Translational Oncology 14 (2021) 101222

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Original Research

Improving differential diagnosis of pulmonary large cell neuroendocrine carcinoma and small cell lung cancer via a transcriptomic, biological pathway-based machine learning model

Junhong Guo [a,1], Likun Hou [a,1], Wei Zhang [a], Zhengwei Dong [a], Lei Zhang [b,*], Chunyan Wu [a,*]

[a] Department of Pathology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, PR China
[b] Department of Thoracic Surgery, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai 200433, PR China

ARTICLE INFO

Keywords:
Pulmonary large cell neuroendocrine carcinoma (LCNEC)
Small cell lung cancer (SCLC)
Differential diagnosis, Molecular profile
Machine learning

ABSTRACT

Background: Accurately differentiating between pulmonary large cell neuroendocrine carcinomas (LCNEC) and small cell lung cancer (SCLC) is crucial to make appropriate therapeutic decisions. Here, a classifier was constructed based on transcriptome data to improve the diagnostic accuracy for LCNEC and SCLC.
Methods: 13,959 genes mapped to 186 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were included. Gene Set Variation Analysis (GSVA) algorithm was used to enrich and score each KEGG pathway from RNA-sequencing data of each sample. A prediction model based on GSVA score was constructed and trained via ridge regression based on RNA-sequencing datasets from 3 published studies. It was validated by another independent RNA-sequencing dataset. Clinical feasibility was tested by comparing model predicated result using RNA-sequencing data derived from hard-to-diagnose samples of lung neuroendocrine cancer to conventional histology-based diagnosis.
Results: This model achieved a ROC-AUC of 0.949 and a concordance rate of 0.75 for the entire prediction efficiency. Of the 27 borderline samples, 17/27 (63.0%) were predicted as LCNEC, 7/27 were predicted as SCLC, and the remainder was NSCLC. Only 8 cases (29.6%) with LCNEC were diagnosed by pathologists, which was significantly lower than the results predicted by the model. Furthermore, cases with predicted LCNEC by the model had a significant longer disease-free survival than those where the model predicted SCLC ($P = 0.0043$).
Conclusion: This model was able to give an accurate prediction of LCNEC and SCLC. It may assist clinicians to make the optimal decision for patients with pulmonary neuroendocrine tumors in choosing appropriate treatment.

Introduction

Pulmonary large cell neuroendocrine carcinoma (LCNEC) is a rare type of tumor, with poorly differentiated or undifferentiated neuroendocrine morphology, accounting for approximately 3% of all lung cancers [1]. In 2004, LCNEC was classified as a subtype of large cell lung cancer by the World Health Organization (WHO). In 2015, it was classified as a neuroendocrine tumor of the lung along with small cell lung cancer (SCLC), atypical carcinoids and typical carcinoids [2]. Local and systemic metastases are common in LCNEC, and the cure rate and overall prognosis of LCNEC are poor, with 5-year survival rates of 13–57% for all patients, 27–62% for early stage, and < 5% for advanced stage patients [3]. For advanced LCNEC, there is an increasing evidence suggesting LCNEC should be treated with a platinum/etoposide–based regimen in clinical practice due to similar neuroendocrine lineage origin [4,5].

Radical surgical resection is recommended to localized LCNEC [3]. Further, Raman V and colleagues, in a study of 6092 LCNEC patients, suggested a survival benefit of surgery in patients with both early and locally advanced (stage IIIA) LCNEC [6]. However, unlike LCNEC, surgical treatment is recommended only for stage I/II N0 SCLC [7]. Therefore, more in-depth exploration is needed to find effective differential/correct diagnosis methods to screen LCNEC patients so as to obtain effective treatment strategies.

* Corresponding authors.
1 These authors contributed equally to this Work.
E-mail addresses: zhanglei_fkyy@163.com (L. Zhang), wuchunyan581@sina.com (C. Wu).

https://doi.org/10.1016/j.tranon.2021.101222
Received 31 August 2021; Accepted 7 September 2021

Conventional diagnosis of LCNEC is mainly based on histopathology and immunohistochemistry. Usually, an accurate diagnosis can be made on a surgically resected tumor specimen; however, it is difficult to perform morphological and immunohistochemical examination for small or crushed biopsy specimens or cytological samples, which often lack a well-preserved morphology. Thus it is difficult to obtain the accurate diagnosis for patients who can only provide these types of specimens [8]. Furthermore, some subsets of SCLC and LCNEC (borderline samples) have similar characteristics, make it difficult to establish accurate pathologic diagnosis [9,10]. The other issue with LCNEC pathological diagnosis is that it has only a moderate reproducibility. Travis et al. reported that 3 out of 5 pathologists agreed on SCLC in 90% of the cases, but only in 50% of the LCNEC cases [11]. Two other previous studies have data in the same range: consensus (5/9 pathologists) for SCLC / LCNEC /NSCLC in 69% of the cases [12] and consensus for SCLC in 83%, for LCNEC in 78% of the cases [13]. Therefore, developing an objective method based on tumor molecular markers to accurately diagnose LCNEC will have important clinical value.

Recently, the molecular characteristics of LCNEC have been reported [14] through next-generation sequencing (NGS). George et al. performed whole-genome sequencing on 60 LCNEC tissues and transcriptome sequencing on 69 LCNEC tissues [15]. Compared with normal tissue, significant mutations were found in *TP53, RB1, STK11, KEAP1* and *RAS* (*KRAS/NRAS/HRAS*) pathway. Other studies have found that the mutation rate of the *NOTCH* family genes in LCNEC was relatively high, which causes a change of the NOTCH pathway and ultimately affects the differentiation of neuroendocrine tissue [16,17]. In a small study ($n = 45$), it was found that 18 LCNEC samples had genetic characteristics of SCLC (with co-mutation/loss of *TP53* and *RB1, MYCL* amplification, and other SCLC-type alterations), 25 samples had genetic characteristics of NSCLC (lack of co-altered *TP53* and *RB1*, but with *STK11, KRAS*, and *KEAP1* mutations), and 2 samples had carcinoid-like genetic characteristics (with *MEN1* mutations and low mutation burden) [16]. However, the molecular typing studies are all carried out in patients who had been clearly diagnosed as LCNEC. No study has been conducted to distinguish LCNEC patients from SCLC and other lung neuroendocrine carcinomas based on the molecular characteristics.

Therefore, this study is first to construct and validate a prediction model, that can distinguish LCNEC from SCLC, through comprehensive transcriptomic profiling and machine learning approach.

Methods

Data collection

In order to build a classifier that predicts the patients' tumor subtype, highly selected groups of patients were included in the study. To achieve sample balance, 60 NSCLC cases, 66 LCNEC cases and 52 SCLC cases were selected as training set, the RNA-Seq data of these cases were retrieved from TCGA (Firehose Legacy, https://tcga-data.nci.nih.gov/tcga/), and public data from Julie George et al., Nature Communications 2018 [15], and Julie George et al., Nature 2015 [18], respectively. Another 82 NSCLC and 30 LCNEC cases tested by Amoy Diagnostics Co. Ltd. (Xiamen, China) and 15 SCLC cases published by Martin Peifer et al., Nature Genetics 2012 [19] were included in the validation set. Only coding genes, a total of 13,959, were selected for further study. The expression level of each gene was log (TPM+1) transformed for the downstream analyses. The characteristics of the patient in these cohorts are provided in Table 1.

Additionally, RNA sequencing data of 27 samples that could not be clearly diagnosed as NSCLC or LCNEC or SCLC by histological characteristics and immunohistochemistry (named borderline samples) was performed at Amoy Diagnostics Co Ltd. (Xiamen, China), and the sequencing data were used to further validate the accuracy of the prediction model in differentiating LCNEC or SCLC. The characteristics of these patients are provided in Table 2 and Supplementary Table 1. The study protocol was approved by the Ethics Committee of the Shanghai Pulmonary Hospital. Patient consent was not required because of the retrospective nature of the study. Patient data were anonymized.

Candidate genes model

By consulting the literature [15,18,20–23], a batch of prior candidate genes that differ in these three lung cancer subtypes at the mutation, CNV and transcriptome levels were collected. In order to evaluate the difference among the data, the parallel coordinates for candidate genes were examined in the training set and the validation set. Principal component analysis (PCA) was used to infer the distribution of the training set and the validation set, expecting that the training set and the validation set had the same distribution, and the generalization ability of the model could be better reflected in the validation set.

Considering that the differentiation of lung cancer subtypes is a classification issue, the Logistic Regression model is mainly used for classification and prediction of the data. In order to prevent over-fitting of the model, Lasso (L1) regularization was applied to generate a sparse matrix for feature selection due to the large number of candidate genes. The data of the training set were divided into 10 parts in proportion, K-fold cross-validation was used to estimate the most appropriate regularization coefficient. The threshold value of the regularization coefficient was a logarithmic geometric sequence of (−2, 2). At the same time, the coordinate descent method ('liblinear') was used to optimize the parameter iteration, and the highest number of iterations was set as

Table 1
Demographics of study subjects in training and validation cohort.

Characteristics	Training Cohort				Validation Cohort			
	LCNEC	SCLC	NSCLC	*P* value	LCNEC	SCLC	NSCLC	*P* value
Number	66	52	60		30	15	82	
Median age (years, range)	63 (45–90)	64 (51–83)	66.5 (38–82)	0.263	65 (42–78)	65 (47–78)	62 (31.7–80)	0.147
Gender				**0.017**				**0.003**
Male	52 (78.8%)	39 (75.0%)	34 (56.7%)		29 (96.7%)	8 (53.3%)	60 (73.2%)	
Female	14 (21.2%)	13 (25.0%)	26 (43.3%)		1 (3.3%)	7 (46.7%)	22 (26.8%)	
Smoking history				**0.001**				**<0.001**
Never	1 (1.5%)	1 (1.9%)	9 (15.0%)		21 (70.0%)	–	31 (37.8%)	
Current/ Former	61 (92.4%)	50 (96.2%)	51 (85.0%)		9 (30.0%)	–	51 (62.2%)	
NA	4 (6.1%)	1 (1.9%)	0		0	15 (100%)	0	
Stage				0.492				**<0.001**
I	26 (39.4%)	22 (42.3%)	30 (50.0%)		15 (50.0%)	–	45 (54.9%)	
II	16 (24.2%)	8 (15.4%)	12 (20.0%)		6 (20.0%)	–	17 (20.7%)	
III	17 (25.8%)	15 (28.8%)	12 (20.0%)		8 (26.7%)	–	15 (18.3%)	
IV	4 (6.1%)	7 (13.5%)	5 (8.3%)		1 (3.3%)	–	5 (6.1%)	
NA	3 (4.5%)	0	1 (1.7%)		0	15 (100%)	0	

LCNEC: large-cell neuroendocrine carcinomas; NSCLC: non-small cell lung cancer; SCLC: small cell lung cancer; NA: not available.

Table 2
Demographics of 27 subjects with borderline samples.

	Total	Model prediction			Pathologist judgement			
		LCNEC	SCLC	NSCLC	LCNEC	SCLC	NSCLC	Not determined
Number	27	17	7	3	8	12	6	1
Median age (years, range)	64 (44–80)	69 (53–80)	62 (44–72)	59 (53–64)	59 (53–79)	70 (64–74)	53 (59–79)	71
Gender								
Male	23 (85.2%)	15 (88.2%)	5 (71.4%)	3 (100%)	7 (87.5%)	10 (83.3%)	5 (83.3%)	1 (100%)
Female	4 (14.8%)	2 (11.8%)	2 (28.6%)	0	1 (12.5%)	2 (16.7%)	1 (16.7%)	0
Smoking history								
Never	2 (7.4%)	1 (5.9%)	1 (14.3%)	0	0	2 (16.7%)	0	0
Current/ Former	25 (92.6%)	16 (94.1%)	6 (85.7%)	3 (100%)	8 (100%)	10 (83.3%)	6 (100%)	1 (100%)
Stage								
I	11 (40.8%)	7 (41.2%)	1 (14.3%)	1 (33.3%)	3 (27.3%)	5 (50%)	2 (40%)	1 (100%)
II	8 (29.6%)	6 (35.3%)	2 (28.6%)	1 (33.3%)	5 (45.4%)	0	3 (60%)	0
III	8 (29.6%)	4 (23.5%)	4 (57.1%)	1 (33.3%)	3 (27.3%)	5 (50%)	0	0

LCNEC: large-cell neuroendocrine carcinomas; NSCLC: non-small cell lung cancer; SCLC: small cell lung cancer.

5000. Then, the optimized parameters were brought into the model to predict the validation set. The performance of the model was evaluated by calculating the receiver operating characteristic (ROC) curve, the value of the area under the curve (AUC), the multi-class confusion matrix and the Kappa coefficient.

GSVA score model of KEGG pathway

The expression value of a single gene fluctuates greatly among different samples. Thus, as a feature tag, the generalization ability of a single gene is not very good. Therefore, to enhance the robustness of the model, we considered looking for a comprehensive index to replace the single gene as an expression signature. Considering that the Kyoto Encyclopedia of Genes and Genomes (KEGG) forms a biologically significant gene set based on metabolic pathways and other cellular processes, Gene Set Variation Analysis (GSVA) algorithm was used to enrich and score each KEGG pathway, and then using the score to perform feature engineering replacement. The GSVA R package (version 1.34.0) [24] was used to calculate normalized enrichment scores (ES). GSVA assumes that the data obey a normal distribution. By fitting the normalized expression value with a Gaussian distribution function, the score was transformed into a probability distribution instead of the original expression value.

In this study, a total of 13,959 genes, mapped to 186 KEGG pathways, were included in the analysis, and the distribution of these pathways in the training and validation set was also examined through the PCA method. Compared with the Candidate Genes Model, the distribution of these pathways in the two data sets was more similar in the GSVA Score Model (Supplementary Fig. 1). Considering that all KEGG pathways have their own biological significance, the pathway heat map was drawn based on the hierarchical clustering method to find the differences in the up-regulated or down-regulated pathways between

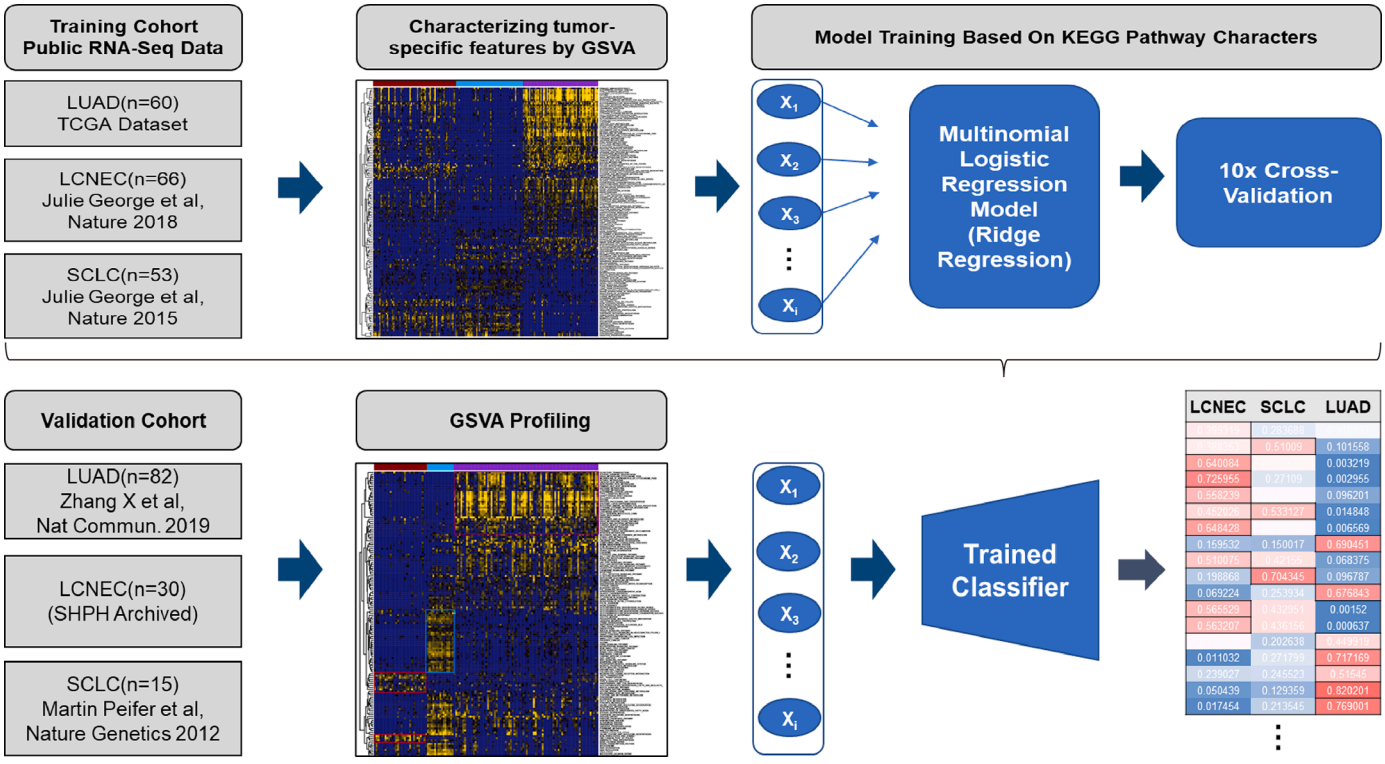

Fig. 1. Flow-chart showing steps for designing the GSVA score prediction model. The development of the classifier was divided into 2 phases. Phase 1 involved selecting the training cohort, charactering tumor-specific features by the GSVA, and the model training based on KEGG pathway characters. Phase 2 involved selecting the validation cohort, the GSVA profiling, and the probability distribution of each patient in LCNEC, SCLC and NSCLC as calculated by the classifier.

LCENC and SCLC in the training set, so as to further clarify the differences in molecular characteristics between the two lung cancer subtypes, and verified in the validation set (Fig. 1).

The logistic regression model was also used to process data, but this time the L2 regularization was adopted. Since each KEGG pathway has its own biological significance and classification effect, the characteristics of all pathways were retained as much as possible, and made the coefficients of pathways with small contributions as small as possible. The subsequent calculation and evaluation process of the GSVA Score Model was the same as of the Candidate Genes Model. The Sklearn Python package was mainly used in the process of model establishment, the Logistic Regression L2 regularization loss function was as follows, $\min_{w,c}\frac{1}{2}w^T w + C\sum_{i=1}^{n}\log(\exp(-y_i(X_i^T w + c)) + 1)$, where "$w$" is the regression coefficient of the variable, "X_i" is the score, "y_i" is the classification tag, and "C" is the coefficient that controls the balance between regularization term and empirical risk function.

Predictions of borderline samples

After the optimized classification model was obtained, the GSVA score of 27 borderline samples was substituted into the prediction classification. The Kaplan-Meier survival curves for different lung cancer subtypes were determined by pathologists or the model was compared to evaluate the classification effectiveness of the GSVA Score Model.

Immunohistochemistry

Formalin-fixed, paraffin-embedded tumor samples were analyzed by immunohistochemistry (IHC) using antibodies to CD56 (Dako, clone 123C3, 1:100), Syn (Dako, clone DAK-SYNAP, 1:50), INSM1 (Santa Cruz Biotechnology, clone A-8, 1:400), and CgA (Dako, clone DAK-A3, 1:100). The experimental procedure was performed by following the manufacturer's instructions, and IHC stains were evaluated by two pathologists. Expression of each neuroendocrine marker was semi-quantified using H-scores (range 0–300), which incorporate the staining intensity (range 0–3+) and the percentage of positively-stained tumor cells (range 0–100%).

Statistical analysis

Statistical calculations were performed with R (version 4.0.4; https://www.r-project.org/). Disease Free Survival (DFS) curves were described by Kaplan-Meier method and compared by Log-rank test. Receiver operating curve (ROC) analyses were performed to assess the classification prediction effectiveness of the model. P values < 0.05 were considered significant.

Results

GSVA score model building and validation

Both LCNEC and SCLC belong to pulmonary neuroendocrine tumors. Due to their similarities in pathological phenotypes and molecular characteristics it may be difficult to differentiate between both tumor types and make an appropriate treatment decision for the patients. Therefore, this study was planned to find the different genes between LCNEC and SCLC and to construct a classifier to assist clinical diagnosis. First, a Candidate Genes Model was designed based on the different genes collected from published literature. Although the AUC of the Candidate Genes Model was high, the results of the confusion matrix analysis showed that its classification effect was not ideal (the accuracy of judging SCLC in the validation set was only 2/15, internal data). In addition, PCA showed that there was a batch effect in the Candidate Genes Model, so it was not the optimal choice to construct a model with gene tags.

Considering that GSVA is an intermediate universal tool, it provides summaries of pathway activities for a more open-ended biological analysis. In this study, the GSVA method was further used to construct the classifier. After screening, a total of 13,959 genes, mapped to 186 KEGG pathways, were included in the analysis, the pathway was scored by GSVA, then the score was used as the feature tag for classification selection. Both in training and validation set, the GSVA confirmed the KEGG terms, including non-homologous end-joining, one carbon pool by folate, notch signaling pathway, taste transduction, and so on were most enriched in LCNEC. Long term potentiation, oocyte meiosis, neurotrophin signaling pathway, mTOR signaling pathway, RNA polymerase, and so on were most enriched in SCLC. Allograft rejection, asthma, apoptosis, autoimmune thyroid disease, antigen processing, presentation, and so on were most enriched in NSCLC (Supplementary Fig. 3). From the heatmap, there are obvious differences in the enrichment of up-regulated and down-regulated expression pathways among LCNEC, SCLC, and NSCLC patients in the training set (Supplementary Fig. 2) and the validation set (Fig. 2a, b). The pattern of up-regulated and down-regulated expression pathways between the two data sets were similar. The prediction power of the model was examined using the confusion matrix (Fig. 2c) and the ROC curve (Fig. 2d). The confusion matrix algorithm has a Kappa coefficient of 0.75, indicating a high degree of consistency between the model prediction and clinical diagnosis, particularly in SCLC prediction, which was completely consistent with the clinical diagnosis. In addition, the ROC curve was used to evaluate the specificity and sensitivity of the model prediction results. The area under the ROC curve calculated on the independent validation cohort data was 0.94, indicating that the GSVA Score Model was reliable for classifying LCNEC and SCLC.

GSVA score model performance

To validate the constructed GSVA Score Model, 27 borderline samples were classified using the model. The Sankey diagram showed the difference of pathological subtypes of the 27 borderline samples determined by clinical diagnosis and model prediction (Fig. 3a). 7 samples who were pathologically judged to be SCLC and 5 samples who were pathologically judged to be NSCLC with neuroendocrine characteristics were predicted as LCNEC by the model, and 5 samples were judged to be LCNEC by both pathologists and the model. 1 sample without a definite diagnosis by a pathologist was predicted as SCLC by the model, 1 NSCLC sample with neuroendocrine characteristics judged by a pathologist was predicted as SCLC by the model, and another 5 samples were judged to be SCLC by both a pathologist and the model. 2 samples with pathologically judged LCNEC and 1 sample with pathologically judged SCLC were predicted as NSCLC by the model.

Detailed clinical characteristics, the clinical diagnosis, the model prediction results, and the IHC scores for each neuroendocrine as well as shared genetic markers between LCNECs and SCLC of the 27 borderline samples are shown in Fig. 3b. The difference of expression of neuroendocrine markers in borderline samples was not obvious. Both LCNEC and SCLC as predicted by the model detect neuroendocrine markers, among which, the number of LCNEC and SCLC samples positive for neural cell adhesion molecule (NCAM-1/CD56) was 12/17 and 7/7 respectively, the number of LCNEC and SCLC samples positive for synaptophysin (Syn) was 11/17 and 4/7 respectively, the number of LCNEC and SCLC samples positive for chromogranin A (CgA) was 9/17 and 5/7 respectively, the number of LCNEC and SCLC samples positive for insulinoma-associated protein 1 (INSM1) was 13/17 and 5/7 respectively, and the number of LCNEC and SCLC samples positive for thyroid transcription factor-1 (TTF-1) was 7/17 and 4/7, respectively. However, these markers were negative in all three of the predicted NSCLC patients, except for one patient who had a positive CD56 test. For genetic markers, the number of samples with positive detection of P53 in LCNEC, SCLC and NSCLC samples predicted by the model was 12/17, 5/7, and 1/3 respectively; most of the LCNEC (13/17) and SCLC (5/7)

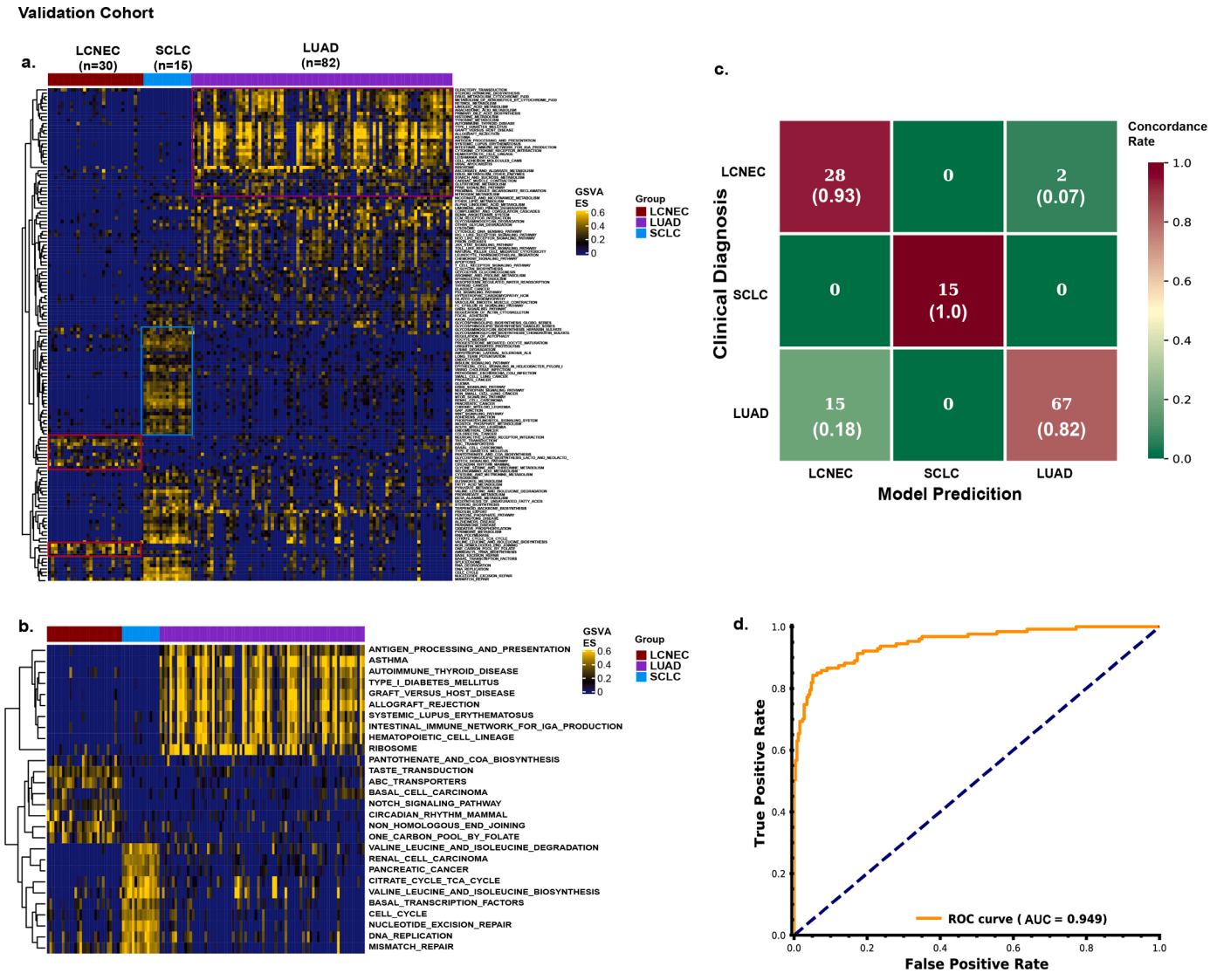

Fig. 2. The GSVA score prediction model construction and evaluation. (a) A total of 13,959 genes were mapped to 186 pathways, heatmap of the enrichment scores (ES) of these pathways in LCNEC, SCLC and NSCLC from validation cohort. (b) Heatmap of the ES for top 28 differentially expressed pathways between LCNEC and SCLC from validation cohort. (c) Confusion matrix showing the model prediction results of validation cohort. (d) ROC curve evaluating the classification performance of the model for LCNEC and SCLC patients.

samples were negative for retinoblastoma protein (RB1) staining; positive staining of neurogenic locus notch homolog (NOTCH) was detected in 3/17 of LCNEC, 2/3 of NSCLC and 4/7 of SCLC samples and occurred in a mutually exclusive fashion with DLL3 (10/17 in LCNEC, 0/3 in NSCLC and 3/7 in SCLC); p40 was identified only in one NSCLC sample.

Three representative examples of H&E staining and IHC results from borderline samples are illustrated in Fig. 4. In sample 1, the diagnostic results were consistent with the pathologists' judgments and model prediction: it showed the classic LCNEC morphology, including a nested growth pattern with peripheral palisading, abundant eosinophilic cytoplasm, coarsely granular chromatin, presented with necrosis, and strong/diffuse staining for four neuroendocrine markers (Fig. 4a). In sample 2, the pathologists diagnosed SCLC, but the model predicted LCNEC, it had the morphological characteristics of SCLC, including tumor cells with scant cytoplasm and round or oval nuclei, and trabecular arrangement. IHC illustrated intense/diffuse staining for CD56, Syn, and INSM1, but weak/focal CgA staining (Fig. 4b). In sample 3, the pathologists diagnosed LCNEC, but the model predicted SCLC, it had untypical morphological characteristics of neuroendocrine carcinoma, but tumor cells with a moderate amount of cytoplasm, with weak CD56 expression, and complete loss of Syn, INSM1, and CgA expression (Fig. 4c).

Clinical outcome of the GSVA score model predicted classification

27 patients with borderline samples that within the model were predicted to have LCNEC had a significant longer DFS than those with the model predicting SCLC (59.0 vs. 6.0 months, $P = 0.0043$, Fig. 5a). No statistically significant difference in DFS was observed between patients with pathologists judging LCNEC or SCLC (20.0 vs. 14.0 months, $P = 0.69$, Fig. 5b), The DFS was significantly longer for LCNEC predicted by the model and treated with a SCLC-regimen than for SCLC predicted by the model and treated with a SCLC-regimen (20.0 vs. 7.0 months, $P = 0.012$, Fig. 5c). However, no difference of DFS was observed between LCNEC treated with SCLC-regimen judged by pathologists and SCLC treated with SCLC-regimen judged by pathologists (20.0 vs. 12.0 months, $P = 0.064$, Fig. 5d) (the treatment regimens are listed in supplementary Table 1). Thus, the results suggested that patients with LCNEC predicted by the model may benefit more from SCLC-regimen. Additionally, limited by the small sample size, no significant

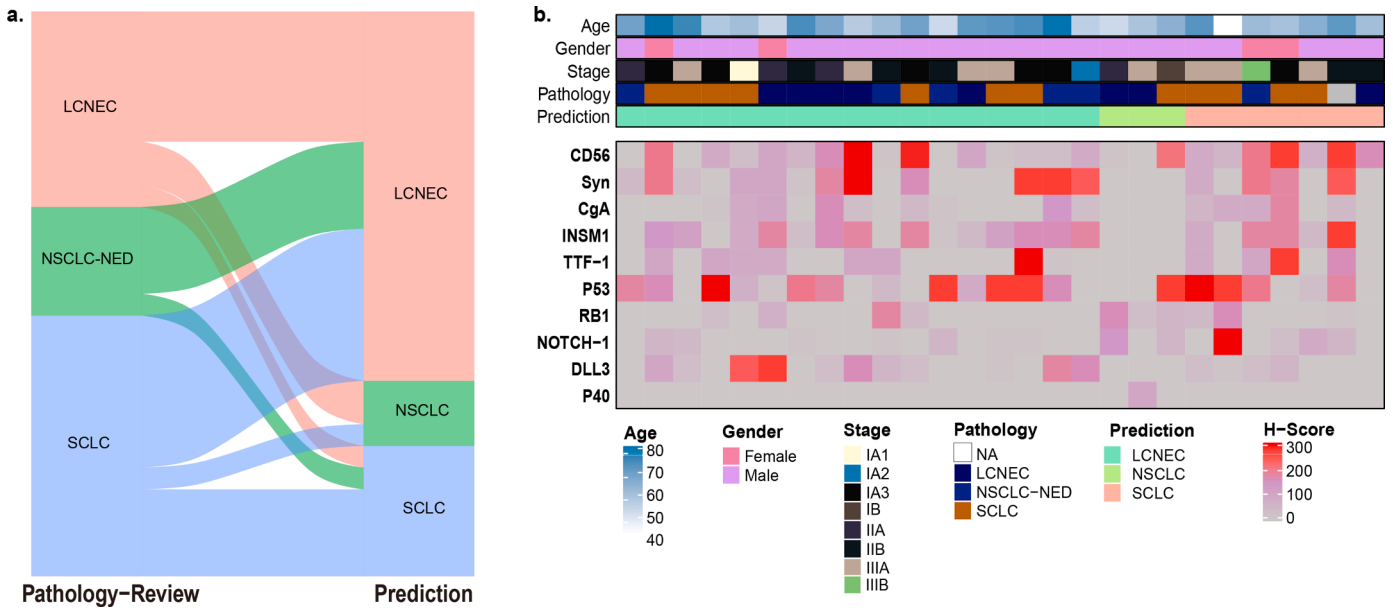

Fig. 3. The Model performance evaluated by borderline samples. (a) The Sankey plot showing the subtype reclassification of 27 pulmonary neuroendocrine tumors between clinicopathological assessment and model prediction. (b) Overview of clinical characteristics, clinical diagnosis and model prediction results, and IHC scores for each neuroendocrine as well as shared genetic marker between LCNEC and SCLC of the 27 borderline samples.

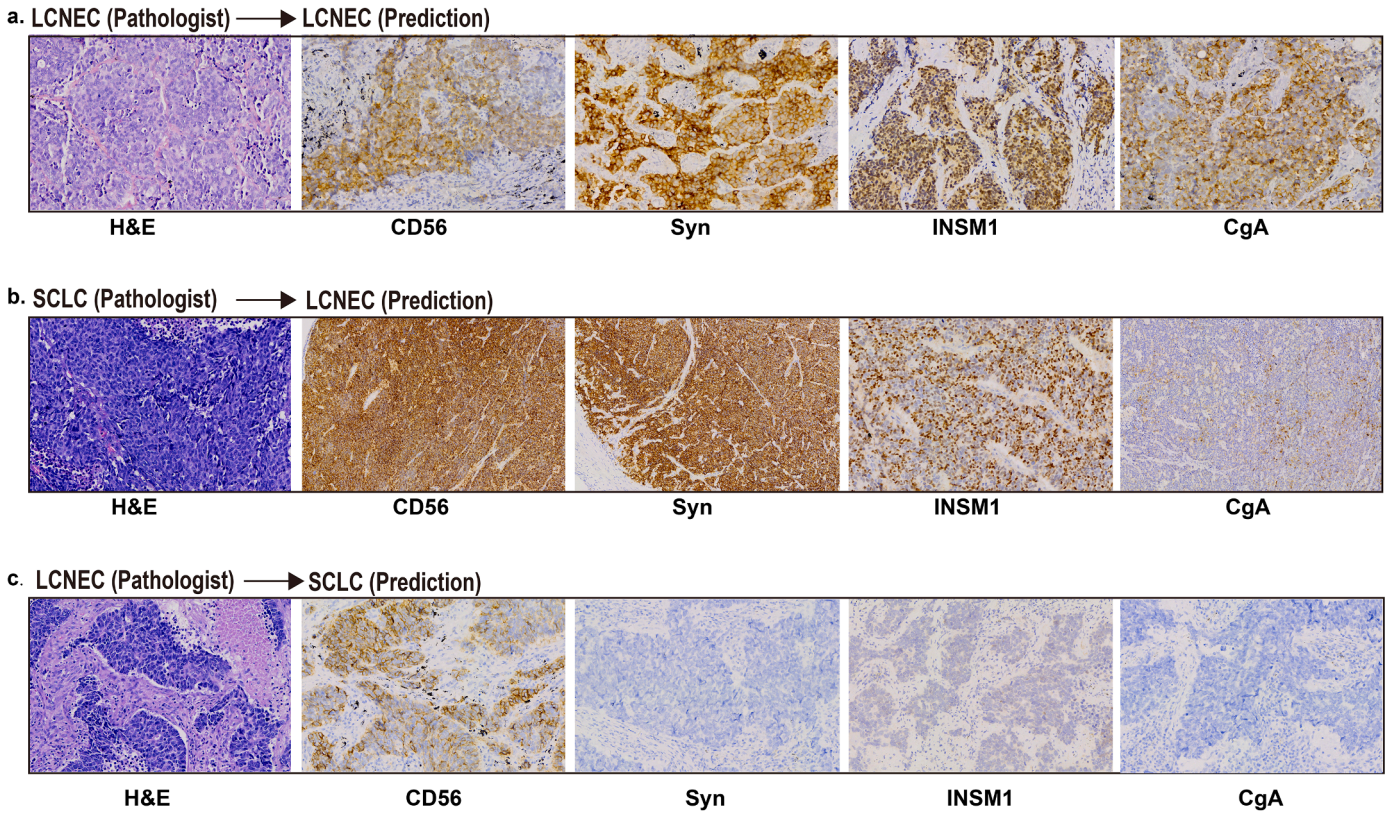

Fig. 4. Morphologic and immunohistochemical features of tree representative examples. (a) sample 1, LCNEC was determined by both pathologist and the model: CD56, Syn, INSM1 and CgA were positive. (b) sample 2, pathological diagnosis was SCLC while the model prediction was LCNEC: CD56, Syn, and INSM1 were positive, CgA was weakly positive. (c) sample 3, pathological diagnosis was LCNEC while the model prediction was SCLC: CD56 was positive, Syn, INSM1 and CgA were negative.

difference in DFS was observed between model-predicted LCNEC with pStage III and model-predicted SCLC with pStage III (10.5 vs. 50.0 months, $P = 0.17$, Fig. 5e). No difference of DFS was observed between LCNEC with pStage III and SCLC with pStage III as judged by pathologists (5.0 vs. 14.0 months, $P = 0.24$, Fig. 5f).

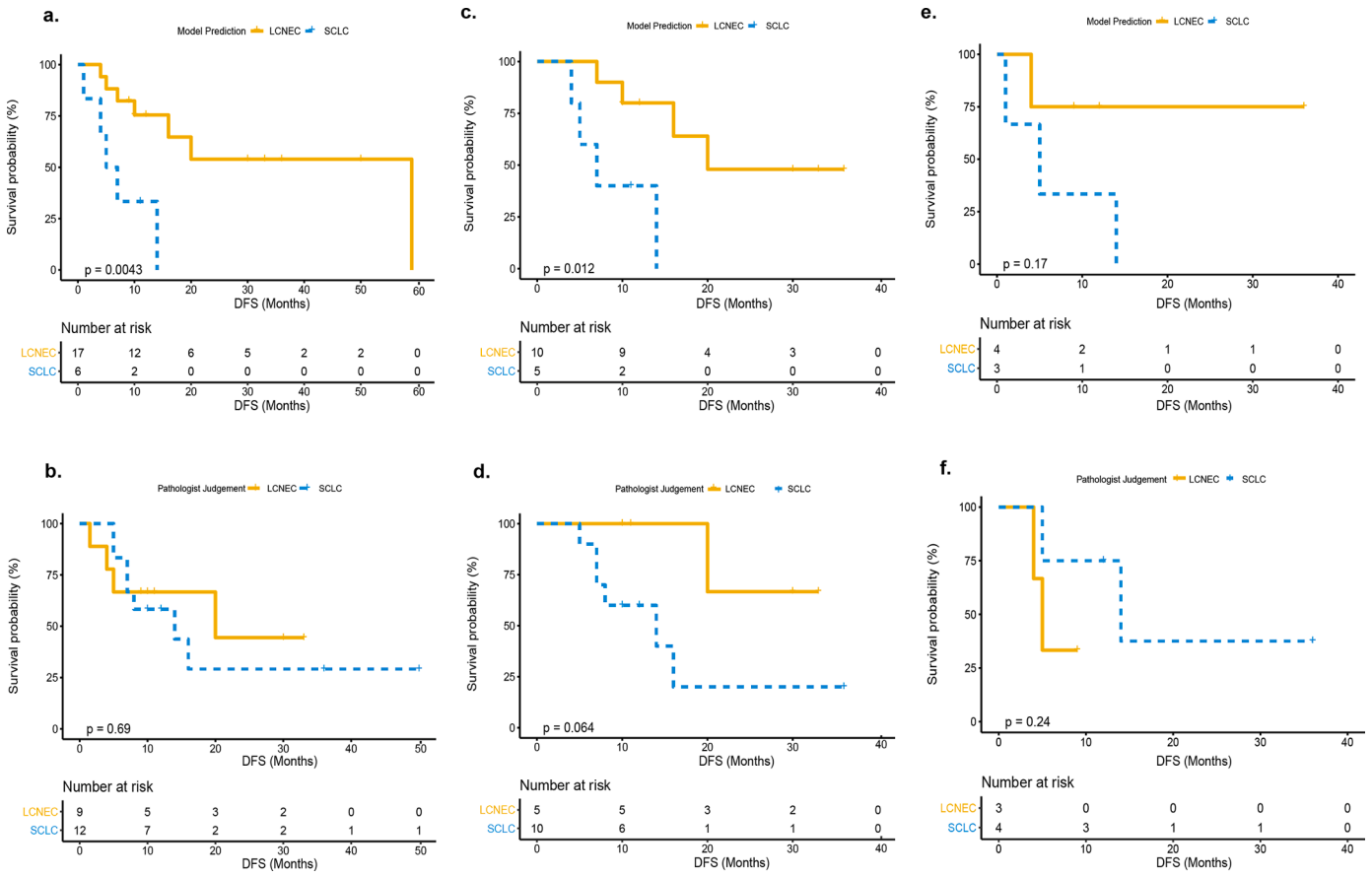

Fig. 5. Disease free survival (DFS) for LCNEC and SCLC in 27 borderline samples subtypes determined by model prediction or pathologists' judgment. (a) DFS for model predicted LCNEC and SCLC. (b) DFS for pathologists judged LCNEC and SCLC. (c) DFS for model predicted LCNEC and SCLC treated with SCLC-regimen. (d) DFS for pathologists judged LCNEC and SCLC treated with SCLC-regimen. (e) DFS for model predicted LCNEC and SCLC with pStage III. (f) DFS for pathologists judged LCNEC and SCLC with pStage III.

Discussion

LCNEC is a type of high-grade neuroendocrine carcinomas (HGNEC) with is highly heterogeneous both histologically and biologically. The differential diagnosis of LCNEC with classical pathological methods can be challenging in some borderline cases as well as in some small biopsy or cytology specimens. Recently, molecular profiling with NGS has revealed a few biologically distinct subsets of LCNEC, and confirmed that chemotherapy regimens have different efficacy in different molecular subtypes of LCNEC. In this study, a transcriptome sequencing data-based classifier was developed to predict the tumor subtype, which enabled prediction of LCNEC from borderline samples efficiently.

There are multiple studies that have shown similarities and differences in molecular alterations between LCNEC and other lung neuroendocrine carcinomas. Researchers hope to find the molecular markers with high specificity and sensitivity to assist the pathological diagnosis of LCNEC. At an individual gene level, inactivation of p16 has been reported in LCNEC and borderline HGNEC, but it is rare in SCLC [25,26]. SKT11, KRAS, KEAP1, LAMA1, PCLO, and MEGF8 mutations are more frequent in LCNEC than in SCLC [16,27]. Conversely, the prevalence of RB1 mutations is lower in LCNEC than in SCLC [27]. As for the expression of biomarkers, Bari et al. performed a gene expression profile using frozen tissues of 8 SCLC and 8 LCNEC samples, and found that the expression of caudal type homeobox 2 (CDX2), Villin 1 (VIL1), and brain-specific angiogenesis inhibitor 3 (BAI3) were significantly different in the two different tumors [28]. Morise et al. analyzed the expression of tumor stem cell-related markers in 60 cases of SCLC and 45 cases of LCNEC. He and his team found that the expression of SOX2 and CD166 in SCLC and LCNEC was significantly different, and SCLC and LCNEC could be distinguished by the expression of these two molecules [29]. It is now known that DLL3 is a downstream target gene of the achaete-scute complex homologue 1 (ASCL1), and is involved in the neuroendocrine differentiation of lung neuroendocrine tumors [30, 31]. Hermans et al. found that the expression of DLL3 was related to ASLC1 with a high rate of 74% (70/94), and the expression of DLL3 in TP53 wild-type and TP53 mutant LCNEC was different [32]. These previous studies showed that a lot of molecular markers were related to the molecular subtype and neuroendocrine spectrum of LCNEC, and may have the potential to differentiate LCNEC from other lung neuroendocrine carcinomas. Therefore, in this study, a batch of prior candidate genes with differences in LCNEC, SCLC and NSCLC at the level of mutation, CNV and transcriptome were selected after literature review to construct the Candidate Genes Model in order to distinguish these three types of tumors. However, the classification effect of this model was not ideal, for instance, the prediction accuracy rate of SCLC for the validation cohort was only 13.3% (2/15). This may be related to the spatio-temporal specificity and random fluctuation of gene expression. In addition, PCA showed that the prediction effect of this model had a batch effect. Therefore, it is not the most ideal choice to construct the prediction model with gene tags.

To further develop an efficiency model to differentiate LCNEC with SCLC, the GSVA algorithm was considered. The GSVA method replaces the original expression value of each gene with the rank order of gene expression, which reduces the error caused by random fluctuation of gene expression. It considers the up-regulation and down-regulation of a pathway as a whole; it is theoretically more robust than that of a single

gene as the classification feature tag [24]. Here, a total of 13,959 genes included in the construction of the Candidate Genes Model were mapped to 186 KEGG pathways. These pathways were scored using the GSVA algorithm to construct the GSVA Score Model. Compared with the Candidate Genes Model, the GSVA Score Model had a higher reliability (AUC: 0.949).

According to the comprehensive integrative genomic and transcriptomic profiling of 75 LCNECs, George et al. has revealed two molecular subgroups: type I LCNECs (37%), characterized by biallelic TP53 and STK11 and/or KEAP1 alterations, and a neuroendocrine differentiation phenotype (ASLC1^{high}/DLL3^{high}/NOTCH^{low}). type II LCNECs (42%), which have biallelic inactivation of TP53 but also RB1, are ASLC1^{low}/DLL3^{low}/NOTCH^{high}, and have upregulation of immune-related pathways [15]. In a retrospective study, LCNEC patients with co-mutant TP53 and RB1 were defined as SCLC-like LCNEC, and those without co-mutant TP53 and RB1 were defined as NSCLC-like LCNEC. The prognosis of SCLC-like patients was worse than of NSCLC-like patients, but the difference was not significant. For SCLC-like patients, compared to gemcitabine/taxane-platinum regimen, etoposide-platinum regimen was associated with longer OS but without statistical significance. Treatment with a gemcitabine/taxane-platinum regimen caused a shorter survival compared to etoposide-platinum or pemetrexed-platinum regimen in NSCLC-like patients [33]. Another study showed that in LCNEC patients with wild-type RB1 gene the gemcitabine/paclitaxel-platinum regimen can significantly prolong the OS of patients compared with an etoposide-platinum based regimen; however, in patients with RB1 mutant LCNEC, no difference in efficacy between the two regimens was observed [34]. Thus, from the aforementioned studies it can be concluded that predicting the prognosis of patients based on molecular typing is still somewhat controversial.

Neuroendocrine morphology and positive staining of neuroendocrine markers are two criteria for LCNEC diagnosis. In cases where the morphology is equivocal, a useful consideration in differential diagnosis is the extent and number of positive neuroendocrine markers. It has been reported that about 20% of LCNECs only show one neuroendocrine marker, and in some cases only show local immunoreactivity of a single marker [35,36]. In contrast, the expression of neuroendocrine markers in adenocarcinoma/large cell carcinoma is usually focal and usually limited to only one marker [8]. However, 1–4% of adenocarcinomas have been reported to express two neuroendocrine markers [37,38]. Here, among the 27 borderline samples, pathological results of 18 samples were inconsistent with the model prediction results, including some samples showing only one neuroendocrine marker, and some cases expressed three to four standard neuroendocrine markers but lack of typical morphological features of LCNEC. In contrast, using the GSVA Score Model, 17 LCNEC, 7 SCLC and 3 NSCLC were predicted from 27 borderline samples. Patients with model-predicted LCNEC had a significant longer DFS those with model-predicted SCLC, and DFS was also significantly longer for model-predicted LCNEC treated with NSCLC-regimen than model-predicted SCLC treated with NSCLC-regimen. Additionally, independent of a treatment with a NSCLC-regimen or a SCLC-regimen, when LCNEC predicted by the model was consistent with the pathologists' judgment, the prognosis of these patients was better. However, when the model prediction was inconsistent with the pathologists' judgment, the prognosis of these patients was worse. Therefore, the prediction results of the GSVA Score Model could be a good supplement to the pathological diagnosis for LCNEC. Further large-scale investigations are warranted to confirm the accuracy of the GSVA Score Model.

Furthermore, to this date, the optimal chemotherapeutic regimen for LCNEC remains unclear (NSCLC vs SCLC regimens). A randomized phase III study on adjuvant chemotherapy in HGNEC using either cisplatin-etoposide or cisplatin-irinotecan was reported recently [39]. The subjects in this study were patients with SCLC including combined SCLC or LCNEC including combined LCNEC. The adjusted primary endpoint relapse-free survival did not show superiority for the irinotecan-based regimen. Although it has not been proven in a specific study that the use of molecular characteristics to distinguishing LCNEC from SCLC can achieve better efficacy, from our analysis and previous studies, molecular-based LCNEC subtypes may have different response to chemotherapy regimens, with the SCLC subtype being more responsive to an etoposide-based and the NSCLC subtype to a gemcitabine/taxanes-based regimen.

Conclusion

In conclusion, this study showed the possibility of predicting LCNEC from borderline samples based on a model constructed by KEGG pathways and a GVSA algorithm. This classification model should be validated and optimized in larger cohorts in order to improve the accuracy of the clinical diagnosis of LCNEC and select patients better to benefit from different treatment regimens.

CRediT authorship contribution statement

Junhong Guo: Visualization, Formal analysis, Writing – original draft. **Likun Hou:** Visualization, Data curation, Formal analysis, Writing – original draft. **Wei Zhang:** . **Zhengwei Dong:** . **Lei Zhang:** Visualization, Data curation. **Chunyan Wu:** Visualization, Data curation.

Declaration of Competing Interest

None declared.

Acknowledgments

This study was partially supported by Shanghai Municipal Commission of Health and Family Planning (No.20184Y0222), "Dream Mentor" training program of Shanghai Pulmonary Hospital (No. fkxr1903), and Shanghai Shenkang Hospital Development Center (SHDC2020CR3047B). We thank Wuzhou Yuan, Zhan Huang, and Changbin Zhu (Amoy Diagnostics Co., Ltd., Xiamen, China) for data extraction, statistical analysis and proofreading the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2021.101222.

References

[1] M. Fasano, C.M. Della Corte, F. Papaccio, F. Ciardiello, F. Morgillo, Pulmonary large-cell neuroendocrine carcinoma: from epidemiology to therapy, J. Thorac. Oncol. Off. Publ. Int. Assoc. Stud. Lung Cancer 10 (2015) 1133–1141, https://doi.org/10.1097/JTO.0000000000000589.

[2] W.D. Travis, E. Brambilla, A.G. Nicholson, Y. Yatabe, J.H.M. Austin, M.B. Beasley, L.R. Chirieac, S. Dacic, E. Duhig, D.B. Flieder, et al., The 2015 world health organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification, J. Thorac. Oncol. Off. Publ. Int. Assoc. Stud. Lung Cancer 10 (2015) 1243–1260, https://doi.org/10.1097/JTO.0000000000000630.

[3] J. Naidoo, M.L. Santos-Zabala, T. Iyriboz, K.M. Woo, C.S. Sima, J.J. Fiore, M. G. Kris, G.J. Riely, P. Lito, A. Iqbal, et al., Large cell neuroendocrine carcinoma of the lung: clinico-pathologic features, treatment, and outcomes, Clin. Lung Cancer 17 (2016) e121–e129, https://doi.org/10.1016/j.cllc.2016.01.003.

[4] S. Yamazaki, I. Sekine, Y. Matsuno, H. Takei, N. Yamamoto, H. Kunitoh, Y. Ohe, T. Tamura, T. Kodama, H. Asamura, et al., Clinical responses of large cell neuroendocrine carcinoma of the lung to cisplatin-based chemotherapy, Lung Cancer 49 (2005) 217–223, https://doi.org/10.1016/j.lungcan.2005.01.008.

[5] J. Le Treut, M.C. Sault, H. Lena, P.J. Souquet, A. Vergnenegre, H. Le Caer, H. Berard, S. Boffa, I. Monnet, D. Damotte, et al., Multicentre phase II study of cisplatin-etoposide chemotherapy for advanced large-cell neuroendocrine lung carcinoma: the GFPC 0302 study, Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. 24 (2013) 1548–1552, https://doi.org/10.1093/annonc/mdt009.

[6] V. Raman, O.K. Jawitz, C.J. Yang, S.L. Voigt, B.C. Tong, T.A. D'Amico, D. H. Harpole, Outcomes for surgery in large cell lung neuroendocrine cancer, J. Thorac. Oncol. Off. Publ. Int. Assoc. Stud. Lung Cancer 14 (2019) 2143–2151, https://doi.org/10.1016/j.jtho.2019.09.005.

[7] D. Uprety, L. Arjyal, Y. Vallatharasu, A. Bista, A. Borgert, A.J. Fitzsimmons, B. M. Parsons, Utilization of surgery and its impact on survival in patients with early stage small-cell lung cancer in the United States, Clin. Lung Cancer 21 (2020) 186–193, https://doi.org/10.1016/j.cllc.2019.07.013, e182.

[8] M.K. Baine, N. Rekhtman, Multiple faces of pulmonary large cell neuroendocrine carcinoma: update with a focus on practical approach to diagnosis, Transl. Lung Cancer Res. 9 (2020) 860–878, https://doi.org/10.21037/tlcr.2020.02.13.

[9] D. Sonkin, A. Thomas, B.A. Teicher, Are neuroendocrine negative small cell lung cancer and large cell neuroendocrine carcinoma with WT RB1 two faces of the same entity? Lung Cancer Manag. 8 (2019) https://doi.org/10.2217/lmt-2019-0005. LMT13.

[10] E. Thunnissen, A.C. Borczuk, D.B. Flieder, B. Witte, M.B. Beasley, J.H. Chung, S. Dacic, S. Lantuejoul, P.A. Russell, M. den Bakker, et al., The use of immunohistochemistry improves the diagnosis of small cell lung cancer and its differential diagnosis. an international reproducibility study in a demanding set of cases, J. Thorac. Oncol. Off. Publ. Int. Assoc. Stud. Lung Cancer 12 (2017) 334–346, https://doi.org/10.1016/j.jtho.2016.12.004.

[11] W.D. Travis, A.A. Gal, T.V. Colby, D.S. Klimstra, R. Falk, M.N. Koss, Reproducibility of neuroendocrine lung tumor classification, Hum. Pathol. 29 (1998) 272–279, https://doi.org/10.1016/s0046-8177(98)90047-8.

[12] M.A. den Bakker, S. Willemsen, K. Grunberg, L.A. Noorduijn, M.F. van Oosterhout, R.J. van Suylen, W. Timens, B. Vrugt, A. Wiersma-van Tilburg, F.B. Thunnissen, Small cell carcinoma of the lung and large cell neuroendocrine carcinoma interobserver variability, Histopathology 56 (2010) 356–363, https://doi.org/10.1111/j.1365-2559.2010.03486.x.

[13] S.Y. Ha, J. Han, W.S. Kim, B.S. Suh, M.S. Roh, Interobserver variability in diagnosing high-grade neuroendocrine carcinoma of the lung and comparing it with the morphometric analysis, Korean J. Pathol. 46 (2012) 42–47, https://doi.org/10.4132/KoreanJPathol.2012.46.1.42.

[14] J.L. Derks, N. Leblay, S. Lantuejoul, A.C. Dingemans, E.M. Speel, L. Fernandez-Cuesta, New insights into the molecular characteristics of pulmonary carcinoids and large cell neuroendocrine carcinomas, and the impact on their clinical management, J. Thorac. Oncol. Off. Publ. Int. Assoc. Stud. Lung Cancer 13 (2018) 752–766, https://doi.org/10.1016/j.jtho.2018.02.002.

[15] J. George, V. Walter, M. Peifer, L.B. Alexandrov, D. Seidel, F. Leenders, L. Maas, C. Muller, I. Dahmen, T.M. Delhomme, et al., Integrative genomic profiling of large-cell neuroendocrine carcinomas reveals distinct subtypes of high-grade neuroendocrine lung tumors, Nat. Commun. 9 (2018) 1048, https://doi.org/10.1038/s41467-018-03099-x.

[16] N. Rekhtman, M.C. Pietanza, M.D. Hellmann, J. Naidoo, A. Arora, H. Won, D. F. Halpenny, H. Wang, S.K. Tian, A.M. Litvak, et al., Next-generation sequencing of pulmonary large cell neuroendocrine carcinoma reveals small cell carcinoma-like and non-small cell carcinoma-like subsets, Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 22 (2016) 3618–3629, https://doi.org/10.1158/1078-0432.CCR-15-2946.

[17] L. Meder, K. Konig, L. Ozretic, A.M. Schultheis, F. Ueckeroth, C.P. Ade, K. Albus, D. Boehm, U. Rommerscheidt-Fuss, A. Florin, et al., NOTCH, ASCL1, p53 and RB alterations define an alternative pathway driving neuroendocrine and small cell lung carcinomas, Int. J. Cancer 138 (2016) 927–938, https://doi.org/10.1002/ijc.29835.

[18] J. George, J.S. Lim, S.J. Jang, Y. Cun, L. Ozretic, G. Kong, F. Leenders, X. Lu, L. Fernandez-Cuesta, G. Bosco, et al., Comprehensive genomic profiles of small cell lung cancer, Nature 524 (2015) 47–53, https://doi.org/10.1038/nature14664.

[19] M. Peifer, L. Fernandez-Cuesta, M.L. Sos, J. George, D. Seidel, L.H. Kasper, D. Plenker, F. Leenders, R. Sun, T. Zander, et al., Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer, Nat. Genet. 44 (2012) 1104–1110, https://doi.org/10.1038/ng.2396.

[20] C.M. Rudin, S. Durinck, E.W. Stawiski, J.T. Poirier, Z. Modrusan, D.S. Shames, E. A. Bergbower, Y. Guan, J. Shin, J. Guillory, et al., Comprehensive genomic analysis identifies SOX_2 as a frequently amplified gene in small-cell lung cancer, Nat. Genet. 44 (2012) 1111–1116, https://doi.org/10.1038/ng.2405.

[21] J.D. Campbell, A. Alexandrov, J. Kim, J. Wala, A.H. Berger, C.S. Pedamallu, S. A. Shukla, G. Guo, A.N. Brooks, B.A. Murray, et al., Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas, Nat. Genet. 48 (2016) 607–616, https://doi.org/10.1038/ng.3564.

[22] Cancer Genome Atlas Research N, Comprehensive molecular profiling of lung adenocarcinoma, Nature 511 (2014) 543–550, https://doi.org/10.1038/nature13385.

[23] Clinical Lung Cancer Genome P, Network Genomic M, A genomics-based classification of human lung tumors, Sci. Transl. Med. 5 (2013), https://doi.org/10.1126/scitranslmed.3006802, 209ra153.

[24] S. Hanzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7, https://doi.org/10.1186/1471-2105-14-7.

[25] O. Washimi, M. Nagatake, H. Osada, R. Ueda, T. Koshikawa, T. Seki, T. Takahashi, T. Takahashi, *In vivo* occurrence of p16 (MTS1) and p15 (MTS2) alterations preferentially in non-small cell lung cancers, Cancer Res. 55 (1995) 514–517.

[26] K. Hiroshima, A. Iyoda, K. Shibuya, Y. Haga, T. Toyozaki, T. Iizasa, T. Nakayama, T. Fujisawa, H. Ohwada, Genetic alterations in early-stage pulmonary large cell neuroendocrine carcinoma, Cancer 100 (2004) 1190–1198, https://doi.org/10.1002/cncr.20108.

[27] T. Miyoshi, S. Umemura, Y. Matsumura, S. Mimaki, S. Tada, H. Makinoshima, G. Ishii, H. Udagawa, S. Matsumoto, K. Yoh, et al., Genomic profiling of large-cell neuroendocrine carcinoma of the lung, Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 23 (2017) 757–765, https://doi.org/10.1158/1078-0432.CCR-16-0355.

[28] M.F. Bari, H. Brown, A.G. Nicholson, K.M. Kerr, J.R. Gosney, W.A. Wallace, I. Soomro, S. Muller, D. Peat, J.D. Moore, et al., BAI3, CDX2 and VIL1: a panel of three antibodies to distinguish small cell from large cell neuroendocrine lung carcinomas, Histopathology 64 (2014) 547–556, https://doi.org/10.1111/his.12278.

[29] M. Morise, T. Hishida, A. Takahashi, J. Yoshida, Y. Ohe, K. Nagai, G. Ishii, Clinicopathological significance of cancer stem-like cell markers in high-grade neuroendocrine carcinoma of the lung, J. Cancer Res. Clin. Oncol. 141 (2015) 2121–2130, https://doi.org/10.1007/s00432-015-1985-3.

[30] T. Jiang, B.J. Collins, N. Jin, D.N. Watkins, M.V. Brock, W. Matsui, B.D. Nelkin, D. W. Ball, Achaete-scute complex homologue 1 regulates tumor-initiating capacity in human small cell lung cancer, Cancer Res. 69 (2009) 845–854, https://doi.org/10.1158/0008-5472.CAN-08-2762.

[31] R.M. Henke, D.M. Meredith, M.D. Borromeo, T.K. Savage, J.E. Johnson, Ascl1 and Neurog2 form novel complexes and regulate delta-like3 (Dll3) expression in the neural tube, Dev. Biol. 328 (2009) 529–540, https://doi.org/10.1016/j.ydbio.2009.01.007.

[32] B.C.M. Hermans, J.L. Derks, E. Thunnissen, R.J. van Suylen, M.A. den Bakker, H.J. M. Groen, E.F. Smit, R.A. Damhuis, E.C. van den Broek, et al., DLL3 expression in large cell neuroendocrine carcinoma (LCNEC) and association with molecular subtypes and neuroendocrine profile, Lung Cancer 138 (2019) 102–108, https://doi.org/10.1016/j.lungcan.2019.10.010.

[33] M. Zhuo, Y. Guan, X. Yang, L. Hong, Y. Wang, Z. Li, R. Chen, H.A. Abbas, L. Chang, Y. Gong, et al., The prognostic and therapeutic role of genomic subtyping by sequencing tumor or cell-free DNA in pulmonary large-cell neuroendocrine carcinoma, Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 26 (2020) 892–901, https://doi.org/10.1158/1078-0432.CCR-19-0556.

[34] J.L. Derks, N. Leblay, E. Thunnissen, R.J. van Suylen, M. den Bakker, H.J.M. Groen, E.F. Smit, R. Damhuis, E.C. van den Broek, A. Charbrier, et al., Molecular subtypes of pulmonary large-cell neuroendocrine carcinoma predict chemotherapy treatment outcome, Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 24 (2018) 33–42, https://doi.org/10.1158/1078-0432.CCR-17-1921.

[35] M.K. Baine, J.H. Sinard, G. Cai, R.J. Homer, A semiquantitative scoring system may allow biopsy diagnosis of pulmonary large cell neuroendocrine carcinoma, Am. J. Clin. Pathol. 153 (2020) 165–174, https://doi.org/10.1093/ajcp/aqz149.

[36] J.L. Derks, A.C. Dingemans, R.J. van Suylen, M.A. den Bakker, R.A.M. Damhuis, E. C. van den Broek, E.J. Speel, E. Thunnissen, Is the sum of positive neuroendocrine immunohistochemical stains useful for diagnosis of large cell neuroendocrine carcinoma (LCNEC) on biopsy specimens? Histopathology 74 (2019) 555–566, https://doi.org/10.1111/his.13800.

[37] B. Ye, J. Cappel, J. Findeis-Hosey, L. McMahon, Q. Yang, G.Q. Xiao, H. Xu, F. Li, hASH1 is a specific immunohistochemical marker for lung neuroendocrine tumors, Hum. Pathol. 48 (2016) 142–147, https://doi.org/10.1016/j.humpath.2015.09.019.

[38] D.N. Ionescu, D. Treaba, C.B. Gilks, S. Leung, D. Renouf, J. Laskin, R. Wood-Baker, A.M. Gown, Nonsmall cell lung carcinoma with neuroendocrine differentiation–an entity of no clinical or prognostic significance, Am. J. Surg. Pathol. 31 (2007) 26–32, https://doi.org/10.1097/01.pas.0000213319.04919.97.

[39] H. Kenmotsu, S. Niho, M. Tsuboi, M. Wakabayashi, G. Ishii, K. Nakagawa, H. Daga, H. Tanaka, H. Saito, K. Aokage, et al., Randomized phase III study of irinotecan plus cisplatin versus etoposide plus cisplatin for completely resected high-grade neuroendocrine carcinoma of the lung: JCOG1205/1206, J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. 38 (2020) 4292–4301, https://doi.org/10.1200/JCO.20.01806.