidence of prognostic power (presence of peritoneal disease in the mesentery and diffuse—non-mass forming—peritoneal involvement, as found in the study by Vargas et al [21]; number of sites with peritoneal disease; presence of peritoneal disease in the liver; presence of Pouch of Douglas implants; presence of supradiaphragmatic lymphadenopathy; and non-visualization of a discrete ovarian mass). For binary image features, these associations were assessed using a Mann-Whitney U test to compare the expression levels of a specific protein among patients with the feature

to those among patients without the feature. The area under the receiver operating characteristic curve (AUC) was reported as a metric of the strength and direction of this association. For quantitative features, these associations were assessed through inferences on the Kendall tau rank correlation coefficient.

Univariate associations between the texture radiomic features (cluster site entropy, cluster standard deviation, and cluster dissimilarity) and protein relative abundance were assessed through inferences on the Kendall tau rank correlation coefficient. The Benjamini-Hochberg procedure was used to correct for multiple hypothesis testing. All tests were performed at the $\alpha = 0.05$ level.

Results

Patient characteristics and distribution of disease

The patients' characteristics are shown in Table 3. The median number of tumor sites was six (range, 1–11). The median tumor volume was 192 cm³ (range, 12–2574 cm³), the highest being the tumor volume of the ovarian masses, with a median of 80 cm³ (range, 0–600 cm³). Figure 3 illustrates the tumor volumes for each anatomical sub-region.

Inter-reader agreement of CT-based imaging traits

There was a good to excellent inter-reader agreement for CT-based imaging traits, with an absolute agreement ranging between 75% ($\kappa = 0.32$) for the variable mass internal architecture to 100% ($\kappa = 1.0$) agreement for the majority of variables (Table 1).

Table 3 Clinical characteristics of the 20 patients used in the analysis

Characteristic	
Median age (year)	60.5 (49–80)
FIGO stage III	8 (40%)
FIGO stage IV	12 (60%)
Residual disease following primary debulking	
None	5 (25%)
Less than 1 cm	11 (55%)
Greater than 1 cm	3 (15%)
No. of patients with missing data	1 (5%)
Median TTP (day)	426 (9–1475)
Median OS	1469 (9–2749)

Age, TTP, and OS are shown as median (range), while the remaining variables are shown as number of patients; TTP, time to progression; OS, overall survival

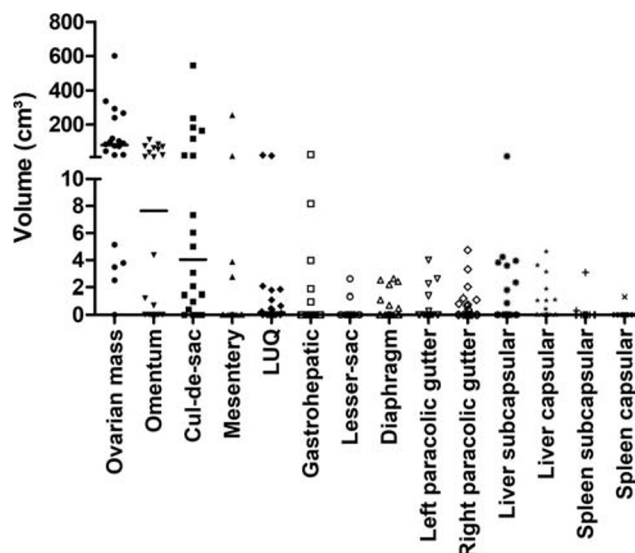


Fig. 3 Tumor volume of the sub-regions of the 20 patients included in the final analysis is shown. LUQ, left upper quadrant; the horizontal line indicates the median value

CT imaging traits are associated with an abundance of several proteins

Four CT imaging traits were associated with an abundance of several proteins (Table 4). For example, disease in the mesentery was associated with reduced levels of CRIP2 ($p = 0.0002$, AUC = 0.05) and GPI ($p = 0.004$, AUC = 0.22). Supradiaphragmatic lymphadenopathy was associated with decreased abundance of ALDH3A2 ($p = 0.02$, AUC 0.192) and increased abundance of MAGEA4 ($p = 0.046$, AUC = 0.768). The number of sites with peritoneal disease was correlated with increased levels of ALDH3A2 ($p = 0.03$, $\tau = 0.36$) and CRIP2 ($p = 0.01$, $\tau = 0.378$). The shape of peritoneal disease was associated with the abundance of MAGEA4 ($\tau = 0.04$, $\tau = -0.343$). After correction for multiple testing through the Benjamini-Hochberg procedure, the correlation between CRIP2 and the presence of supradiaphragmatic lymphadenopathy remained significant (corrected $p = 0.03$). Patients with without or diffuse peritoneal disease had higher levels of MAGEA4 (median 2.31 and 2.31 respectively) than patients with predominantly diffuse or predominantly nodular disease (median -0.233 and -1.51 respectively).

CT-based intra- and inter-site heterogeneity metrics were associated with an abundance of several proteins

The protein abundance of three proteins was associated with intra- and inter-site tumor heterogeneity texture metrics (Table 5). For example, cluster site entropy was positively correlated with the abundance of STXBP2 ($p = 0.007$, $\tau = 0.432$) and negatively with ASS1 ($p = 0.011$, $\tau = -0.364$). Cluster standard deviation was positively correlated with an

Table 4 Associations between protein abundance and CT-based imaging traits previously found to be prognostic in patients with HGSOc. Only associations for which the *p* value was less than 0.05 were included in this table. AUC values greater than 0.5 indicate a positive association between the imaging trait and the level of

expression of the protein whereas AUC values less than 0.5 indicate a negative association. Positive values of τ indicate positive association between the imaging trait and the level of expression of the protein whereas negative values of indicate negative association

Image trait	Protein	<i>p</i> value	Corr. <i>p</i> value	Point estimate of metric of association
Disease in mesentery	CRIP2	0.0002	0.03	AUC = 0.05
	GPI	0.04	0.51	AUC = 0.22
Supradiaphragmatic lymphadenopathy	ALDH3A2	0.02	0.45	AUC = 0.192
	MAGEA4	0.046	0.52	AUC = 0.768
Number of sites with peritoneal disease	ALDH3A2	0.03	0.51	$\tau = 0.36$
	CRIP2	0.01	0.29	$\tau = 0.439$
Shape of peritoneal disease	MAGEA4	0.04	0.52	

STXBP2 ($p = 0.05$, $\tau = 0.453$). Cluster dissimilarity was positively correlated with STXBP2 ($p = 0.03$, $\tau = 0.368$) and negatively with ASS1 ($p = 0.009$, $\tau = -0.427$) and CBK ($p = 0.047$, $\tau = -0.326$). However, none of these associations were significant after correction for multiple testing through the Benjamini-Hochberg procedure.

Discussion

In this hypothesis-generating study, we investigated the relationship of CT imaging traits and CT-based texture measures of tumor heterogeneity to protein abundance in patients with HGSOc. This study is part of the efforts of the NCI initiative to combine proteomic data available from TCGA with patient-matched CT images collected as part of the TCIA effort in ovarian cancer. Our results provide preliminary evidence that suggests possible associations between imaging traits, CT-based texture measures of tumor heterogeneity, and the abundance of several proteins. These results are a step forward to the development of models that integrate clinical, proteomic, and radiomic data to predict meaningful clinical endpoints and facilitate tailored therapies.

Tumor involvement of the mesentery, which is a known important limiting factor in primary debulking surgery, was

negatively correlated with the protein abundance of cysteine rich protein 2 (CRIP2). This correlation was even significant after correction for multiple testing. CRIP2 regulates cell proliferation, and acts as a tumor suppressor [33, 34]. In addition, the presence of supradiaphragmatic lymphadenopathy that causes patients to be up-staged to stage IV disease was positively associated with the protein abundance of MAGE family member A4 (MAGE4). Increased MAGE4 expression in ovarian cancer cells is an independent predictor for mortality in associated with worse overall survival [35].

This study also provides the first insights into possible associations between measures of CT-based tumor heterogeneity and protein abundance. The most interesting associations was with argininosuccinate synthase 1 (ASS1). HGSOc cells are known to express high levels of ASS1 [36, 37], and reduced levels of this protein correlate with platinum-based drug resistance in vitro and in vivo, and with worse prognosis [38]. We found that more heterogeneous tumors had lower ASS1 expression in with more homogenous tumors. This would support previous finding higher tumor heterogeneity measured by CT is predictive for worse survival [18].

A few studies, thus far, have applied radiomics in patients with ovarian cancer, mainly evaluating the ability of texture features to characterize tumor tissue [15, 17, 39, 40] and predict outcome [14, 15, 17, 18, 21, 22, 39]. Best to our knowledge,

Table 5 Association between CT-based texture metrics and protein abundance in patients with HGSOc. Only associations for which the *p* value was less than 0.05 were included in this table. Positive values of τ

indicate positive association between the texture metric and the level of expression of the protein whereas negative values of τ indicate negative association

Texture feature	Protein	<i>p</i> value	Corr. <i>p</i> value	Point estimate of metric of association
Cluster site entropy	ASS1	0.03	0.45	$\tau = -0.364$
	STXBP2	0.007	0.28	$\tau = 0.432$
Cluster standard deviation	STXBP2	0.005	0.20	$\tau = 0.453$
Cluster dissimilarity	ASS1	0.009	0.29	$\tau = -0.427$
	CKB	0.047	0.52	$\tau = -0.326$
	STXBP2	0.03	0.45	$\tau = 0.368$

our work is the first that combines proteomic and radiomic data in patients with HGSOE, thereby facilitating the integration of multiple layers of data that is needed for the realization of precision medicine in oncological patients [20].

Our study was primarily limited by the retrospective design, the small study population, and the lack of a validation cohort. Due to the small sample size, there are inevitable biases in extracting CT textural features, with the risk of both false-positive and false-negative results. Therefore, we a priori selected only three texture metrics to assess IISTH, which had already been validated in larger studies to be predictive of treatment response and survival in patients with HGSOE [18, 19]. In addition, the process of selecting proteins may have screened out proteins that may be important. However, we think that this approach is justified, as the main aim of this study was to reduce the number of selected proteins to avoid false-positive correlations between protein expression and CT data. A larger study is needed for a more definitive analysis on the association between imaging traits and texture metric and CLOVAR subtypes, as well as from integrating the imaging, genomics, and proteomics data to predict response to treatment and outcome. However, since the TCGA/TCIA has been the only publicly available cohort of HGSOE patients with imaging, proteomics, and genomics data to date, this study was designed as a small hypothesis-generating study rather than a confirmative study.

In conclusion, this study provides the first insights into the potential associations between CT imaging traits and CT-based texture measures of tumor burden heterogeneity and the abundance of several tumor-associated proteins. Future, larger, prospective studies, such as that being conducted in the Applied Proteogenomic Organizational Learning and Outcomes (APOLLO) Research Network, are needed to validate our findings and enable the integration of clinical data, proteomics and radiomics to guide tailored therapies.

Funding information This work was supported by The Mark Foundation for Cancer Research and Cancer Research UK Cambridge Centre (C9685/A25177) and the US Department of Defense - Uniformed Services University of the Health Sciences (HU0001-16-2-0006). In addition, this project has been funded in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E. Data used in this publication were generated by the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). The results published here are in whole or part based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Prof. Evis Sala.

Conflict of interest The authors of this manuscript declare relationships with the following companies: ES: Siemens Speakers Bureau. The other authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors has significant statistical expertise (EH).

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Study subjects or cohorts overlap Some study subjects or cohorts have been previously reported in:

Vargas HA, Miccò M, Hong SI et al (2014) 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.

Vargas HA, Veeraraghavan H, Micco M et al (2017) A novel representation of inter-site tumor heterogeneity from pre-treatment computed tomography textures classifies ovarian cancers by clinical outcome. *Eur Radiol* 27:3991–4001.

Vargas HA, Huang EP, Lakhman Y et al (2017) Radiogenomics of High-Grade Serous Ovarian Cancer: Multireader Multi-Institutional Study from the Cancer Genome Atlas Ovarian Cancer Imaging Research Group. *Radiology* 285:482–492.

Harini Veeraraghavan, Hebert Alberto Vargas, Alejandro Jimenez Sanchez et al (2019) Computed Tomography Measures of Inter-site tumor Heterogeneity for Classifying Outcomes in High-Grade Serous Ovarian Carcinoma: a Retrospective Study. *bioRxiv* 531,046; <https://doi.org/10.1101/531046>

Methodology

- Retrospective
- Observational
- Multicenter study

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.


References

1. Siegel RL, Miller KD, Jemal A (2017) Cancer statistics, 2017. *CA Cancer J Clin* 67:7–30
2. Winter WE 3rd, Maxwell GL, Tian C et al (2007) Prognostic factors for stage III epithelial ovarian cancer: a gynecologic oncology group study. *J Clin Oncol* 25:3621–3627
3. Agarwal R, Kaye SB (2005) Prognostic factors in ovarian cancer: how close are we to a complete picture? *Ann Oncol* 16:4–6
4. Yoshihara K, Tsunoda T, Shigemizu D et al (2012) High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by downregulation of antigen presentation pathway. *Clin Cancer Res* 18:1374–1385
5. Kang J, D'Andrea AD, Kozono D (2012) A DNA repair pathway-focused score for prediction of outcomes in ovarian cancer treated

- with platinum-based chemotherapy. *J Natl Cancer Inst* 104:670–681
6. Verhaak RG, Tamayo P, Yang JY et al (2013) Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. *J Clin Invest* 123:517–525
 7. Parkinson CA, Gale D, Piskorz AM et al (2016) Exploratory analysis of TP53 mutations in circulating tumour DNA as biomarkers of treatment response for patients with relapsed high-grade serous ovarian carcinoma: a retrospective study. *PLoS Med* 13:e1002198
 8. Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature* 474:609
 9. Ahmed AA, Etemadmoghadam D, Temple J et al (2010) Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J Pathol* 221:49–56
 10. Yang JY, Yoshihara K, Tanaka K et al (2013) Predicting time to ovarian carcinoma recurrence using protein markers. *J Clin Invest* 123:3740–3750
 11. Zhang H, Liu T, Zhang Z et al (2016) Integrated proteogenomic characterization of human high-grade serous ovarian cancer. *Cell* 166:755–765
 12. Zhang H, Mao Y, Chen X et al (2019) Magnetic resonance imaging radiomics in categorizing ovarian masses and predicting clinical outcome: a preliminary study. *Eur Radiol* 29:3358–3371
 13. Danala G, Thai T, Gunderson CC et al (2017) Applying quantitative CT image feature analysis to predict response of ovarian cancer patients to chemotherapy. *Acad Radiol* 24:1233–1239
 14. Rizzo S, Botta F, Raimondi S et al (2018) Radiomics of high-grade serous ovarian cancer: association between quantitative CT features, residual tumour and disease progression within 12 months. *Eur Radiol* 28:4849–4859
 15. Lu H, Arshad M, Thornton A et al (2019) A mathematical-descriptor of tumor-mesoscopic-structure from computed-tomography images annotates prognostic- and molecular-phenotypes of epithelial ovarian cancer. *Nat Commun* 10:764
 16. Wei W, Liu Z, Rong Y et al (2019) A computed tomography-based radiomic prognostic marker of advanced high-grade serous ovarian cancer recurrence: a multicenter study. *Front Oncol* 9:255
 17. Vargas HA, Miccò M, Hong SI et al (2014) 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
 18. Vargas HA, Veeraraghavan H, Micco M et al (2017) A novel representation of inter-site tumour heterogeneity from pre-treatment computed tomography textures classifies ovarian cancers by clinical outcome. *Eur Radiol* 27:3991–4001
 19. Himoto Y, Veeraraghavan H, Zheng J et al (2019) Computed tomography-derived radiomic metrics can identify responders to immunotherapy in ovarian cancer. *JCO Precis Oncol*. <https://doi.org/10.1200/po.19.00038>
 20. Pinker K, Chin J, Melsaether AN, Morris EA, Moy L (2018) Precision medicine and radiogenomics in breast cancer: new approaches toward diagnosis and treatment. *Radiology* 287:732–747
 21. Vargas HA, Huang EP, Lakhman Y et al (2017) Radiogenomics of high-grade serous ovarian cancer: multireader multi-institutional study from the cancer Genome Atlas Ovarian Cancer Imaging Research Group. *Radiology* 285:482–492
 22. Veeraraghavan H, Vargas HA, Sanchez AJ et al (2019) Computed tomography measures of inter-site tumor heterogeneity for classifying outcomes in high-grade serous ovarian carcinoma: a retrospective study. *bioRxiv*. <https://doi.org/10.1101/531046>
 23. Holback C, Jarosz R, Prior F et al (2016) Radiology data from The Cancer Genome Atlas Ovarian Cancer [TCGA-OV] collection. The Cancer Imaging Archive
 24. Clark K, Vendt B, Smith K et al (2013) The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 26:1045–1057
 25. Mongkolwat P, Kleper V, Talbot S, Rubin D (2014) The National Cancer Informatics Program (NCIP) Annotation and Image Markup (AIM) Foundation model. *J Digit Imaging* 27:692–701
 26. Beer L, Sahin H, Blazic I et al (2019) Data from integration of CT-based qualitative and radiomic features with proteomic variables in patients with high-grade serous ovarian cancer: an exploratory analysis. The Cancer Imaging Archive. <https://doi.org/10.7937/TCIA.2019.9STOINF1>
 27. Fedorov A, Beichel R, Kalpathy-Cramer J et al (2012) 3D slicer as an image computing platform for the quantitative imaging Network. *Magn Reson Imaging* 30:1323–1341
 28. Yoo TS, Ackerman MJ, Lorensen WE et al (2002) Engineering and algorithm design for an image processing Api: a technical report on ITK—the insight Toolkit. *Stud Health Technol Inform* 85:586–592
 29. Dhillon IS, Guan Y, Kulis B (2004) Kernel k-means: spectral clustering and normalized cuts Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 551–556
 30. Jimenez-Sanchez A, Memon D, Pourpe S et al (2017) Heterogeneous tumor-immune microenvironments among differentially growing metastases in an ovarian cancer patient. *Cell* 170(927–938):e920
 31. Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature* 474:609–615
 32. Karnovsky A, Weymouth T, Hull T et al (2011) Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. *Bioinformatics* 28:373–380
 33. Cheung AKL, Ko JM, Lung HL et al (2011) Cysteine-rich intestinal protein 2 (CRIP2) acts as a repressor of NF- κ B-mediated proangiogenic cytokine transcription to suppress tumorigenesis and angiogenesis. *Proc Natl Acad Sci U S A* 108:8390–8395
 34. Zhou L, Wang Y, Zhou M et al (2018) HOXA9 inhibits HIF-1 α -mediated glycolysis through interacting with CRIP2 to repress cutaneous squamous cell carcinoma development. *Nat Commun* 9:1480
 35. Yakirevich E, Sabo E, Lavie O, Mazareb S, Spagnoli GC, Resnick MB (2003) Expression of the MAGE-A4 and NY-ESO-1 cancer-testis antigens in serous ovarian neoplasms. *Clin Cancer Res* 9:6453–6460
 36. Cheon DJ, Walts AE, Beach JA et al (2015) Differential expression of argininosuccinate synthetase in serous and non-serous ovarian carcinomas. *J Pathol Clin Res* 1:41–53
 37. Coscia F, Watters KM, Curtis M et al (2016) Integrative proteomic profiling of ovarian cancer cell lines reveals precursor cell associated proteins and functional status. *Nat Commun* 7:12645
 38. Nicholson LJ, Smith PR, Hiller L et al (2009) Epigenetic silencing of argininosuccinate synthetase confers resistance to platinum-induced cell death but collateral sensitivity to arginine auxotrophy in ovarian cancer. *Int J Cancer* 125:1454–1463
 39. Zhang H, Mao Y, Chen X et al (2019) Magnetic resonance imaging radiomics in categorizing ovarian masses and predicting clinical outcome: a preliminary study. *Eur Radiol*. <https://doi.org/10.1007/s00330-019-06124-9>
 40. Meier A, Veeraraghavan H, Nougaret S et al (2018) Association between CT-texture-derived tumor heterogeneity, outcomes, and BRCA mutation status in patients with high-grade serous ovarian cancer. *Abdom Radiol (NY)*. <https://doi.org/10.1007/s00261-018-1840-5>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Lucian Beer¹ · Hilal Sahin¹ · Nicholas W. Bateman^{2,3} · Ivana Blazic⁴ · Hebert Alberto Vargas⁵ · Harini Veeraraghavan⁶ · Justin Kirby⁷ · Brenda Fevrier-Sullivan⁷ · John B. Freymann⁷ · C. Carl Jaffe⁸ · James Brenton^{9,10} · Maura Miccò¹¹ · Stephanie Nougaret¹² · Kathleen M. Darcy^{2,3} · G. Larry Maxwell^{2,3,13} · Thomas P. Conrads^{2,3,13,14} · Erich Huang¹⁵ · Evis Sala^{1,16} 

¹ Department of Radiology, Cancer Research UK Cambridge Center, Cambridge CB2 0QQ, UK

² Department of Obstetrics and Gynecology, Gynecologic Cancer Center of Excellence, Walter Reed National Military Medical Center, Uniformed Services University, 8901 Wisconsin Avenue, Bethesda, MD 20889, USA

³ The John P. Murtha Cancer Center, Walter Reed National Military Medical Center, Uniformed Services University, 8901 Wisconsin Avenue, Bethesda, MD 20889, USA

⁴ Department of Radiology, Clinical Hospital Center Zemun, Vukova 9, Belgrade 11080, Serbia

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

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

⁷ Cancer Imaging Informatics Lab, Frederick National Laboratory for Cancer Research, Frederick, MD, USA

⁸ Department of Radiology, Boston University School of Medicine, Boston, MA 02118, USA

⁹ Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge, Cambridgeshire, UK

¹⁰ Cancer Research UK Cambridge Centre, Cambridge, Cambridgeshire, UK

¹¹ Dipartimento Diagnostica per Immagini, Radiologia Diagnostica e Interventistica Generale, Area Diagnostica per Immagini, Radioterapia Oncologica ed Ematologia, Rome, Italy

¹² Department of Radiology, Montpellier Cancer Institute, INSERM, University of Montpellier, Montpellier, France

¹³ Department of Obstetrics and Gynecology, Inova Fairfax Medical Campus, 3300 Gallows Rd., Falls Church, VA 22042, USA

¹⁴ Inova Center for Personalized Health, Inova Schar Cancer Institute, 3300 Gallows Rd., Falls Church, VA 22042, USA

¹⁵ Biometric Research Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, NIH, Rockville, MD 20850, USA

¹⁶ Department of Radiology, University of Cambridge, Box 218, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK