Radiogenomic analysis of cervical lymph node metastasis in papillary thyroid carcinoma: A preliminary study

Yuyang Tong

Department of Medical Ultrasound, Fudan University Shanghai Cancer Centre, Shanghai 200032, China
Department of Surgical Oncology, The Ohio State University, 460 West 12th Avenue, Columbus, Ohio, 43210, USA <https://orcid.org/0000-0001-8209-4271>

Peixuan Sun

Diagnostic Imaging Centre, Shanghai Children's Medical Centre, Shanghai Jiao Tong University School of Medicine, Shanghai, 200127, China

Yunxia Huang

Department of Medical Ultrasound, Fudan University Shanghai Cancer Centre, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China

Jin Zhou

Department of Medical Ultrasound, Fudan University Shanghai Cancer Centre, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China

Juanjuan Yong

Department of Pathology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, 107 West Yanjiang Road, Guangzhou, Guangdong, 510120, China

Yi Guo

Department of Electronic Engineering, Fudan University and Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention of Shanghai, Shanghai, 200433, China

Jinhua Yu

Department of Electronic Engineering, Fudan University and Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention of Shanghai, Shanghai, 200433, China

Yuan Yuan Wang

Department of Electronic Engineering, Fudan University and Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention of Shanghai, Shanghai, 200433, China

Shichong Zhou (✉ sczhou@hotmail.com)

Department of Medical Ultrasound, Fudan University Shanghai Cancer Centre, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China

Cai Chang

Department of Medical Ultrasound, Fudan University Shanghai Cancer Centre, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China

Research

Keywords: Radiogenomic, Papillary thyroid carcinoma, Cervical lymph node metastasis, Ultrasound

Posted Date: May 29th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-31054/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Papillary thyroid carcinoma (PTC) is characterized by frequent metastasis to cervical lymph nodes (CLNs), and the presence of lymph node metastasis at diagnosis has a significant impact on the surgical approach. Therefore, we established a radiomic signature to predict the CLN status of PTC patients using preoperative thyroid ultrasound and investigated the association between the radiomic features and underlying molecular characteristics of PTC tumours.

Methods

A radiogenomic map linking radiomics features to gene modules was constructed, and immunohistochemistry was performed to validate key associations. In all, 180 patients were enrolled in this prospective study, and 47 radiomic features, including tumour size, shape, position, margin, echo pattern and calcification, were extracted. Total protein extracted from 49 tumour samples was analysed with LC/MS and iTRAQ technology. Gene modules acquired by clustering were chosen for their diagnostic significance. A radiogenomic map linking radiomic features to gene modules was constructed with the Spearman correlation matrix. Immunohistochemistry was performed to validate key associations between radiomic features and gene modules.

Results

The diagnostic performance of radiomics signature was better than that of the ultrasound-based method in predicting CLN status. Weighted gene co-expression network analysis generated 16 gene modules, and a radiogenomic map with nine significant correlations between radiomics features and gene modules was created. For example, the feature 'minimum calcification area' was significantly associated with module MEblue, which represents cell-cell adhesion and glycolysis. Immunohistochemistry showed that LAMC1 and THBS1 expression was associated with several radiomics features.

Conclusions

The radiomic signature proposed here has the potential to noninvasively predict the CLN status in PTC patients. Merging imaging phenotypes with genomic data could allow the noninvasive identification of the molecular properties of PTC tumours, which would support clinical decision making with personalized management.

Trail registration:

This prospective study was approved by the ethics committee of the Fudan University Shanghai Cancer Centre (1809191-1-NSFC) in July 2018.

Background

Papillary thyroid carcinoma (PTC), as the most common malignant thyroid cancer, accounts for approximately 90% of all thyroid malignancies [1]. PTC is characterized by frequent metastasis to cervical lymph nodes (CLNs), which occurs in approximately 30–90% of PTC cases [2]. Cervical lymph node metastasis (CLNM) increases the likelihood of local recurrence and is associated with decreased survival in the high-risk group [3, 4]. The presence of LNM at diagnosis has a significant impact on the surgical approach [5]. Therefore, preoperative evaluation of the CLNM status is of paramount significance in designing an optimal therapeutic schedule and assessing the prognosis of PTC patients [6].

Fine-needle aspiration (FNA) is an alternative way to confirm LN involvement; however, FNA is an invasive procedure with a risk of bleeding and infection, in addition to being expensive and time consuming [7]. Ultrasound (US) examination of the CLNs before surgery is recommended by the American Thyroid Association for patients with known or suspected thyroid nodules [8]. Unfortunately, preoperative US has a relatively low sensitivity for detecting CLNM. In particular, some small metastases or metastases in the central compartment may not appear as abnormal findings on US images [9]. It is universally acknowledged that the development of metastasis depends on the biological behaviour of the primary tumour. Thus, preoperatively predicting the CLNM status by investigating the biological characteristics of PTC tumours should be pursued.

An emerging discipline called radiogenomics aims to integrate radiological imaging data with genomic data, which provides an opportunity to investigate the relationship between radiomic features and the corresponding molecular profile [10]. These noninvasive imaging biomarkers may act as surrogates for molecularly defined features [11]. To date, radiogenomic strategies have been applied in clear renal cell carcinoma [12], lung cancer [13], liver cancer [14] and breast cancer [15] and have attempted to correlate the radiomic features of tumours with DNA mutations, mRNA expression, and underlying signalling pathways. To the best of our knowledge, no published studies have investigated the association between the radiomic features of PTC and the underlying molecular genotype.

In this study, we hypothesized that the radiomic features could potentially reflect the gene expression patterns in patients with PTC. The purposes of our study were to 1) establish a radiomic signature to predict the CLN status of patients with PTC using preoperative thyroid US images; 2) investigate the associations between the radiomic features and the underlying molecular characteristics of the tumour; and 3) validate the key associations via immunohistochemistry (IHC).

Methods

Patients

This prospective study was approved by the ethics committee of the Fudan University Shanghai Cancer Centre (1809191-1-NSFC). Written informed consent was provided by each patient. This study was performed in accordance with the Declaration of Helsinki. From February 2019 to June 2019, 406 consecutive patients with newly diagnosed thyroid tumours at our centre took part in this study. All patients underwent neck ultrasound examination before surgery and obtained the final pathological diagnosis of PTC from surgical specimen. All of the study patients underwent subtotal or total thyroidectomy. Dissection of the central CLNs was routinely performed, and if suspicious lateral LNs were indicated by FNA, preoperative CT or US, lateral LN dissection was performed. The exclusion criteria included the followings: i) preoperative therapy (radiofrequency or microwave ablation, radiotherapy, and chemotherapy); ii) patients with multifocal lesions or bilateral disease; iii) poor imaging quality; iv) pathological reports and follow-up information were not available in clinical medical records. In all, 180 patients were enrolled in this study. The pathological results of the LNs reviewed by senior pathologists with at least 6 years of experience were used as the gold standard for determining the LN status. Figure 1 summarizes the workflow of the study design.

US evaluation of the CLNs

The status of the CLNs was assessed according to the ACR Thyroid Imaging, Reporting and Data System (TI-RADS) [16]. In brief, a globular shape, the lack of a normal echogenic hilum, the presence of peripheral flow, heterogeneity with cystic components, and microcalcification were abnormal findings suggestive of CLNM. Based on these criteria, sonographic evaluation of the CLNs yielded a normal finding or suspected metastasis.

US imaging and segmentation

US images were acquired using ultrasonic equipment from Aixplorer (SuperSonic Imagine, Aix-en-Provence, France) with a 5–14 MHz linear transducer operated by radiologists with more than 6 years of experience. Thyroid US images were recorded and saved in DICOM format, in which tumour images were manually segmented by a radiologist with at least 8 years of experience. For each nodule, only the grey-scale image of the longitudinal section along the longest axis was subjected to segmentation. MATLAB software (MathWorks, Natick, MA, USA) was used for manual segmentation.

Feature extraction

All 47 features recorded on the B-mode US images comprised demographic information and tumour parenchyma-related features, which can reflect the size, shape, position, margin, echo pattern and calcification of the tumour. The texture complexity was represented by grey-level co-occurrence matrix (GLCM)-based features. Details are listed in Additional file 1 (Table S1). All features were extracted from the original thyroid sonographic images, and all image and data processing were performed with MATLAB R2015b software (MathWorks, Inc., Natick, MA, USA). Redundant radiomic features were

eliminated using a correlation matrix method in the radiogenomic correlation part. For each individual feature, the mean absolute correlation based on pairwise correlation was calculated. If a pairwise correlation was beyond 0.8, features with the higher mean absolute correlation were removed. This process was conducted by using the 'findCorrelation' function in the caret package in *R* software.

Classification

There are many common classifiers used in radiomics, and the support vector machine (SVM) is currently widely used for its robustness and stability. In this study, we used an SVM for classification, and a specific mathematical model of the SVM can be written as follows:

$$\min_{\omega, b, \xi} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i$$

$$s. t. y_i(\omega x_i + b) \geq 1 - \xi_i, i = 1, 2, \dots, N,$$

$$\xi_i \geq 0, i = 1, 2, \dots, N,$$

where ξ_i is the slack variable, b is the bias vector, ω is the weight vector, and C is the penalty parameter. In addition, $x_i \in X \in \mathbb{R}^n$, $y_i \in Y \in \mathbb{R}^n$, $i=1, 2, \dots, N$ where x_i is the i th feature, and y_i is the label of x_i .

A receiver operating characteristic (ROC) curve was created to assess the overall performance of the radiomic signature and the US-based method. Other criteria, such as accuracy (ACC), sensitivity (SEN), specificity (SPEC) and area under the ROC (AUC), were calculated to evaluate the ability of the model to discriminate the LN status.

Cross-validation

We evaluated the model by leave-one-out cross-validation (LOOCV). Every time the LOOCV method takes one sample as a testing sample, all the remaining samples are used as a training set. This process was repeated until all samples were tested individually.

Samples and protein extraction

Fifty-five tumour samples were obtained intraoperatively as the specimen was resected from the patient. A tumour section from the largest diameter of the specimen was harvested (the slice thickness varied depending on the overall dimensions of the tumour). For larger nodules, we avoided any areas of overt central necrosis. Samples collected after surgery were frozen in liquid nitrogen until protein extraction. Each sample was homogenized in RIPA buffer and then subjected to ultrasonication. The resulting homogenates were held on ice for 30 min and centrifuged to remove the precipitates. A bicinchoninic acid (BCA) assay was performed for every sample to quantify the protein concentration. Protein digestion was performed according to the filter-aided sample preparation (FASP) strategy described in a previous report [17]. Briefly, 150 µg of protein from each group was reduced, alkylated, and digested in a centrifuge tube.

After the peptide solutions were digested at 37 °C overnight, they were centrifuged, and the filtrates were collected. Finally, the digested peptides were dried by vacuum centrifugation and stored at -80 °C until further use.

iTRAQ labelling

Seventy-five micrograms of peptides from each sample were used for iTRAQ labelling (AB SCIEX, Darmstadt, Germany). The labelling procedure was conducted in line with the manufacturer's protocol with some modifications. After reconstitution in dissolution buffer, the digested peptides were incubated with a specific iTRAQ for 3 hours at room temperature. The labelled samples were homogenously mixed and dried by SpeedVac before they were redissolved in 60 µL of 5 mM ammonium formate. Next, 50 µL of the sample was prefractionated by high-pH reverse-phase liquid chromatography on an ACQUITY UPLC H-Class Bio system (Waters, Milford, MA, USA), and, finally, 10 consolidated fractions were obtained. Labelled peptides in each fraction were dried and redissolved in 30 µL of 2% acetonitrile /0.1% formic acid for LC-MS/MS analysis.

Nano-LC-MS/MS

The labelled peptide mixtures were separated using an EASY-nLC 1000 system (Thermo Fisher Scientific, USA) and then trapped on a PepMap100 C18, 3 µm, 75 µm × 20 mm column (Thermo Fisher Scientific, USA). Then, the peptides were separated on a PepMap100, C18 2 µm 75 µm × 150 mm analytic column (NanoViper, Thermo Fisher Dionex, USA) with a 105 min mobile phase gradient from 5–35%. Mass spectra were recorded on a Q Exactive mass spectrometer equipped with a Nano-ESI source (Thermo Fisher Scientific, USA). Full-scan mass spectra were acquired in the 300–1600 m/z range at a resolution of 70,000, the top 20 precursors were selected for high-energy collision-induced dissociation (HCD) with a collision energy of 27%, and the productions were detected at a resolution of 17,500 on the data-dependent acquisition mode.

iTRAQ data analysis

Protein sequences were searched using the MASCOT engine (version 2.3.2, Matrix Science, London, UK) embedded into Proteome Discoverer Software 1.4 (Thermo Electron, San Jose, USA). For search parameters, the following options were used: 1) database, SwissProt; 2) Taxonomy, *Homo sapiens*; 3) Enzyme, trypsin; 4) Fixed modifications, carbamidomethyl (C), iTRAQ 8-plex (N-term), iTRAQ 8-plex (K); 5) Variable modifications, oxidation (M) and iTRAQ 8-plex (Y); 6) Max missed cleavages, 2; 7) Peptide charge state, 2+, 3+, and 4+; 8) Peptide mass tolerance, 10 ppm; and 9) MS/MS tolerance, ± 0.02 Da. All reported data are based on 99% confidence for protein identification as determined by a false discovery rate (FDR) of ≤ 1%. The procedure was described in more detail by Plubell et al. [18] The data were analysed with the support of Wayen Biotechnologies Co., Ltd. (Shanghai, China).

Identification of gene modules of tumour tissues

The large amount of gene expression data, complex potential combinations and multiple testing issues make traditional analytical methods infeasible. Weighted gene co-expression network analysis (WGCNA)

is a well-established method and has become the standard in systems biology to identify modules—that is, clusters of genes with highly similar expression data. Compared with conventional clustering approaches, WGCNA has been demonstrated to be more robust and reliable because it considers both topological and correlative information. We employed the WGCNA package (version 1.64) in *R* software to identify gene modules. A detailed procedure for WGCNA was published in a previous study [14]. Each gene module, also known as the module eigengene (ME), was summarized by its first principal component of the scaled (standardized) module expression profiles, and the minimum number of genes per module was set to 20.

Clinical relevance of gene modules

To investigate the correlation between gene modules and LNM, ROC curve analysis was used for gene expression in tumour samples. Gene modules that had a relatively high diagnostic performance for LNM (area under ROC curve (AUC) > 0.5) were selected for further analyses. MedCalc software (version 19.0.4, Mariakerke, Belgium) was employed to analyse the ROC curves.

Radiogenomic correlations and biological function annotations

The correlation matrix between the diagnostic gene modules and the selected radiomic features was built using Spearman rank correlations. Consequently, significant pairwise correlations between the radiomic features and the gene modules were obtained, and only the significant correlations were used in subsequent analyses.

To understand the biological functions of the genes listed in the gene modules, the Gene Ontology (GO) database was used to determine the biological properties of the identified genes (DAVID, <http://david.abcc.ncifcrf.gov/>). GO can be divided into three categories, as follows: biological process, cellular component and molecular function. Pathway analysis was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. For correcting the multiple hypothesis tests, the Benjamin-Hochberg method was used to control the FDR. The protein-protein interaction (PPI) network within the two modules was constructed with Cytoscape software using the online STRING database (version 10.5; <http://string-db.org/>). The top five most connected genes in the PPI network were defined as hub genes.

Immunohistochemical staining for LAMC1 and THBS1

In 40 cases, resected PTC tissue was fixed, embedded in paraffin, section at 5 µm, and subjected to IHC. Sections were deparaffinized and quenched in a 3% hydrogen peroxide aqueous solution to block endogenous peroxidase activity. Then, antigen retrieval was performed using an antigen retrieval unmasking solution (Vector Laboratories, US). The expression levels of laminin subunit gamma 1 (LAMC1) and thrombospondin 1 (THBS1) were detected with anti-LAMC1 (HPA001908, Sigma-Aldrich, US) and THBS1 (ab1823, Abcam, US) antibodies, respectively. The sections were then counterstained

with haematoxylin, dehydrated and mounted. Images of representative fields were visualized and photographed using a microscope (Leica, Germany) at $\times 10$ and $\times 40$ magnification.

Evaluation of the immunostained sections

After immunohistochemical staining, the slides were reviewed and scored by an experienced pathologist. The intensity of LAMC1 and THBS1 expression was graded as follows: negative (-), weakly positive (+), moderately positive (++), and strongly positive (+++). Quantitative analysis of the captured images was performed with ImageJ software using the publicly available IHC toolbox plugin. The ratio of the positively stained area to the total area of each image was recorded as the IHC staining score.

Statistical analysis

Continuous variables are expressed as the mean $\pm$ standard deviation (SD), and Student's t-test was used for comparisons between two groups. Categorical variables were compared by Mann-Whitney U or chi-squared test. The Pearson correlation coefficient r was used to assess the correlation between the value of radiomic features and the IHC staining score. To evaluate the diagnostic performance of radiomics and grey-scale US in predicting CLNM, ROC analysis was performed, and the AUC was compared using Delong's test. All statistical analyses were performed using SPSS software 23.0 for Mac (IBM Corporation, Armonk, NY) or R software (R Foundation for Statistical Computing, Vienna, Austria). A two-tailed p value less than 0.05 was considered statistically significant.

Results

Patient characteristics

As shown in Table 1, 180 patients were included in this study, and they were classified into two groups (metastasis group, $n = 81$; non-metastasis group, $n = 99$) according to their CLN status. Demographic and clinical information was acquired from the medical records of the patients and included age, sex, tumour location, and tumour diameter. The mean age of the patients with and without LNM was significantly different (38.81 ± 11.42 years and 43.90 ± 10.19 years, respectively; $p = 0.001$). A significant difference was also identified between the two groups regarding tumour diameter ($p = 0.002$), but there were no significant differences between the two groups in sex or tumour location.

Table 1
Clinical characteristics of PTC patients with or without LNM

Characteristics	Metastasis (n = 81)	Non-Metastasis (n = 99)	P value
Age			0.001
Mean ± SD	38.81 ± 11.42	43.90 ± 10.19	
Range	20–68	18–65	
Diameter			0.002
Mean ± SD	1.18 ± 0.61	0.90 ± 0.44	
Range	0.3-3.0	0.3–2.5	
Sex			0.062
Male	30 (37.0)	24 (24.2)	
Female	51(63.0)	75 (75.8)	
Location			0.694
Left	28 (37.8)	40 (40.4)	
Right	49 (57.2)	54 (54.5)	
Isthmus	4 (5.0)	5 (5.1)	
PTC indicates papillary thyroid carcinoma; LNM indicates lymph node metastases; SD indicates standard deviation; Significant differences are highlighted in boldface.			

Prediction of CLN status based on radiomic evaluation

The 47 features extracted from the thyroid nodule images were used to construct the radiomic signature via the SVM with LOOCV and were applied to predict the LN status in our cohort. The diagnostic performance of the radiomic signature was assessed by ROC curve analysis (Fig. 2a). The radiomic signature yielded an AUC of 0.873, and the ACC, SEN, and SPEC values were 82.22%, 83.59%, and 80.81%, respectively. In contrast, assessment of the CLN status based solely on grey-scale US yielded an AUC of 0.586 (Fig. 2a), with ACC, SEN, and SPEC values of 60.56%, 39.51%, and 77.78%, respectively. The AUC of the radiomic signature was significantly higher than that of the US-based method ($p < 0.001$).

Genomic module clustering and ROC analysis

Twenty-six PTC tissues with CLNM and twenty-nine PTC tissues without CLNM collected from 55 patients were used for iTRAQ assessment. However, due to protein degradation or uncertainty of the pathological report, 6 samples were excluded from further analysis with subsequent bioinformatic analysis focused on the remaining 49 tumour samples. After the tissues were processed for LC-MS, a

total of 2339 proteins were identified after rejecting proteins with unreliable expression. Sixteen gene modules were generated by WGCNA. Figure 2b and c shows the cluster dendrogram and gene counts for each gene module. ROC analysis was employed to evaluate the discriminatory power of each gene module for predicting CLNM. The AUC values for each gene module were between 0.354 and 0.807 (Table 2). Among the 16 modules, eight had good discriminatory performance (AUC greater than 0.5) and were selected for further study: MEmidnightblue, MELightcyan, MEpink, MEMagenta, MESalmon, MEblue, MEtan, and MEcyan.

Table 2
AUC for the gene modules in assessing cervical lymph node metastasis

Variable	AUC	95%CI
MEmidnightblue	0.807 ± 0.079	0.636–0.922
MELightcyan	0.804 ± 0.078	0.632–0.919
MEpink	0.709 ± 0.093	0.528–0.851
MEMagenta	0.719 ± 0.093	0.539–0.859
MESalmon	0.688 ± 0.010	0.506–0.835
MEblue	0.691 ± 0.095	0.510–0.838
MEtan	0.614 ± 0.106	0.432–0.775
MEcyan	0.589 ± 0.109	0.408–0.755
MEpurple	0.470 ± 0.101	0.272–0.668
MEblack	0.467 ± 0.103	0.264–0.669
MEgreenyellow	0.435 ± 0.100	0.240–0.630
MEturquoise	0.425 ± 0.105	0.219–0.630
MERed	0.421 ± 0.106	0.214–0.629
MEgreen	0.418 ± 0.102	0.217–0.618
MEbrown	0.397 ± 0.099	0.203–0.590
MEyellow	0.354 ± 0.096	0.165–0.543
AUC indicates area under curve; CI indicates confidence interval.		

Associations between radiomic features and gene modules

Among the 47 extracted radiomic features, 28 remained after redundancy removal. The correlation matrix of the remaining radiomic features is shown in Fig. 2d. A radiogenomic map was constructed by correlating the selected radiomic features and the gene modules. Nine pairwise correlations were

statistically significant in the radiogenomic correlation map, as identified by the asterisk (Fig. 2e). Among these, there were 5 gene modules correlated with 6 radiomic features.

According to the radiogenomic correlation map, the margin feature ‘rectlike’ was negatively correlated with the gene modules MEpink ($p = 0.001$) and MEmagenta ($p < 0.001$). The shape feature ‘Eccentricity’ was negatively correlated with the module MEsalmon ($p = 0.006$). The position feature ‘Distance to capsule’ and the echo pattern feature ‘mean tumour contrast’ were positively correlated with MEsalmon. Another echo pattern feature, ‘deviation ratio of tumour and normal thyroid gland’, was positively correlated with the modules MEpink ($p < 0.001$) and MEmagenta ($p < 0.001$). The calcification feature ‘minimum calcification area’ was positively correlated with the modules MElightcyan ($p = 0.005$) and MEblue ($p < 0.001$).

The MEmagenta module had a strong correlation with ‘rectlike’ ($r = -0.61$) and ‘deviation ratio of tumour and normal thyroid gland’ ($r = 0.70$), and MEblue had a strong correlation with ‘minimum calcification area’ ($r = 0.64$). Based on these data, we focused on the modules MEmagenta and MEblue.

GO and KEGG pathway enrichment analyses of the gene modules

GO analysis was applied to determine the biological process, cellular component, and molecular functions of the MEmagenta (Fig. 3a-c) and MEblue (Fig. 4a-c) modules, and KEGG analysis was performed to investigate the involved signalling pathways of the two modules. The most significantly enriched pathways are shown in Figs. 3d and 4d, and the PPI network of the two gene modules is shown in Fig. 5. In the MEmagenta module, the hub genes were *CCT3*, *TCP1*, *CCT4*, *LAMB1* and *LAMC1*. In the MEblue module, the hub genes were *RAD23A*, *PGD*, *ITPA*, *IDH1* and *THBS1*. Table 3 lists the annotations of the three other gene modules as GO terms and KEGG pathway outputs.

Table 3. Annotations of significant gene modules in GO terms and KEGG pathway

Gene module	GO term (biological Process) (adjust p<0.05)	GO term (Cellular Component) (adjust p<0.05)	Go term (Molecular Function) (adjust p<0.05)	KEGG pathway (adjust p<0.05)	Abbreviation
MElightcyan	Regulation of complement activation Complement activation (alternative pathway) Complement activation (classical pathway)	Membrane attack complex Extracellular region Spectrin-associated cytoskeleton	Phospholipid binding Structural constituent of cytoskeleton Receptor binding	Complement and coagulation cascades Proteoglycans in cancer PPAR signaling pathway	Complement activation Protein binding
MEpink	DNA topological change Chromatin remodeling mRNA splicing	Nucleoplasm Membrane Nucleolus	RNA binding Protein binding Chromatin binding	Ribosome biogenesis in eukaryotes Complement and coagulation cascades RNA transportation	Cell division
MEsalmon	Cell adhesion Extracellular matrix organization Endothelial cell migration	Extracellular matrix Extracellular space Extracellular region	Heparin binding Collagen binding Extracellular matrix binding	Proteoglycans in cancer ECM-receptor interaction Focal adhesion	Cell adhesion and migration

Validation of bioinformatic analysis results by IHC

Next, we performed IHC on 40 of the samples tested by iTRAQ and LC/MS to further validate the putative connections between the radiomic features and biological information. *LAMC1* and *THBS1* were hub genes for MEmagenta and MEblue, respectively. These two gene modules were both inclusive of the PI3K/AKT signalling pathway, and *LAMC1* and *THBS1* are both involved in the PI3K/AKT pathway. Therefore, we selected these two representative genes from each module for further validation (Fig. 6).

The results of immunohistochemical staining for *LAMC1* and *THBS1* are shown in Table 4, respectively. There were significant differences in *LAMC1* and *THBS1* staining intensity between PTC tissues with and without CLNM ($U = 93.0$, $p = 0.003$; $U = 90.0$, $p = 0.002$).

Table 4
LAMC1 and THBS1 expression in PTC with or without cervical lymph node metastasis

Gene	Pathology	-	+	++	+++	Total	U	P value
LAMC1	Non-metastasis	7	8	3	2	20	93	0.003
	Metastasis	1	5	7	7	20		
THBS1	Non-metastasis	1	4	8	7	20	90	0.002
	Metastasis	6	9	3	2	20		
PTC indicates papillary thyroid carcinoma.								
The values of the radiomic features ‘Rectlike’ and ‘deviation ratio of tumour and normal thyroid gland’ were significantly correlated with the expression of LAMC1 (r=-0.658, p < 0.001; r = 0.715, p < 0.001); in addition, the value of the feature ‘minimum calcification area’ was highly correlated with the expression of THBS1 (r=-0.756, p < 0.001).								

Discussion

Predicting the CLN status is important for guiding the clinical management of patients with PTC. In this study, we integrated the radiomic features of conventional thyroid US images with the gene expression profile to identify potential radiogenomic biomarkers for PTC. We successfully identified 9 significant pairwise associations between radiomic features and gene modules annotated by functional gene enrichment analysis. These associations demonstrated the feasibility of the noninvasive molecular characterization of PTC by using radiogenomic methods, which can provide complementary information for the noninvasive classification and management of PTC.

High-resolution US is regarded as the first choice for the preoperative assessment of the CLN status. However, the efficiency of US in detecting nonpalpable CLNM is unacceptable. In particular, metastatic nodes in the central cervical compartment are not easily detected by US examination because they are obscured by the thyroid gland and complex structures in the central neck. In our study, diagnosing CLNM based solely on conventional US yielded an AUC of only 0.586, with relatively low ACC and SEN. Radiomics is a new technique that converts medical images into high-dimensional mineable data by quantitatively extracting high-throughput features [13, 19]. Radiomics sometimes applies machine learning methods and has attracted the research interest of many scholars who seek to explore the association between diagnostic information and imaging features [20]. The AUC of our radiomic signature was improved to 0.873, which indicated that compared with the model relying solely on the US features of LNs, the radiomic model had better performance in predicting CLNM, which is consistent with our previous work [21].

The radiomic feature 'Rectlike' measures the smoothness of the tumour margin, and a higher 'Rectlike' value indicates a smoother margin. An ill-defined PTC margin was reported to be significantly associated with CLNM [7, 22]; therefore, patients with a higher 'Rectlike' value tended to have a lower risk of CLNM.

'Rectlike' was negatively correlated with the modules MEpink and MEmagenta. The gene module MEpink was enriched in biological processes, cellular components, and molecular functions and pathways relating to cell division, while the module MEmagenta represented the upregulation of telomere maintenance via telomerase and cell-cell adhesion. Cancer progression is achieved by uncontrolled cell division, invasion, and, eventually, metastasis [23]. Many anti-cancer agents have been designed to regulate cell division to curb cancer progression and metastasis [24]. Cell division causes the telomere to shorten gradually; hence, a key event in the acquisition of cellular immortality is upregulation of the telomere maintenance mechanism [25]. Telomerase is an enzyme responsible for telomere maintenance. An important mechanism of telomerase activation includes mutations in the promoter region (C228T and C250T) of the telomere reverse transcriptase (*TERT*) gene [26]. Clinically, mutations in the TERT promoter are frequently examined from FNA samples [27]. PTC tumours harbouring TERT mutations showed more frequent regional LNM spread than did PTC tumours with a wild-type TERT promoter [28], and TERT mutations are markers for metastatic behaviour [29].

The shape feature 'Eccentricity' measures the ratio of the longitudinal axis to the horizontal axis of the tumour. Tumours with a higher 'Eccentricity' value tend to be taller than wide in shape. Previous studies have demonstrated that a taller-than-wide shape on US is an independent predictor for the absence of CLNM [30, 31]. 'Eccentricity' was negatively correlated with the module MEsalmon, which was related to cell-cell adhesion and extracellular matrix (ECM) in the GO and KEGG enrichment analyses. Cell-cell adhesion is essential for cell-cell cooperation, multicellular polarity, tissue homeostasis, and collective cell movement [32]. Cancer cell migration, the basis for metastatic dissemination, is a plastic and adaptive process integrating cell-cell adhesion, cytoskeletal dynamics, and ECM remodelling. In single-cell migration, the loss of adhesion between cells triggers a dynamic change in the actin cytoskeleton, which endows the cells with motility and alters cell polarity to form spindle-shaped cells. These newly formed mesenchymal-like cells invade the basal ECM and migrate to the underlying tissues [33]. Degradation and remodelling of the ECM, including the basement membrane, by tumour-secreted proteolytic enzymes are also crucial steps in the process of cancer cell intra- and extravasation and colonization at distant sites [34].

The position feature 'Distance to capsule' indicates the distance between the nodule and the nearest thyroid capsule. Evidence has shown that the tumour being in close proximity to or within the capsule was significantly more indicative of CLNM in PTC [22, 35]. Closer proximity to the thyroid capsule may offer more chances for tumours to encounter lymphatic vessels, therefore increasing the likelihood of metastasis within the lymphatic system [31]. 'Distance to capsule' was also correlated with MEsalmon.

Both 'mean tumour contrast' and 'deviation ratio of tumour and normal thyroid gland' are features that reflect the internal echogenicity of PTC. Tumours with higher 'mean tumour contrast' and 'deviation ratio of tumour and normal thyroid gland' values tend to be more hypoechoic or even markedly hypoechoic. Lee et al. [36] demonstrated that hypoechoic and markedly hypoechoic tumours on US were independent risk factors for CLNM, and these malignant US appearances suggested a more invasive biological behaviour, including CLNM [37]. 'Mean tumour contrast' was correlated with the module MEsalmon as

well, and 'deviation ratio of tumour and normal thyroid gland' was correlated with the modules MEpink and MEmagenta.

'Minimum calcification area' is a feature that measures the extent of microcalcification of PTC. The higher the 'minimum calcification area' value, the more microcalcifications there are within the tumour. Microcalcification has been recognized as an independent predictive factor for CLNM [30]. 'Minimum calcification area' was positively correlated with the modules MElightcyan and MEblue. Upon conducting GO and KEGG enrichment analyses, we found that the module MElightcyan was related to complement activation, while MEblue represented cell adhesion and glycolysis. Accumulating evidence has shown that complement activation in the tumour microenvironment promotes tumour growth, suppresses antitumour immunity, and increases metastasis [38]. Enhanced glycolysis has been considered to be the dominant metabolic alteration in malignant tumours [39] and the primary source of ATP for tumour survival upon detachment and during metastasis [40].

Alterations in the expression levels of the genes of interest were confirmed by immunohistochemical staining, indicating that these molecules play an important role in the process of PTC metastasis. Laminins are heterotrimeric ECM proteins that are composed of the alpha, beta, and gamma chains, which mediate cellular signaling via interaction with cell membrane receptors. The *LAMC1* gene, encoding the laminin subunit gamma 1 protein, has been reported to be involved in the progression of various malignant tumours. Previous study showed that overexpression of LAMC1 in endometrial carcinoma was related to aggressive histological types, LN metastasis, advanced International Federation of Gynecology and Obstetrics (FIGO) stage; and LAMC1 knockdown suppressed cell motile and invasive properties in endometrial cancer cells [41]. Besides, treatment with the specific LAMC1 peptide enhanced pulmonary metastasis of B16 melanoma cells and induced the production of matrix metalloproteinase (MMP)-9 from B16 cells [42]. Some studies have suggested that LAMC1 functions to promote metastasis and might be a novel therapeutic target in the treatment of human cancer. Consistent with these previous studies, we found that LAMC1 was expressed at a relatively higher level in samples from patients with metastasis.

THBS1 is a secreted glycoprotein involved in tumour progression via the regulation of ECM remodelling and angiogenesis. The role of THBS1 as an antiangiogenic factor is well documented; however, its effect on tumour progression and metastasis remains controversial [43]. On the one hand, Giuseppe and colleagues found a significant reduction in THBS1 expression associated with patients presenting with thyroid carcinoma metastasis [44]. In vitro, THBS1 could inhibit the migration of clear cell renal carcinoma cells [45]. On the other hand, THBS1 has been reported to promote human follicular thyroid carcinoma cell invasion through the upregulation of urokinase-dependent activity [46] and metastasis to the lungs in a transgenic mouse model of breast cancer [47]. Our present data indicate that *THBS1* may function as a suppressor gene in PTC.

Both the gene modules MEmagenta and MEblue were inclusive of the PI3K/AKT pathway, and both of LAMC1 and THBS1 could activate this signalling pathway. Therefore, we hypothesized that this might be

a key pathway involved in the process of PTC metastasis. The PI3K/AKT pathway is a key regulator of cellular processes involved in cell proliferation/the cell cycle, metabolism, apoptosis, and neovascularization [48]. The PI3K/AKT axis could promote tumour metastasis by enhancing cell motility and angiogenesis; for instance, activated AKT could increase the bioactivities of NF- κ B and VEGF, which in turn promote enhanced cell motility and angiogenesis. The PI3K/AKT pathway could also upregulate the expression of MMPs, which degrade the ECM and promote tumour metastasis and invasion [49].

Radiomics could provide a more accurate and robust method to predict CLN involvement in patients with PTC because the proteomic pattern is expressed in terms of image-based features. A previous study evaluated the association between the gene expression signature and CLNM in PTC [50]. However, this molecular approach is limited by the requirement of invasive surgery or biopsy and its high cost. The imaging features in radiomic studies generally lack biological interpretations, which might hinder their clinical application. Thus, linking imaging characteristics with molecular signatures is a growing research trend that provides additional value to conventional clinical imaging with relevant molecular biological information. In this study, we identified radiomic features that were associated with gene modules, annotated the molecular and physiological effects of the relevant gene modules, and validated the associations and the results of the bioinformatic analysis, all of which lend convincing support to the proposed method. Overall, this approach highlights that imaging phenotypes allow for the noninvasive assessment of the molecular activity of PTC lesions with potential implications for treatment. Moreover, the radiogenomic map could be extended to other solid tumours and might predict the therapeutic response of existing agents through the use of gene signatures.

Several limitations were encountered in our study. First, in research on radiomics, the application of multimodal imaging is common, enlightening us to use elastography and contrast-enhanced US as other modalities to make our radiomic signature more robust and discriminative. Second, a small number of patients were enrolled in this study, and the examination of a larger cohort with available imaging, genetic and survival data is warranted. Our data come from only a single institution, and the results should be validated with a multicentre study. Finally, we focused on molecular analysis at the protein level. In the future, by incorporating other types of ‘-omic’ data, for instance, copy number variations and genetic mutations, we could provide a more complete picture of the molecular characteristics of tumours.

Conclusions

In summary, our results suggest that the radiomic signature proposed in this article has potential for clinical application to noninvasively predict the CLN status in patients with PTC based on preoperative US images. The radiogenomic map allows for a better understanding of the pathophysiological structure of PTC and how molecular processes manifest in a macromolecular way as captured by imaging features. Merging imaging phenotypes with genomic data in the future might further facilitate the application of radiomics in clinical practice for cancer patients and support clinical decision making with personalized management.

Declarations

Ethics approval and consent to participate

This prospective study was approved by the ethics committee of the Fudan University Shanghai Cancer Centre (1809191-1-NSFC). Written informed consent was provided by each patient. This study was performed in accordance with the Declaration of Helsinki.

Consent for publication

All authors have agreed on the contents of the manuscript and provided consent.

Availability of data and materials

The data that support the findings of this study, including the clinical information and proteome data, are available from the ProteomeXchange database. The datasets generated and analysed during the current study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

National Natural Science Foundation of China (81401422), Shanghai Science and Technology Foundation of China (17411963300)

Authors' contributions

Y.T. and P.S. contributed to performing the experiments, interpreting the data, and writing the paper. Y.H. and J.Z. contributed to collecting the clinical data and the ultrasound images. J.Y. contributed to evaluating all the pathology slides. Y.G., J.Y. and Y.W. contributed to feature extraction and model construction. S.Z. and C.C. conceived the idea and provided funding support.

Acknowledgements

We thank Dr. Leyin Li for her generous help with radiomic feature extraction. We also appreciate Dr. Jiangnan Wu for his specialized consultation on statistical analysis. This study was supported by the National Natural Science Foundation of China (81401422) and the Shanghai Science and Technology Foundation of China (17411963300).

References

1. Noone AM, Cronin KA, Altekruze SF, Howlader N, Lewis DR, Petkov VI, et al. Cancer incidence and survival trends by subtype using data from the surveillance epidemiology and end results program, 1992–2013. *Cancer Epidemiol Biomarkers Prev.* 2017;26:632–41.
2. Roh JL, Park JY, Kim JM, Song CJ. Use of preoperative ultrasonography as guidance for neck dissection in patients with papillary thyroid carcinoma. *J Surg Oncol.* 2009;99:28–31.
3. Baek SK, Jung KY, Kang SM, Kwon SY, Woo JS, Cho SH, et al. Clinical risk factors associated with cervical lymph node recurrence in papillary thyroid carcinoma. *Thyroid.* 2010;20:147–52.
4. Lundgren CI, Hall P, Dickman PW, Zedenius J. Clinically significant prognostic factors for differentiated thyroid carcinoma: a population-based, nested case-control study. *Cancer.* 2006;106:524–31.
5. Hong YR, Yan CX, Mo GQ, Luo ZY, Zhang Y, Wang Y, et al. Conventional US, elastography, and contrast enhanced US features of papillary thyroid microcarcinoma predict central compartment lymph node metastases. *Sci Rep.* 2015;5:7748.
6. Zhang Y, Luo YK, Zhang MB, Li J, Li CT, Tang J, et al. Values of ultrasound features and MMP-9 of papillary thyroid carcinoma in predicting cervical lymph node metastases. *Sci Rep.* 2017;7:6670.
7. Ardakani AA, Reiazi R, Mohammadi A. A clinical decision support system using ultrasound textures and radiologic features to distinguish metastasis from tumor-free cervical lymph nodes in patients with papillary thyroid carcinoma. *J Ultrasound Med.* 2018;37:2527–35.
8. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid.* 2016;26:1–133.
9. Stulak JM, Grant CS, Farley DR, Thompson GB, van Heerden JA, Hay ID, et al. Value of preoperative ultrasonography in the surgical management of initial and reoperative papillary thyroid cancer. *Arch Surg.* 2006;141:489 – 94; discussion 94 – 86..
10. Zhu Y, Li H, Guo W, Drukker K, Lan L, Giger ML, et al. Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma. *Sci Rep.* 2015;5:17787.
11. Zhou M, Leung A, Echegaray S, Gentles A, Shrager JB, Jensen KC, et al. Non-small cell lung cancer radiogenomics map identifies relationships between molecular and imaging phenotypes with prognostic implications. *Radiology.* 2018;286:307–15.
12. Karlo CA, Di Paolo PL, Chaim J, Hakimi AA, Ostrovnaya I, Russo P, et al. Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations. *Radiology.* 2014;270:464–71.
13. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:4006.

14. Xia W, Chen Y, Zhang R, Yan Z, Zhou X, Zhang B, et al. Radiogenomics of hepatocellular carcinoma: multiregion analysis-based identification of prognostic imaging biomarkers by integrating gene data—a preliminary study. *Phys Med Biol*. 2018;63:035044.
15. Wu J, Li B, Sun X, Cao G, Rubin DL, Napel S, et al. Heterogeneous enhancement patterns of tumor-adjacent parenchyma at MR imaging are associated with dysregulated signaling pathways and poor survival in breast cancer. *Radiology*. 2017;285:401–13.
16. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. *J Am Coll Radiol*. 2017;14:587–95.
17. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods*. 2009;6:359–62.
18. Plubell DL, Wilmarth PA, Zhao Y, Fenton AM, Minnier J, Reddy AP, et al. Extended multiplexing of tandem mass tags (TMT) labeling reveals age and high fat diet specific proteome changes in mouse epididymal adipose tissue. *Mol Cell Proteomics*. 2017;16:873–90.
19. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278:563–77.
20. Wu S, Zheng J, Li Y, Yu H, Shi S, Xie W, et al. A radiomics nomogram for the preoperative prediction of lymph node metastasis in bladder cancer. *Clin Cancer Res*. 2017;23:6904–911.
21. Liu T, Zhou S, Yu J, Guo Y, Wang Y, Zhou J, et al. Prediction of lymph node metastasis in patients with papillary thyroid carcinoma: a radiomics method based on preoperative ultrasound images. *Technol Cancer Res Treat*. 2019;18:1533033819831713.
22. Guo L, Ma YQ, Yao Y, Wu M, Deng ZH, Zhu FW, et al. Role of ultrasonographic features and quantified BRAFV600E mutation in lymph node metastasis in Chinese patients with papillary thyroid carcinoma. *Sci Rep*. 2019;9:75.
23. Riaz SK, Iqbal Y, Malik MF. Diagnostic and therapeutic implications of the vascular endothelial growth factor family in cancer. *Asian Pac J Cancer Prev*. 2015;16:1677–82.
24. Avnet S, Cortini M. Role of pericellular matrix in the regulation of cancer stemness. *Stem Cell Rev Rep*. 2016;12:464–75.
25. Reddel RR. Telomere maintenance mechanisms in cancer: clinical implications. *Curr Pharm Des*. 2014;20:6361–74.
26. Nasirden A, Saito T, Fukumura Y, Hara K, Akaike K, Kurisaki-Arakawa A, et al. In Japanese patients with papillary thyroid carcinoma, TERT promoter mutation is associated with poor prognosis, in contrast to BRAF (V600E) mutation. *Virchows Arch*. 2016;469:687–96.
27. Ferrari SM, Fallahi P, Ruffilli I, Elia G, Ragusa F, Paparo SR, et al. Molecular testing in the diagnosis of differentiated thyroid carcinomas. *Gland Surg*. 2018;7:19–29.
28. Bu R, Siraj AK, Divya SP, Kong Y, Parvathareddy SK, Al-Rasheed M, et al. Telomerase reverse transcriptase mutations are independent predictor of disease-free survival in Middle Eastern papillary thyroid cancer. *Int J Cancer*. 2018;142:2028–39.

29. Donati B, Ciarrocchi A. Telomerase and telomeres biology in thyroid cancer. *Int J Mol Sci.* 2019;20:E2887.
30. Xu JM, Xu XH, Xu HX, Zhang YF, Guo LH, Liu LN, et al. Prediction of cervical lymph node metastasis in patients with papillary thyroid cancer using combined conventional ultrasound, strain elastography, and acoustic radiation force impulse (ARFI) elastography. *Eur Radiol.* 2016;26:2611–22.
31. Wang QC, Cheng W, Wen X, Li JB, Jing H, Nie CL. Shorter distance between the nodule and capsule has greater risk of cervical lymph node metastasis in papillary thyroid carcinoma. *Asian Pac J Cancer Prev.* 2014;15:855–60.
32. Odenthal J, Takes R, Friedl P. Plasticity of tumor cell invasion: governance by growth factors and cytokines. *Carcinogenesis.* 2016;37:1117–28.
33. Gonzalez DM, Medici D. Signaling mechanisms of the epithelial-mesenchymal transition. *Sci Signal.* 2014;7:re8.
34. Bogenrieder T, Herlyn M. Axis of evil: molecular mechanisms of cancer metastasis. *Oncogene.* 2003;22:6524–36.
35. Wu Q, Zhang YM, Sun S, Li JJ, Wu J, Li X, et al. Clinical and sonographic assessment of cervical lymph node metastasis in papillary thyroid carcinoma. *J Huazhong Univ Sci Technol Med Sci.* 2016;36:823–27.
36. Lee YS, Lim YS, Lee JC, Wang SG, Son SM, Kim SS, et al. Ultrasonographic findings relating to lymph node metastasis in single micropapillary thyroid cancer. *World J Surg Oncol.* 2014;12:273.
37. Nam SY, Shin JH, Han BK, Ko EY, Ko ES, Hahn SY, et al. Preoperative ultrasonographic features of papillary thyroid carcinoma predict biological behavior. *J Clin Endocrinol Metab.* 2013;98:1476–82.
38. Afshar-Kharghan V. The role of the complement system in cancer. *J Clin Invest.* 2017;127:780–9.
39. Yang J, Ren B, Yang G, Wang H, Chen G, You L, et al. The enhancement of glycolysis regulates pancreatic cancer metastasis. *Cell Mol Life Sci.* 2019;77:305–21.
40. Payen VL, Porporato PE, Baselet B, Sonveaux P. Metabolic changes associated with tumor metastasis, part 1: tumor pH, glycolysis and the pentose phosphate pathway. *Cell Mol Life Sci.* 2015;73:1333–48.
41. Long R, Liu Z, Li J, Yu H. COL6A6 interacted with P4HA3 to suppress the growth and metastasis of pituitary adenoma via blocking PI3K-Akt pathway. *Aging.* 2019;11:8845–59.
42. Yeh MH, Tzeng YJ, Fu TY, You JJ, Chang HT, Ger LP, et al. Extracellular matrix-receptor interaction signaling genes associated with inferior breast cancer survival. *Anticancer Res.* 2018;38:4593–605.
43. Sepiashvili L, Hui A, Ignatchenko V, Shi W, Su S, Xu W, et al. Potentially novel candidate biomarkers for head and neck squamous cell carcinoma identified using an integrated cell line-based discovery strategy. *Mol Cell Proteomics.* 2012;11:1404–15.
44. Bunone G, Vigneri P, Mariani L, Butó S, Collini P, Pilotti S, et al. Expression of angiogenesis stimulators and inhibitors in human thyroid tumors and correlation with clinical pathological

features. *Am J Pathol.* 1999;155:1967–76.

45. Bienes-Martinez R, Ordonez A, Feijoo-Cuaresma M, Corral-Escariz M, Mateo G, Stenina O, et al. Autocrine stimulation of clear-cell renal carcinoma cell migration in hypoxia via HIF-independent suppression of thrombospondin-1. *Sci Rep.* 2012;2:788.
46. Sid B, Langlois B, Sartelet H, Bellon G, Dedieu S, Martiny L. Thrombospondin-1 enhances human thyroid carcinoma cell invasion through urokinase activity. *Int J Biochem Cell Biol.* 2008;40:1890–900.
47. Yee KO, Connolly CM, Duquette M, Kazerounian S, Washington R, Lawler J. The effect of thrombospondin-1 on breast cancer metastasis. *Breast Cancer Res Treat.* 2009;114:85–96.
48. Hamzehzadeh L, Atkin SL, Majeed M, Butler AE, Sahebkar A. The versatile role of curcumin in cancer prevention and treatment: a focus on PI3K/AKT pathway. *J Cell Physiol.* 2018;233:6530–37.
49. Jin S, Borkhuu O, Bao W, Yang YT. Signaling pathways in thyroid cancer and their therapeutic implications. *J Clin Med Res.* 2016;8:284–96.
50. Zhai T, Muhanhali D, Jia X, Wu Z, Cai Z, Ling Y. Identification of gene co-expression modules and hub genes associated with lymph node metastasis of papillary thyroid cancer. *Endocrine.* 2019;66:573–84.

Figures

Figure 1

The figure shows the workflow and the data processing of the radiogenomic pipeline. a) US images were acquired before surgery, and these images were manually segmented; radiomic features were extracted from segmented thyroid US images; a support vector machine was used to build the final radiomic signature. b) A slice of tumour tissue was harvested during surgery and analysed by iTRAQ and LC-MS. The whole bulk expression data were reduced to gene modules using WGCNA. c) These modules were correlated with the radiomic features; then, the modules were further investigated by GO and KEGG pathway analyses. In the final step, IHC was performed to validate the results of the bioinformatic analysis. US indicates ultrasound; iTRAQ indicates isobaric tags for relative and absolute quantitation; LC-MS indicates liquid chromatograph-mass spectrometry; WGCNA indicates weighted gene co-expression network analysis; GO indicates Gene Ontology; KEGG indicates Kyoto Encyclopedia of Genes and Genomes; IHC indicates immunohistochemistry.

Figure 2

WGCNA analysis and the radiogenomics correlation map. a) The ROC curves of the radiomic signature and the US-based method. The purple line is the curve of the radiomic signature, with an AUC of 0.873;

the green line is the curve of the US-based method, with an AUC of 0.586; and the yellow line is the reference line, with AUC=0.500. Cluster dendrogram of 2356 proteins for 49 PTC patients created by applying the WGCNA method. b) Gene dendrogram after clustering; each colour represents one gene module. c) Counts for each gene module. d) The correlation matrix of 28 radiomic features. e) Radiogenomic correlation map between selected radiomic features and gene modules created using the Spearman rank correlation method; significant correlations ($p < 0.05$) are identified with asterisks. The corresponding coefficients are displayed in a heatmap, where red and blue indicate positive and negative correlations, respectively. * indicates a statistically significant association. AUC indicates area under the receiver operating characteristic curve; US indicates ultrasound; PTC indicates papillary thyroid carcinoma; ROC indicates receiver operating characteristic; WGCNA indicates weighted gene co-expression network analysis.

Figure 3

Top significantly enriched GO terms of the MEmagenta module, including a) biological process; b) cellular component; and c) molecular function. Significantly enriched pathways identified by KEGG pathway analysis. d) Pathways enriched in the MEmagenta module. GO indicates Gene Ontology; KEGG indicates Kyoto Encyclopedia of Genes and Genomes.

Figure 4

Top significantly enriched GO terms of the MEblue module, including a) biological process; b) cellular component; and c) molecular function. Significantly enriched pathways identified by KEGG pathway analysis. d) Pathways enriched in the MEblue module. GO indicates Gene Ontology; KEGG indicates Kyoto Encyclopedia of Genes and Genomes.

Figure 5

Protein-protein interaction (PPI) network of the genes listed in the modules MEmegenta (a) and MEblue (b). Protein selected for further immunohistochemistry validation is shown in purple.

Figure 6

Four representative cases are shown: a) shows a lesion with a smooth margin, relatively hyperechoic signal, and low LAMC1 expression level, which was confirmed by immunohistochemical staining; b) shows a lesion with a lobulated/irregular margin, hypoechoic signal and relatively high LAMC1 expression level; c) shows a lesion with little microcalcification and a relatively high THBS1 expression level, which was confirmed by immunohistochemical staining; d) shows a lesion with much more microcalcification and relatively low THBS1 expression. In each case, images of the tumour are displayed at 10X and 40X magnification.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [additionalfile1.docx](#)