scientific reports

Check for updates

OPEN

Mapping the immune landscape in small cell lung cancer by analysing expression of immuno-modulators in tissue biopsies and paired blood samples

Rimlee Dutta[1], Amber Rathor[1], Hanuman Prasad Sharma[2], Hem Chandra Pandey[3], Prabhat Singh Malik[4], Anant Mohan[5], Aruna Nambirajan[1], Rajeev Kumar[6] & Deepali Jain[1]✉

Small cell lung carcinomas (SCLC) are aggressive tumors with high propensity to metastasize. Recent NCCN guidelines have incorporated immunotherapy in extensive stage SCLC. Limited benefit in few patients compounded by side effects of unwonted immune-checkpoint-inhibitor (ICPI) usage necessitates identification of potential biomarkers predicting response to ICPIs. Attempting this, we analysed expression of various immunoregulatory molecules in tissue biopsies and paired blood samples of SCLC patients. In 40 cases, immunohistochemistry for expression of immune inhibitory receptors CTLA-4, PD-L1 and IDO1 was performed. Matched blood samples were quantified for IFN-γ, IL-2, TNF-α and sCTLA-4 levels using immunoassay and additionally for IDO1 activity (Kynurenine/Tryptophan ratio) using LC–MS. Immunopositivity for PD-L1, IDO1 and CTLA-4 was identified in 9.3%, 6.2% and 71.8% cases, respectively. Concentration of serum IFN-γ (p-value < 0.001), TNF-α (p-value = 0.025) and s-CTLA4 (p-value = 0.08) were higher in SCLC patients while IL-2 was lower (p-value = 0.003) as compared to healthy controls. IDO1 activity was significantly elevated in SCLC cohort (p-value = 0.007). We proffer that SCLC patients show immune suppressive milieu in their peripheral circulation. Analysis of CTLA4 immunohistochemical expression along with s-CTLA4 levels appears prospective as biomarkers for predicting responsiveness to ICPIs. Additionally, evaluation of IDO1 appears cogent both as prognostic marker and potential therapeutic target as well.

Accounting for about 15% of lung cancers globally, small cell lung cancer (SCLC) is a belligerent tumor subtype with approximately 70% of cases presenting with extensive disease at diagnosis[1]. Conventional tumor staging divides SCLC into limited stage (LS-SCLC) and extensive stage (ES-SCLC) disease. In most patients, for many years, platinum based chemotherapy has formed the standard first line of therapy for ES-SCLC with overall response rate of 40–70%, however responses to second line chemotherapy has been not so sanguine[2]. Despite therapy, SCLC are aggressive with a high propensity to metastasize or recur. This has led to exploration of newer modalities of therapy, with added emphasis on targeted therapy and immunotherapy. The relevance of immunogenicity of SCLC, has been evinced by recent studies, based on tumor infiltration by numerous T-lymphocytes along with occurrence of frequent paraneoplastic syndromes owing to autoantibody production and the close relation of SCLC with smoking mutational signature, characterized by high tumor mutational burden[3]. Though relative benefit was observed with immunotherapeutic agents like nivolumab and ipilimumab on comparison with chemotherapy in randomized clinical trials, studies have also shown that these benefits are limited to a small portion of patients; ideating that SCLC may be associated with a different immunological microenvironment[4].

[1]Department of Pathology, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India. [2]Bioanalytics Facility, Centralized Core Research Facility, All India Institute of Medical Sciences, New Delhi, India. [3]Department of Transfusion Medicine, All India Institute of Medical Sciences, New Delhi, India. [4]Department of Medical Oncology, Dr B.R.A. Institute Rotary Cancer Hospital, All India Institute of Medical Sciences, New Delhi, India. [5]Department of Pulmonary Medicine, All India Institute of Medical Sciences, New Delhi, India. [6]Delhi Cancer Registry, Dr B.R.A. Institute Rotary Cancer Hospital, All India Institute of Medical Sciences, New Delhi, India. ✉email: deepalijain76@gmail.com

Evaluation of biomarkers, such as smoking status, tumor mutation burden (TMB), and programmed cell death-ligand 1 (PD-L1) expression, have till date not been able to predict outcome[5]. Currently, SCLC is no longer considered a single-disease entity, with subtypes being classified into neuroendocrine (NE)-high and NE-low tumors, defined by distinct RNA gene expression profiles, which can possibly translate into different immunogenic profiles[6]. NE-high tumors, characterized by decreased immune cell infiltration are defined as cold or 'immune desert' phenotype, based on low levels of immune cell-related RNA expression while NE-low tumors are associated with increased immunogenicity, classified as 'hot' or 'immune oasis' phenotype, possibly more likely to respond to immunotherapies[7]. As most SCLCs are inoperable at the time of diagnosis, and relapses are frequently witnessed within the first year of treatment, use of newer modalities like immunotherapy appears cogent. The recent NCCN guidelines have also incorporated immunotherapy in ES-SCLC and SCLC with recurrence within six months of therapy[8]. However, in practical scenario owing to small improvement in overall survival and limited benefit extended to small number of patients, there arises an overwhelming need to identify biomarkers able to predict response to immune checkpoint inhibitors (ICPI) in SCLC patients.

Indoleamine-2,3-dioxygenase 1 (IDO1), which is a cytosolic enzyme catalyzing the rate-limiting step of tryptophan (Trp) catabolism to kynurenine (Kyn), has been found to be a key factor in defining cancer immunogenicity. Accumulation of Trp metabolites promotes the differentiation of T-reg cells and induces apoptosis of effector T-cells with consequent immunosuppression. IDO1 overexpression in tumors exploits immunosuppressive mechanisms to promote their spread and render poor prognosis. Thereby, serum Kyn/Trp concentration ratio has been explored as a potential biomarker of cancer-associated immune suppression in many tumors, including lung carcinomas, particularly non-small cell lung carcinomas (NSCLC)[9]. CTLA-4 (Cytotoxic T-lymphocyte-associated antigen-4, CD152), a CD28 homologue is translocated from intracellular storage to plasma membrane of T-cells, competitively binding to B7 ligands on antigen presenting cells (APCs) with higher affinity, thereby preventing CD28-mediated T-cell activation. However, soluble cytotoxic T-lymphocyte antigen 4 (sCTLA-4), one of the isoforms of CTLA-4, plays an important role in down-regulating the negative signal of CTLA-4 in T-cell responses and studies have suggested the favourable prognostic value of serum sCTLA-4 level in tumor patients[10]. Hence, along with expression of programmed death-1 ligand (PD-L1) on tumor cells, expression of CTLA-4 in the tumor microenvironment and its correlation with serum sCTLA- holds immense therapeutic potential.

Interferon-gamma (IFN-γ), a functionally pleiotropic cytokine modulating the expression of numerous proteins mainly upregulates PD-L1 expression[11]. IFN-γ expression in tumour tissues has been associated with response to treatment and studies have shown direct correlation of IFN-γ levels with response to ICPI in NSCLCs[12]. Levels of interleukin-2 (IL-2), a cytokine that binds to specific receptors expressed on T-cells and natural killer cells and thereby mediates their immune-stimulating effects, have been seen to correlate with long term survival in SCLC patients[13]. The therapeutic value of IL-2, highlighting its requirement in the efficacy of PD-1/L1 ICPI therapy has also been studied in NSCLC patients[14]. Tumor necrosis factor-alpha (TNF-α), an inflammatory cytokine promotes tumor growth and higher serum levels of TNF-α have been reported to be associated with poor prognosis in cancer patients. Preclinical TNF-α blockade has been touted to improve therapeutic effectiveness of ICPIs[15]. Although the prospective utility of all the above biomarkers in relation to ICPIs have been extensively studied and established in NSCLCs, their expression and future therapeutic potential in SCLCs remains largely undetermined.

In this study, we tried to characterize the immune landscape of SCLC by analysing the expression of various immune regulatory molecules on tumor cells as well as on immune cells present in the tumor microenvironment. Additionally, an attempt was also made to evaluate the levels of immune-regulatory, inflammatory and other protein mediators in blood of these patients, as biomarkers predicting potential response to ICPIs.

Materials and methods

The study was of prospective design and approved by the Institute Ethics Committee of All Institute of Medical Sciences (AIIMS, New Delhi) [IEC-552/2016, RP-20/2016, OP-1/2021]. Cases of SCLC over a period of 3 years (2019–2022) were enlisted in the study. Histologically proven biopsy samples with adequate tissue in the formalin fixed paraffin embedded (FFPE) block were included. Paired peripheral blood samples were collected in all available patients. Respective serum and plasma of these cases were stored at −80 °C for further quantification studies. The study was performed in accordance with relevant guidelines/regulations (in accordance with the Declaration of Helsinki). Informed consent was obtained from all participants in the study, including patients and healthy controls.

Immunohistochemistry was performed on FFPE samples for checking expression of inhibitory receptors (CTLA-4, PD-L1 and IDO1) and scoring was done on basis of the intensity and percentage of tumor cells stained. Matched serum samples collected from SCLC patients were quantified for following markers (IFN-γ, IL-2 and TNF-α) using immunoassay and additionally, for soluble CTLA-4 (sCTLA-4) using enzyme linked immunosorbent assay (ELISA). For comparison of above four mentioned serum markers, serum samples from 30 healthy individuals matched for age and gender were collected from donors in blood bank after ethical approval and informed consent, and processed in a similar manner. For determining plasma IDO1 activity, available plasma samples from SCLC patients were assessed for Kyn/Trp ratio, using liquid chromatography mass spectroscopy (LC–MS/MS) method. 40 plasma samples of age and gender-matched healthy controls were also analysed for comparison (collected after ethical approval and informed consent). A brief workflow of the study along with available samples for analyses has been shown in Fig. 1.

Immunohistochemical (IHC) analysis of PD-L1, CTLA-4 and IDO1 in SCLC biopsies. 5 micron sections from selected FFPE blocks were subjected to IHC for PD-L1 (clone SP263) which was performed on Ventana benchmark GX platform optimized with the Optiview DAB IHC detection kit. Sections of placenta

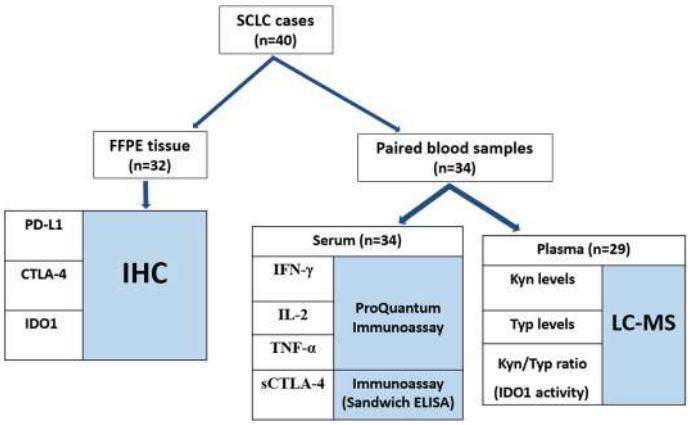

Figure 1. Workflow of the study along with available samples for analyses.

were included as positive controls for endogenous PD-L1 expression. The guidelines for interpretation of PD-L1 IHC were adopted from the interpretation guide by Ventana Inc. for PD-L1 (SP263) assay staining of NSCLC. Any amount of membranous staining (whether discontinuous, circumferential or baso-lateral) of tumor cells at any intensity (greater than background staining) was considered positive while any cytoplasmic staining was disregarded. Tumor proportion score (TPS) was determined by assessing percentage of cells stained amongst viable tumor cells and 1% positivity was considered as the cut-off for PD-L1 positivity. Immunopositivity in the background immune cells like lymphocytes and intra-alveolar macrophages were not included in the scoring criteria. Other IHC markers such as CTLA-4 (sc-376016 clone, 1:50 dilution, Santa cruz) and IDO1 (VIN3IDO clone, 1:2000 dilution, eBiosciences) were performed manually. Sections from lymph nodes were used as controls. Immunohistochemical expression were scored based on the percentage of immunolabelled cells as: 0%—score 0, 1% to 33%—score 1, 34% to 66%—score 2 and 67% to 100%—score 3. Intensity of immunoexpression were scored as: score 1—mild, score 2—moderate, score 3—strong staining.

Quantification of IFN-γ, IL-2 and TNF-α in serum. 34 matched serum samples collected from SCLC patients were quantified for following markers (IFN-γ, IL-2 and TNF-α) using ProQuantum immunoassay kits (Thermofisher Scientific). ProQuantum high sensitivity immunoassays utilize a matched pair of target-specific antibodies, each conjugated to a DNA oligonucleotide. During antibody analyte binding, the two DNA oligos are brought into close proximity, which allows for ligation of the two strands and subsequent creation of a template strand for amplification. Serum samples were added to an antibody-oligo conjugate mix provided in respective kits and incubated at room temperature for 1 h or overnight at 4 °C. Following the initial incubation, master mix and ligase provided in the kit was added to the PCR assay plate followed by the instructed PCR thermal cycling: 20 °C hold for 25 min, 95 °C hold for 2 min, 40 cycles of 95 °C for 15 s and 60 °C for 1 min. After the run was complete, the results of qPCR file were imported using available ProQuantum cloud-based software. The data was then analysed to obtain protein concentration values in pg/mL.

Quantification of sCTLA-4 in serum using ELISA immunoassay. The Human CTLA-4 solid-phase sandwich ELISA kit (enzyme-linked immunosorbent assay) was used to measure the amount of sCTLA-4 in 34 matched serum samples collected from SCLC patients. Appropriate standards diluted as per manufacturer's instructions were added to the coated micro-well plate followed by samples from SCLC patients and healthy controls in 1:2 dilution and allowed to bind to the immobilized (capture) antibody at 37 °C for 2 h. The sandwich formed by the addition of the second (detector) antibody was then allowed to incubate overnight at 4 °C. Finally, substrate solution was added that reacted with the enzyme-antibody-target complex to produce measurable signals (absorbance at 450 nm). The intensity of these signals were then measured which was directly proportional to the concentration of sCTLA-4 present in the original sample.

Assessment of IDO1 activity (Kyn/Trp ratio) in plasma by liquid chromatography mass spectrometry (LC–MS). A previously published LC–MS/MS method with minor modifications was used for the analysis[16].

Instrumentation. A triple quadrupole tandem mass spectrometer (4000 Q-Trap, AB Sciex, CA, USA) coupled with high performance liquid chromatography system (HPLC, Agilent Technologies, 1260 Infinity, CA, USA) was used. The HPLC system consisted of online degasser, quaternary pump, multi-sampler, thermostatted column compartment and variable wavelength UV detector. All the parameters of tandem mass spectrometer and HPLC were controlled by Analyst software, version 1.7.1 (AB Sciex, CA, USA) and Open-LAB control panel software (Agilent Technologies, 1260 Infinity, CA, USA), respectively.

Chromatographic and mass spectrometer conditions. The chromatographic separation was achieved using Zic®-cHILIC column (100 × 4.6 mm, 3 µm, Merck, Germany). An isocratic mobile phase using

acetonitrile with 0.1% formic acid (solvent A) and 5 mM ammonium acetate with 0.1% formic acid (solvent B) at a ratio of 60:40 (A:B) was used. The mobile phase was pumped at a flow rate of 0.5 mL/min with a total run time of 5.5 min. The column oven compartment was kept at ambient and the autosampler tray was maintained at 10 ± 1 °C. The standards and samples were loaded onto 96 well plate and injected at a volume of 2 µL.

All the analytes were ionized using Turbo Ion Spray (ESI) source operated at positive ion mode. Compound related parameters for Kyn, Trp and homatropine (HA) were optimized by infusing 100 ng/ml solution at 0.01 mL/min flow rate through Harvard pump (Harvard Company, Reno, Nevada, USA). HA served as an internal standard (IS) for Kyn and Trp. Multiple reaction monitoring (MRM) was used for the quantification of the analytes. Source dependent parameters like collision associated dissociation (CAD), curtain gas (CUR), ionizing voltage (iV), temperature, gas 1 and gas 2 were kept at 12 psi, 30 psi, 5500 V, 500 °C and 50 psi each, respectively.

Calibration standard preparation. Kyn and Trp were accurately weighed at an amount of 10 mg and transferred to 10 mL amber colored volumetric flask. The powder was dissolved in 1:1 acetonitrile:water (10 mL). The stock solution was stored at 4 °C and used for further analysis. A calibration dilution was prepared in charcoal stripped human plasma where the range for Kyn was from 0.156 µg/mL to 5 µg/mL whereas Trp ranged from 0.625 to 20 µg/mL. 20 µL of each calibration standard was taken into a fresh tube and 200 µL of extraction solvent (80% acetonitrile in water with 0.1% formic acid and 50 ng/mL of HA as internal standard) was added to it. The solution was vortexed for 1 min and centrifuged at 7840*g* for 10 min at 4 °C. The supernatant was transferred to 96 well plate and subjected to LC–MS/MS analysis.

Sample preparation. Plasma from patients as well as healthy controls were thawed at room temperature and vortexed briefly. As described above for the calibration standards, the samples were prepared in a similar manner.

Statistical analysis. Statistical analysis was performed using the statistical software STATA Version 16.1 (STATA Corporation, College Station, TX, USA). Data were expressed as mean ± standard deviation (SD) or median [interquartile range (IQR)] for continuous variables and number (percentage) for categorical data, as appropriate. Shapiro Wilks test was applied to test the normality of distribution. Differences between the control and SCLC group were tested by using Independent t-test or Mann–Whitney test (for continuous variables) or Chi-square/ Fisher's exact test (for categorical variables). The receiver operating characteristic (ROC) curve was applied to compare the strength of the classification of bio-markers between the control and SCLC group. A p-value less than 0.05 was considered statistically significant.

Results

Clinical demographics. A total of 40 SCLC cases were eligible to be included in this study. The age of the patients ranged from 27 to 76 years, with the median age being around 51.5 years. There was a male preponderance (male: female ratio being around 9:1). Smoking index (SI) ranged from 250 to 700 in 25 of 40 prospective cases, three patients exclusively chewed tobacco while data was not available in the rest of the cases. Data regarding tumor stage showed that 29 (72.5%) patients had ES-SCLC disease while 11 (27.5%) patients had LS-SCLC. The ECOG scores were 1, 2 and 3 in 8, 29 and 3 patients respectively; at the time of initial evaluation. 34 patients had progressive disease on latest follow-up visit including three patients who died of disease while 3 patients were lost to follow up (Table 1).

Immunohistochemical (IHC) analysis of PD-L1, CTLA-4 and IDO1 in SCLC biopsies. Only three cases (9.3%, 3/32) showed immunopositivity for PD-L1 in the tumor cells. The TPS in these three positive cases were 1%, 10% and > 50% respectively. Out of 32 cases, IDO1 was found to be positive in only two cases (6.2%). Strong IDO1 immunostaining (score = 3) were seen on approximately 75% and 80% of endothelial cells, intra-tumoral immune cells and peritumoral stromal cells in these two cases respectively. Immunostaining for CTLA-4 was seen in cytoplasm of tumor cells. CTLA-4 immunostaining was present in majority of the cases (71.8%, 23/32). The graphical representation of the intensity and percentage distribution of CTLA-4 immunoexpression across the SCLC cohort have been included in Fig. 2.

IHC analyses could not be performed on 8 cases due to exhaustion of representative tumor tissue in the FFPE block, after performing routine immunohistochemical testing at the time of diagnosis.

Quantification of IFN-γ, IL-2 and TNF-α and sCTLA-4 in serum. For TNF-α, IL-2 and IFN-γ estimation, serum samples of 34 SCLC patients were quantified using ProQuantum immunoassay kit (done in triplicates). Serum concentrations of IFN-γ, IL-2 and TNF-α were interpolated against a standard curve. The normality assumption was violated among all these three biomarkers, the box-whisker plots showing the distribution are presented in Fig. 3. The concentration of IFN-γ was found to be significantly higher in SCLC patients [median (IQR) = 0.813 {0.369 to 2.088} pg/mL; range = 0.099 pg/mL to 576.4 pg/mL; p-value < 0.001] than healthy controls [median = 0.0165 (0.013 to 0.295) pg/mL]. The IL-2 concentrations were found to be significantly lower in SCLC [median (IQR) = 140.43 {3.899 to 407.44} pg/mL; p-value = 0.003] than in healthy controls [median = 519.8 (60.66 to 681.28) pg/mL]. The TNF-α concentration of total 34 samples interpolated against a standard curve showed that concentration in SCLC patients was significantly higher [median (IQR) = 0.269 {0.066 to 0.460} pg/mL; p-value = 0.025) than healthy controls [median (IQR) = 0.0695 {0.054 to 0.107} pg/mL], with exception of two controls who showed high levels of TNF-α (Fig. 3).

Parameter	Number of patients (n)	
Total number of patients	40	
Age range (median—51.5 years)		
< 25 years	0	0%
25–50 years	12	30%
51–75 years	27	67.5%
> 75 years	1	2.5%
Male:female ratio	36 males:4 females	9:1
Smoking index (SI) (available in 25 of 40 cases)		
< 250	2	8%
250–500	10	40%
501–700	13	52%
> 700	0	0%
History of tobacco chewing	3	7.5%
Associated co-morbidities	1 patient had diabetes mellitus	2.5%
ES-SCLC disease	29	72.5%
LS-SCLC disease	11	27.5%
ECOG scores at evaluation		
Score 1	8	20%
Score 2	29	72.5%
Score 3	3	7.5%
Follow up		
DOD	3	7.5%
Progressive disease	34	85%
Lost to follow up	3	7.5%

Table 1. Details of clinical characteristics of SCLC patients. *DOD* died of disease, *ECOG* Eastern Co-operative Oncology Group-performance status scale, *ES-SCLC* extensive stage small cell lung cancer, *LS-SCLC* limited stage small cell lung cancer.

For sCTLA-4 estimation, serum samples of 34 SCLC patients were quantified using ELISA immunoassay (done in duplicates). The sCTLA-4 concentration of total of 34 samples was interpolated against a standard curve and concentrations were found to be normally distributed. The mean of sCTLA-4 was slightly higher in SCLC patients [mean (SD) = 0.199 (0.030) ng/mL; range = 0.142 ng/mL to 0.255 ng/mL] than healthy controls [mean (SD) = 0.182 (0.024) ng/mL] but not statistically significant (p-value = 0.08) (Fig. 3).

The comparison of the area under the ROC curve revealed no significant difference between the discrimination potential of these bio-markers (p-value = 0.224). The respective areas of IFN-γ, IL-2 and TNF-α and sCTLA-4 under the ROC curve and their corresponding 95% confidence interval are presented in Fig. 4.

Quantification could not be carried out in 6 serum samples as they were technically unsuitable.

Assessment of IDO1 activity (Kyn/Trp ratio) in plasma by LC–MS. The hydrophilic compounds Kyn and Trp were separated using hydrophilic interaction liquid chromatography (HILIC). Charcoal stripped plasma was used as the matrix for the preparation of calibration standards and it was free of endogenous Kyn and Trp. The Kyn levels in SCLC patients (n = 29) were found to be higher compared with healthy controls (n = 40). SCLC patients had lower levels of Trp when compared with healthy controls, the difference was also found to be statistically significant ($p < 0.05$). The assessment of IDO1 activity using Kyn/Trp ratio revealed 1.7 times higher values for SCLC patients compared with healthy control (Fig. 5) in a significant manner ($p = 0.007$).

Discussion

In recent times, ICPIs have become the most widely sought after therapeutic regimen, employed alone or in conjunction with chemotherapy, owing to their effective use in many malignancies. Though multiple randomized clinical trials have demonstrated the efficacy of the combination of ICPIs like PD1/PD-L1 inhibitors and anti-CTLA4 monoclonal antibodies along with chemotherapy in the first-line treatment as well as maintenance therapy of ES-SCLC, their practical perquisite in SCLC still remains under speculation. The lack of biomarkers predicting response to therapy compounded by immune related adverse effects owing to unjustified usage of ICPIs acts as hindrances. In this study, we tried to analyse the expression of different biomarkers in SCLC so as to evaluate whether it is possible to dissect a sub-group among the SCLC cohort who can be predicted to be potential responders to ICPIs.

Kyn-pathway controlled IDO1 has been associated with poor prognosis in various cancers. Increased IDO1 levels, reflected by decreased Trp and elevated Kyn concentrations in the peripheral blood has been observed to be related to tumor progression and poor clinical outcome. IDO1 induction in malignant cells and their micro-environment has also been hypothesised as a key mechanism to modulate anti-tumor immune response, helping

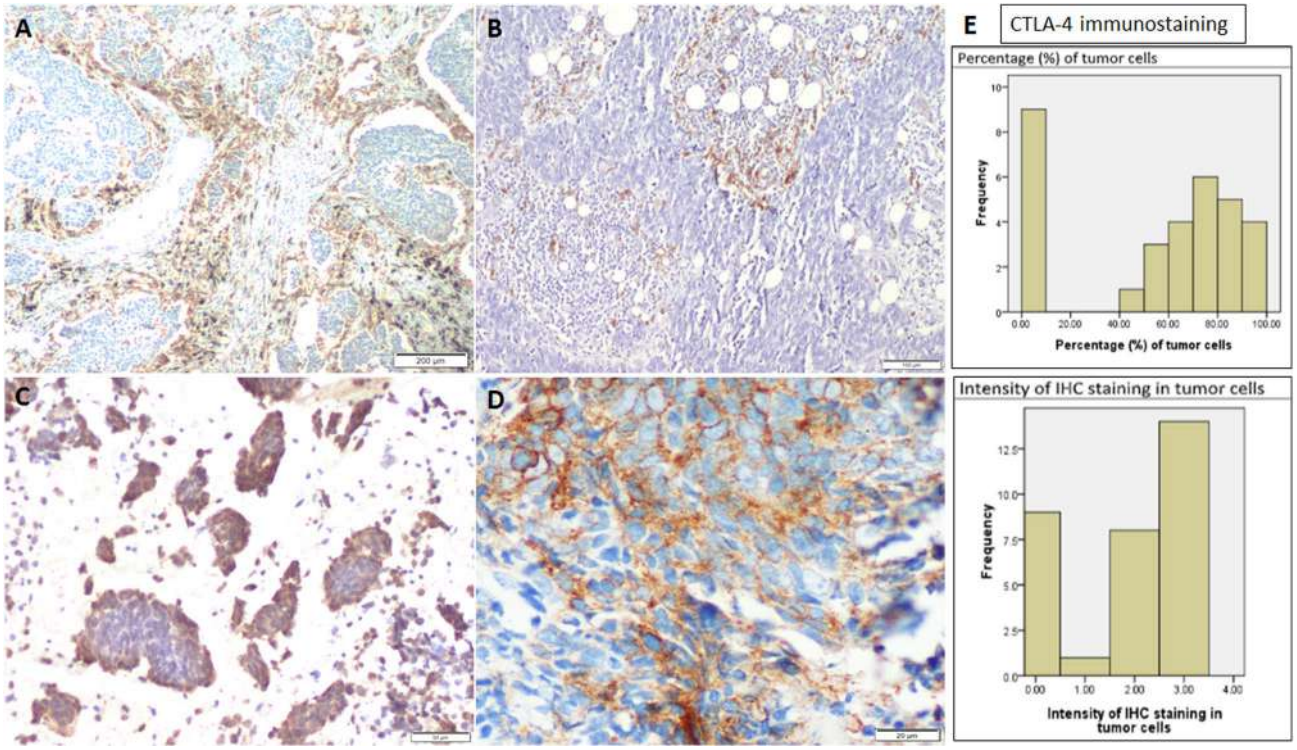

Figure 2. Immunohistochemical analysis of PD-L1, CTLA-4 and IDO1 in SCLC biopsies. (**A,B**) Immunostaining for IDO1 is seen in endothelial cells, intra-tumoral immune cells and stromal cells. (**C**) Immunostaining for CTLA-4 is seen in cytoplasm of tumor cells. (**D**) Strong membranous immunopositivity for PD-L1 is seen only in the tumor cells (TPS > 50%). (**E**) Histogram showing the intensity and percentage distribution of CTLA-4 immunoexpression in tumor cells amongst the SCLC cohort.

in escaping immune surveillance. Decreased Trp availability and accumulation of its metabolites, such as kyn also directly affects immune T-cell proliferation and functions, and induce their apoptosis. Based on these mechanisms, it can be speculated that elevated IDO1 levels can serve as a potential biomarker, highlighting decreased immune cell surveillance and tumor progression. IDO1 activity has been studied in NSCLC where it has been seen to be directly proportional to tumor stage and aggressiveness. Studies in human lung cancer cell lines demonstrated that though IDO mRNA can be constitutively expressed by lung cancer cells, higher IDO1 expression observed in patient samples could be attributed to production of the enzyme by other cells recruited in the tumor microenvironment and the peri-tumoral area, thereby mediating its immune conditioning[17]. Studies trying to correlate IDO1 levels with response to ICPI regimens have found that in certain tumors like stage IV melanoma and renal cell cancer treated with anti-PD-1 therapy, increase in Kyn/Trp during therapy compared to baseline was seen to be associated with significantly reduced progression-free survival[18]. IDO1 immuno-expression has been observed not only in tumor cells but in the surrounding stroma, as well. Thereby IDO1 immunolabelling in tumors has been seen to show three distinct cellular expression patterns which includes expression in tumor cells, interstitial cells in lymphocyte-rich areas in the tumor stroma or endothelial cells[19]. These patterns can be seen individually or in combination. In our study, the levels of kyn/try were assessed using hydrophilic interaction chromatography (HILIC), which offers better separation of hydrophilic analytes[20]. Depleted tryptophan levels and increased levels of kynurenine, indicating raised IDO1 activity were seen in our SCLC cohort as compared to healthy controls and the values were also statistically significant. However, on evaluating IDO1 immunohistochemical expression, we could elicit immunopositivity in only two cases. Nevertheless, a plausible explanation for this contrast in expression, could be due to defective expression of protein at the tissue level.

IDO1 expression is interconnected with other immune checkpoint molecules like CTLA-4. In the local tumor microenvironment, CTLA-4 expression has been seen to upregulate IDO1, which reciprocally promotes T-reg activation. This interplay can also be evinced in peripheral blood, where IDO1 expression has been seen to be associated with increased CTLA-4 + Tregs.

Soluble CTLA-4 (sCTLA-4), produced by alternatively spliced messenger RNA (mRNA) has functions which are contrary to the classical immunosuppressive effect of CTLA-4. Studies have demonstrated that sCTLA-4 functions in down-regulating the action of CTLA-4 in T-cell responses as sCTLA-4 interferes with CD80/CD86:CTLA-4 interactions, thereby blocking the negative signal imparted via CTLA-4 in the later phases of T-cell response. sCTLA-4 in pleural effusions of mesothelioma patients has even been shown to behave as a statistically significant positive prognostic factor[10]. Though sCTLA-4 concentration was higher than controls in our SCLC patients, it was not statistically significant. On the other hand, 71.8% of our cases showed CTLA-4

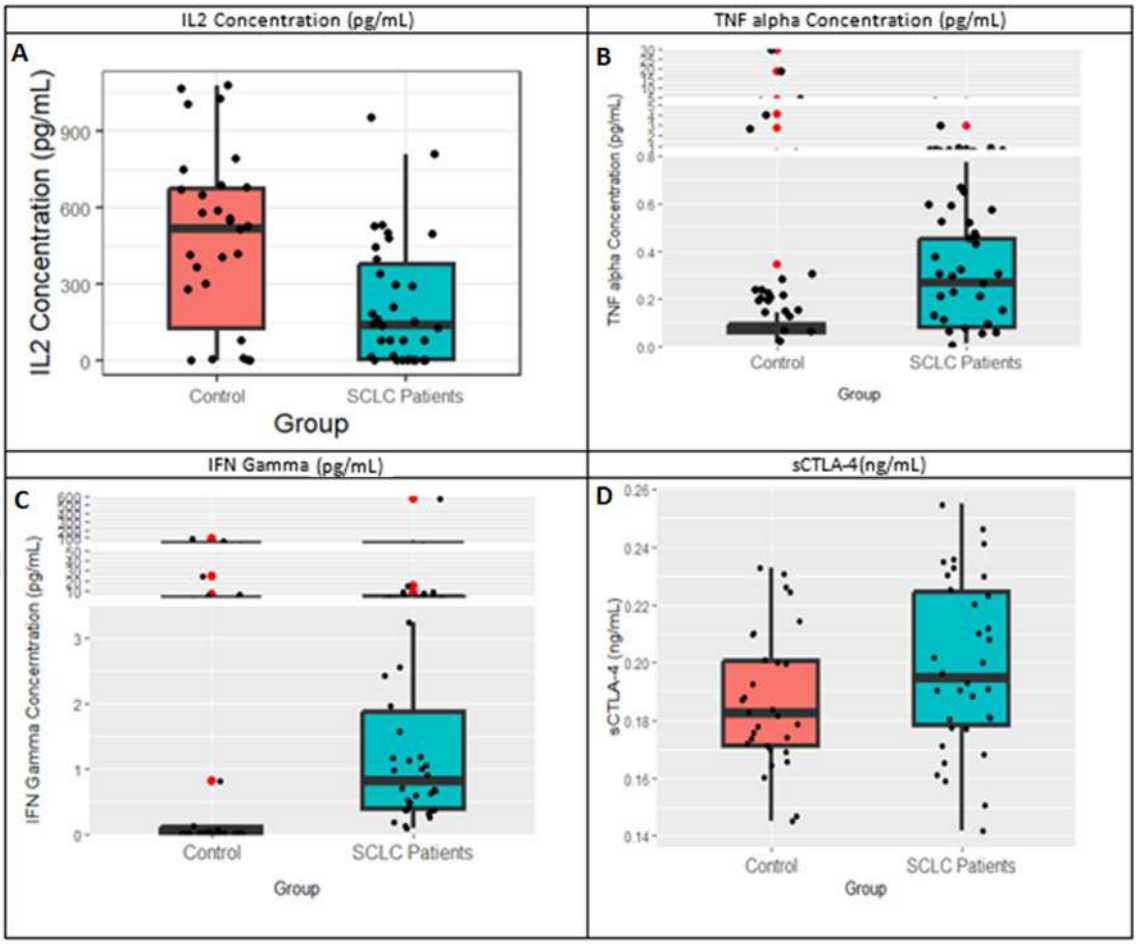

Figure 3. Quantitative estimation of serum cytokine levels and sCTLA-4. Box-whisker plots showing the distribution of serum levels of (**A**) IL-2, (**B**) TNF-α, (**C**) IFN-γ and (**D**) sCTLA-4 in SCLC patients and healthy controls.

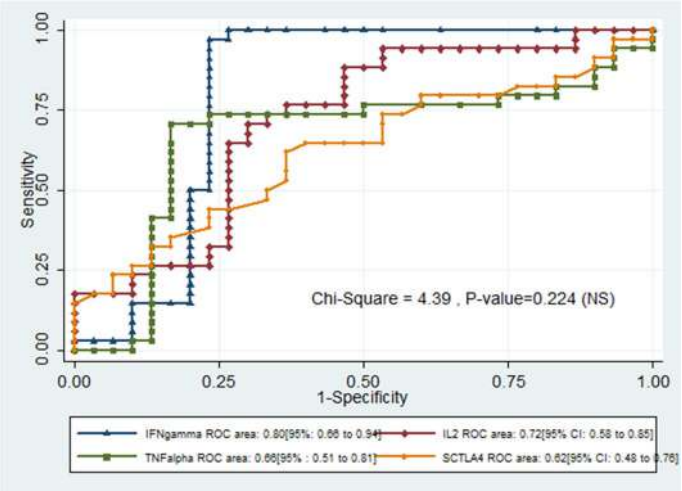

Figure 4. ROC curve comparing the discrimination potential amongst the serum biomarkers.

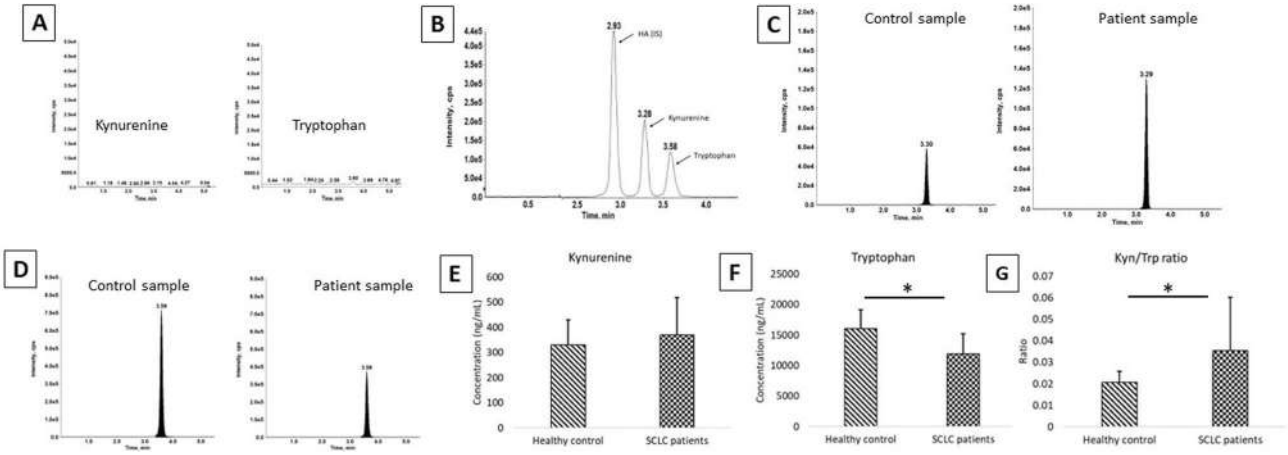

Figure 5. Quantification of kynurenine and tryptophan in healthy control and SCLC patients using LC–MS/MS. (**A**) Stripped plasma blank showing no peaks for endogenous Kyn and Trp. (**B**) Chromatogram showing separation of Kyn, Trp and homatropine (HA) which is used as an internal standard (IS). (**C**) Representative plasma concentration of Kyn and (**D**) Trp in control and patient sample, respectively. (**E**) Kyn levels in the plasma samples of healthy control and SCLC patients. (**F**) Trp levels in the plasma samples of healthy control and SCLC patients. (**G**) Kyn/Trp ratio of healthy control and SCLC patients.

immunohistochemical staining along with significantly raised IDO1 activity. These findings highlighted probable similar interactions between CTLA-4, s-CTLA-4 and IDO-1 as described in studies above.

Interferons are potent IDO1 inducers and part of the IDO1 expression by tumor cells might be attributed to IFN-γ. As tumor-infiltrating lymphocytes (TILs) are a predominant source of IFN-γ, they might upregulate IDO1, thereby potentially contributing to tumor immune escape[12]. IDO1 induction can also be potentiated by other pro-inflammatory cytokines, such as TNF-α. Though studies in NSCLC have highlighted that TNF-α signalling is involved in lung tumor progression and response to immunotherapy, the exact mechanisms modulating it remains to be unearthed[15]. The negative role played by TNF-α secreted by tumor associated macrophages, in promoting cell glycolysis, tumor hypoxia and decreased PD-L1 expression has been highlighted in few studies; while others have observed increased serological levels of TNF-α to be associated with improved response to anti-PD-1 treatment and survival[15]. These observations make the role of TNF-α seem rather cryptic and like a double edged sword. Tryptophan-catabolizing enzyme IDO1 also results in infiltration by FoxP3 + Tregs, thereby mediating suppression of effector T-cell function and causing T cell-intrinsic anergy. This is characterized by production and proliferation of defective IL-2[14]. Fischer et al. in their study on SCLC patients found that IL-2 levels served as independent prognostic factors and high serum levels at baseline directly correlated with long term survival[13]. In our study, IL-2 levels were found to be significantly lower as compared to controls ($p = 0.003$). In addition, serum IFN-γ (p-value < 0.001) and TNF-α levels (p-value = 0.025) were significantly higher in the SCLC cohort compared to healthy controls; suggesting marked immune suppression.

IDO1 induced T-regs are seen to reciprocally upregulate PD-L1 and PD-L2 expression on target dendritic cells[17]. Studies analysing PD-L1 positivity in neuroendocrine carcinomas including SCLC have cited variable percentages ranging from 5.8% to as high as 71.6%[5]. However, this conflicting results could be attributed to use of unconventional clones of PD-L1 for analysing IHC expression and different methodologies adopted for calculation of scores, which includes counting immune-expression in immune cells in addition to tumor. Studies using FDA approved clones and cut-offs for estimating PD-L1 positivity have usually found low rates of PD-L1 immunopositivity in SCLC[5]. In our study, we could elicit PD-L1 immunopositivity in tumor cells in only three cases with TPS scores of 1%, 10% and > 50%. However, we noticed immunopositivity in the immune cells in the tumor stroma in approximately one-fourth (25%) of our cases. In the absence of well-defined criteria to assess PD-L1 status on immune cells alone, using FDA approved clones; whether to interpret this finding as a hint towards immune privilege nature remains a bone of contention.

To summarise, SCLC patients show immune suppressive milieu in their peripheral circulation and can be potentially benefitted by appropriate immuno-modulatory regimen. Owing to various clinical trials advocating the use of anti-PD-L1 drugs in recurrent or ES-SCLC quoting significant difference in survival rates, the need for various predictive biomarkers arises, which might help in decision making regarding sustaining or quitting ICPI regimens. Theragnostic biomarkers like IDO1, s-CTLA-4, IFN-γ, TNF-α and IL-2 have been studied in various tumors, and seems to hold significant potential in mapping the tumor microenvironment and predicting response to ICPIs. However apart from IL-2, till date, there has been hardly any studies exploring the potential of using these biomarkers in SCLC. The findings in our study suggest that these biomarkers are expressed in SCLCs also and might act in a similar manner in predicting prognosis and response to different therapeutic regimens as has been established in other tumors, including NSCLC. However, our study cohort was small in size and more studies incorporating larger groups are warranted to validate whether our findings can be extrapolated to SCLCs per se. Another limitation in our study was that owing to aggressive nature of SCLC and difficulty in following up patients, we could not evaluate the fluctuation of these biomarkers in relation to response to

different therapeutic regimens or carry out survival analyses. Howbeit, we proffer that evaluation of these biomarkers particularly IDO1 at baseline appears cogent not only as a prognostic marker but also hold potential as a therapeutic target. Based on IHC expression of immune co-inhibitory molecules in our cohort, tumor cells of SCLC were found to be poor immuno-expresser of PD-L1 and IDO1, while CTLA-4 immunolabelling was seen in a fair number of cases. In the absence of PD-L1, CTLA-4 immuno-expression in tumor cells and related adjuncts like s-CTLA-4 levels in serum may be studied so as to analyse whether they conform as biomarker for predicting responsiveness to ICPIs.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 25 September 2022; Accepted: 2 March 2023
Published online: 06 March 2023

References

1. Van Meerbeeck, J. P., Fennell, D. A. & De Ruysscher, D. K. Small-cell lung cancer. *Lancet* **378**(9804), 1741–1755 (2011).
2. Esposito, G. *et al.* Immunotherapy in small cell lung cancer. *Cancers* **12**(9), 2522 (2020).
3. Reck, M., Heigener, D. & Reinmuth, N. Immunotherapy for small-cell lung cancer: Emerging evidence. *Future Oncol.* **12**(7), 931–943 (2016).
4. Halmos, B. *et al.* A matching-adjusted indirect comparison of pembrolizumab+ chemotherapy vs nivolumab+ ipilimumab as first-line therapies in patients with PD-L1 TPS≥ 1% metastatic NSCLC. *Cancers* **12**(12), 3648 (2020).
5. Guleria, P., Kumar, S., Malik, P. S. & Jain, D. PD-L1 expression in small cell and large cell neuroendocrine carcinomas of lung: An immunohistochemical study with review of literature. *Pathol. Oncol. Res.* **26**(4), 2363–2370 (2020).
6. Rudin, C. M. *et al.* Molecular subtypes of small cell lung cancer: A synthesis of human and mouse model data. *Nat. Rev. Cancer* **19**(5), 289–297 (2019).
7. Dora, D. *et al.* Neuroendocrine subtypes of small cell lung cancer differ in terms of immune microenvironment and checkpoint molecule distribution. *Mol. Oncol.* **14**(9), 1947–1965 (2020).
8. Ettinger, D. S. *et al.* Non-small cell lung cancer, version 1.2020: Featured updates to the NCCN guidelines. *J. Natl. Comprehens. Cancer Netw.* **17**(12), 1464–1472 (2019).
9. Mandarano, M. *et al.* Kynurenine/tryptophan ratio as a potential blood-based biomarker in non-small cell lung cancer. *Int. J. Mol. Sci.* **9**, 4403 (2021).
10. Liu, Q. *et al.* Soluble cytotoxic T-lymphocyte antigen 4: A favorable predictor in malignant tumors after therapy. *OncoTargets Ther.* **10**, 2147 (2017).
11. Nakamura, Y. Biomarkers for immune checkpoint inhibitor-mediated tumor response and adverse events. *Front. Med.* **6**, 119 (2019).
12. Kanai, T. *et al.* Significance of quantitative interferon-gamma levels in non-small-cell lung cancer patients' response to immune checkpoint inhibitors. *Anticancer Res.* **40**(5), 2787–2793 (2020).
13. Fischer, J. R. *et al.* Long-term survival in small cell lung cancer patients is correlated with high interleukin-2 secretion at diagnosis. *J. Cancer Res. Clin. Oncol.* **126**(12), 730–733 (2000).
14. Monkman, J., Kim, H., Mayer, A., Mehdi, A., Matigian, N., Cumberbatch, M., Bhagat, M., Ladwa, R., Mueller, S.N., Adams, M.N., & O'Byrne, K. IL-2 stromal signatures dissect immunotherapy response groups in non-small cell lung cancer (NSCLC). *medRxiv* (2021).
15. Benoot, T., Piccioni, E., De Ridder, K. & Goyvaerts, C. TNFα and immune checkpoint inhibition: Friend or foe for lung cancer?. *Int. J. Mol. Sci.* **22**(16), 8691 (2021).
16. Chen, R. *et al.* A more rapid, sensitive, and specific HPLC-MS/MS method for nifedipine analysis in human plasma and application to a pharmacokinetic study. *Drug Res.* **63**(01), 38–45 (2013).
17. Karanikas, V. *et al.* Indoleamine 2, 3-dioxygenase (IDO) expression in lung cancer. *Cancer Biol. Ther.* **6**(8), 1269–1278 (2007).
18. Li, H. *et al.* Metabolomic adaptations and correlates of survival to immune checkpoint blockade. *Nat. Commun.* **10**(1), 1–6 (2019).
19. Meireson, A., Devos, M. & Brochez, L. IDO expression in cancer: Different compartment, different functionality?. *Front. Immunol.* **11**, 531491 (2020).
20. Sharma, H. P., Halder, N., Singh, S. B. & Velpandian, T. Involvement of nucleoside transporters in the transcorneal permeation of topically instilled substrates in rabbits in-vivo. *Eur. J. Pharm. Sci.* **114**, 364–371 (2018).

Acknowledgements

The authors are thankful to Mr. Kaustar Yadav; Thoracic Pathology laboratory, AIIMS, New Delhi for his technical contribution.

Author contributions

R.D.: Formal analysis, methodology, data curation, writing-original draft, and writing-review and editing; A.R.: conducting the experiments and data analysis; H.P.S.: conducting the experiments, data analysis and review; H.C.P.: arranging of donor blood samples, P.S.M.: data curation, writing-review and editing; A.M.: data curation, writing-review and editing, A.N.: data curation, writing-review and editing, R.K.: statistical analysis, D.J.: conceptualization, data curation, project administration, writing-review and editing, visualisation and supervision.

Funding

The work is supported by Science and Engineering Research Board (SERB) core research grant number DST 552.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to D.J.

ORIGINAL ARTICLE

NSCLC Subtyping in Conventional Cytology: Results of the International Association for the Study of Lung Cancer Cytology Working Group Survey to Determine Specific Cytomorphologic Criteria for Adenocarcinoma and Squamous Cell Carcinoma

Deepali Jain, MD, FIAC,[a,*] Aruna Nambirajan, MD,[a] Gang Chen, MD,[b] Kim Geisinger, MD,[c] Kenzo Hiroshima, MD, PhD,[d] Lester Layfield, MD,[e] Yuko Minami, MD, PhD,[f] Andre L. Moreira, MD,[g] Noriko Motoi, MD, PhD,[h,i] Mauro Papotti, MD,[j] Natasha Rekhtman, MD, PhD,[k] Prudence A. Russell, FRCPA,[l] Spasenija Savic Prince, MD,[m] Fernando Schmitt, MD, PhD, FIAC,[n] Yasushi Yatabe, MD, PhD,[h] Serenella Eppenberger-Castori, PhD,[m] Lukas Bubendorf, MD,[m] for the IASLC Pathology Committee

[a]*Department of Pathology, All India Institute of Medical Sciences, New Delhi, India*
[b]*Department of Pathology, Zhongshan Hospital, Fudan University, Shanghai, People's Republic of China*

*Corresponding author.

Disclosure: Dr. Layfield reports receiving personal fees from McBride Hall Law Firm Johnson and Kramer Mulholland Law Firm and grants from the University of Missouri, outside of the submitted work; and having a fiduciary role in Papanicolaou Society of Cytopathology. Dr. Motoi reports receiving personal fees from Becton Dickinson Japan, outside of the submitted work. Dr. Papotti reports receiving personal fees from Eli Lilly, Roche, and Pfizer, outside of the submitted work. Dr. Russell reports receiving personal fees from Amgen, outside of the submitted work; and serving on the advisory board for Amgen. Dr. Schmitt reports serving as General Secretary of the International Academy of Cytology. Dr. Yatabe reports receiving grants from ArcherDx, Chugai Pharma, and Thermo Fisher Science; personal fees from AbbVie Inc., Amgen, Bayer, Ono Pharma, Daiichi Sankyo, Eli Lilly, Merck Bio-Pharma, Pfizer, Merck Sharp & Dohme, Novartis, AstraZeneca, Agilent, ArcherDx, Sysmex, Chugai Pharma, Boehringer Ingelheim, Yansen Pharma, Roche/Ventana, Thermo Fisher Science, and Takeda, outside of the submitted work. Dr. Bubendorf reports receiving personal fees from AstraZeneca, Eli Lilly, Bayer, AbbVie, Takeda, Janssen, Bristol Myers Squibb, Thermo Fisher, and Merck Sharp & Dohme; and other support from Roche, Novartis, Thermo Fisher, Systems Oncology, GlaxoSmithKline, and Novartis, outside of the submitted work. The remaining authors declare no conflict of interest.

Other members of the Working Group, IASLC Pathology Committee: Mary Beth Beasley,[1] Sabina Berezowska,[2] Alain Borczuk,[3] Elizabeth Brambilla,[4] Teh-Ying Chou,[5] Jin-Haeng Chung,[6] Wendy Cooper,[7] Sanja Dacic,[8] Yuchen Chan,[9] Fred R. Hirsch,[10] David Hwang,[11] Philippe Joubert,[12] Keith Kerr,[13] Sylvie Lantuejoul,[14] Dongmei Lin,[15] Fernando Lopez-Rios,[16] Daisuke Matsubara,[17] Mari Mino-Kenudson,[18] Andrew Nicholson,[19] Claudia Poleri,[20] Anja Roden,[21] Kurt Schalper,[22] Lynette Sholl,[23] Erik Thunnissen,[24] William D. Travis,[25] Ming Tsao,[26] Ignacio Wistuba,[27].

[1]Department of Pathology, Mount Sinai Medical Center, New York, New York
[2]Department of Pathology, University of Bern, Bern, Switzerland
[3]Department of Pathology, Weill Cornell Medicine, New York, New York
[4]Department of Pathology, CHU Albert Michallon Hospital, Grenoble, France
[5]Department of Pathology, Taipei Veterans General Hospital, Taipei, Taiwan
[6]Department of Pathology, Seoul National University, Bundang Hospital, Seoul, South Korea
[7]Department of Pathology, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia
[8]Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania
[9]Department of Pathology, Shanghai Chest Hospital, Shanghai, People's Republic of China
[10]Center for Thoracic Oncology, Mt. Sinai Cancer Center, New York, New York
[11]Department of Pathology, Sunnybrook Health Sciences Centre, Odette Cancer Centre, Toronto, Ontario, Canada
[12]Department of Pathology, Universite Laval, Laval Hospital, Laval, Quebec, Canada
[13]Department of Pathology, School of Medicine and Dentistry, University of Aberdeen, Aberdeen, Scotland, United Kingdom
[14]Department of Pathology, Centre Leon Berard and Grenoble Alpes University, Grenoble, France
[15]Department of Pathology, Beijing Cancer Center, Beijing, People's Republic of China
[16]Department of Pathology, Hospital Universitario HM Sanchinarro, Madrid, Spain
[17]Department of Pathology, Jichi Medical University, Yakushiji, Shimotsuke-shi, Tochigi-ken, Japan
[18]Department of Pathology, Massachusetts General Hospital, Boston, Massachusetts
[19]Department of Pathology, Royal Brompton and Harefeild NHS Foundation Trust, London, United Kingdom
[20]Pathology Consultants, Buenos Aires, Argentina
[21]Department of Pathology, Mayo Clinic, Rochester, Minnesota
[22]Department of Pathology, Yale Cancer Center, New Haven, Connecticut
[23]Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts
[24]Department of Pathology, VU University Medical Center, Amsterdam, The Netherlands
[25]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York
[26]Department of Pathology, Princess Margaret Hospital, Toronto, Ontario, Canada
[27]Department of Pathology, MD Anderson Cancer Center, Houston, Texas

Address for correspondence: Deepali Jain, MD, FIAC, Department of Pathology, All India Institute of Medical Sciences, New Delhi, India. E-mail: deepalijain76@gmail.com

ISSN: 1556-0864

https://doi.org/10.1016/j.jtho.2022.02.013

[c]Joint Pathology Center, Silver Spring, Maryland
[d]Department of Pathology, Tokyo Women's Medical University Yachiyo Medical Center, Yachiyo, Japan
[e]Department of Pathology and Anatomical Sciences, University of Missouri, Columbia, Missouri
[f]Department of Pathology, National Hospital Organization, Ibaraki Higashi National Hospital, Ibaraki, Japan
[g]Department of Pathology, New York University Langone Health, New York, New York
[h]Department of Diagnostic Pathology, National Cancer Center Hospital, Tokyo, Japan
[i]Department of Pathology, Saitama Cancer Center, Saitama, Japan
[j]Department of Oncology, University of Turin, Turin, Italy
[k]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York
[l]Anatomical Pathology Department, St Vincent's Hospital and the University of Melbourne, Fitzroy, Australia
[m]Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland
[n]Medical Faculty of Porto University/RISE@Cintesis, Porto, Portugal

Received 7 November 2021; revised 3 February 2022; accepted 15 February 2022
Available online - 22 March 2022

ABSTRACT

Introduction: Accurate subtyping of NSCLC into lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) is the cornerstone of NSCLC diagnosis. Cytology samples reveal higher rates of classification failures, that is, subtyping as non–small cell carcinoma—not otherwise specified (NSCC-NOS), as compared with histology specimens. This study aims to identify specific algorithms on the basis of known cytomorphologic features that aid accurate and successful subtyping of NSCLC on cytology.

Methods: A total of 13 expert cytopathologists participated anonymously in an online survey to subtype 119 NSCLC cytology cases (gold standard diagnoses being LUAD in 80 and LUSC in 39) enriched for nonkeratinizing LUSC. They selected from 23 predefined cytomorphologic features that they used in subtyping. Data were analyzed using machine learning algorithms on the basis of random forest method and regression trees.

Results: From 1474 responses recorded, concordant cytology typing was achieved in 53.7% (792 of 1474) responses. NSCC-NOS rates on cytology were similar among gold standard LUAD (36%) and LUSC (38%) cases. Misclassification rates were higher in gold standard LUSC (17.6%) than gold standard LUAD (5.5%; $p < 0.0001$). Keratinization, when present, recognized LUSC with high accuracy. In its absence, the machine learning algorithms developed on the basis of experts' choices were unable to reduce cytology NSCC-NOS rates without increasing misclassification rates.

Conclusions: Suboptimal recognition of LUSC in the absence of keratinization remains the major hurdle in improving cytology subtyping accuracy with such cases either failing classification (NSCC-NOS) or misclassifying as LUAD. NSCC-NOS seems to be an inevitable morphologic diagnosis emphasizing that ancillary immunochemistry is necessary to achieve accurate subtyping on cytology.

Keywords: Non–small cell lung carcinoma; Cytology; Subtyping; Machine learning; Regression tree; IASLC

Introduction

Because the paradigm shift in the treatment and management of NSCLC with molecular targeted therapies becoming the standard of care for advanced lung adenocarcinomas (LUADs),[1] accurate subclassification of NSCLC into squamous cell carcinomas and adenocarcinomas has gained paramount clinical importance. The 2021 WHO classification of lung tumors recommends morphologic subtyping of NSCLC into LUAD and lung squamous cell carcinoma (LUSC) in cytology and small biopsy specimens in the presence of clear-cut glandular and squamous differentiation, respectively; those remaining equivocal on morphology be assessed with additional immunochemistry (IC) with or without mucin stains.[2] Accordingly, these are categorized as non–small cell carcinoma (NSCC)-favor LUAD in the presence of any TTF-1 positivity when using recommended antibody clones (such as 8G7G3/1) regardless of expression of p40, NSCC-favor LUSC when p40 is expressed in the absence of TTF-1 expression, and NSCC—not otherwise specified (NOS) when both TTF-1 and p40 are negative, equivocal, or are not performed owing to technical limitations.[3] In addition to unequivocal LUAD and NSCC-favor LUAD, all NSCC-NOS samples presenting with inoperable disease should be tested for targetable molecular alterations in *EGFR*, *ALK*, *ROS1*, *BRAF*, *MET exon 14*, *RET*, and *KRAS* genes characteristic of LUAD and for programmed death-ligand 1.[1,4]

Conventional or liquid-based cytology offers certain advantages over core biopsies and cell block preparations in the morphologic subtyping of NSCLC: absence of formalin-induced fixation artifacts with better appreciation of cytoplasmic and nuclear details; ability to identify even minimal keratinization as cytoplasmic orangophilia on Papanicolaou stain; lack of obscuring crush artifacts most often found in tissue sections; and presence of distinct architectural clues, such as three-dimensional (3D) clusters and polarized tumor cells with a luminal edge usually not appreciated on histology.[5,6] Multiple studies in the past decade have consistently found high accuracy of NSCLC subtyping as LUSC and LUAD on

pulmonary cytology[5,7–15] with comparable results with corresponding small biopsies[5,11]and with lower need for IC.[5] Misclassifications and failure to classify (i.e., NSCC-NOS) on cytology in these studies are mainly found in exfoliative cytology,[8] biopsy touch preparations,[13] Giemsa-stained smears particularly for LUSC,[17] samples with low cellularity,[5,8] or necrosis,[5,8,9] nonkeratinizing LUSC,[5,8–11] and in rare instances of large cell carcinoma, sarcomatoid carcinoma, or adenosquamous carcinomas which by definition cannot be subtyped on cytology.[8–10]

Despite good accuracy rates for subtyping found in these studies, real-time reporting audits reveal that up to 36% of NSCLC pulmonary cytology samples may not be subtyped resulting in higher rates of NSCC-NOS in cytology.[16] Recent years have witnessed decreasing rates of NSCC-NOS in cytology in part owing to increasing awareness among cytopathologists on the need to subtype,[11,17] and more importantly, use of ancillary IC.[8,9,14,16,18,19] Although IC improves subtyping accuracy considerably, IC on conventional non–cell block cytology preparations is not routine practice in many laboratories leading to higher rates of NSCC-NOS diagnosis in cytology cases, in which a cell block was not made or was inadequate. Such scenarios are not very uncommon, and they also compromise valuable material that may be better triaged for molecular analysis. Interestingly, some authors have observed that morphologic reassessment itself can result in a considerable reduction in NSCC-NOS rate in cytology with high accuracy.[11] Although some studies do mention the specific cytology features that were observed in LUSC and LUAD,[7,9,10] it is not clear how useful these individual features are when applied during real-time reporting. Despite the large number of cytology studies on subtyping, none has systematically evaluated objective cytomorphologic criteria for NSCLC subtyping.

Machine learning (ML) algorithms are important computer-based systems that have been found to be useful to describe and predict biological and medical systems. Their use has increased in the medical field and starts to be of assistance in medical decision-making. There are several types of ML calculations, such as linear model, neuronal network, decision tree, and random forest. A decision regression tree is a model that predicts the output class by answering a series of questions on the input data. The main advantage of regression trees is that they are easy to visualize graphically and be interpreted scientifically. Nevertheless, a small modification in the input data may result in a different algorithm; therefore, in this work, we also apply the random forest method, which builds a classification algorithm on the basis of a combination of regression trees. The last is much more stable, but not possible to be easily depicted.[20]

With the aim of evaluating means of improving NSCLC subtyping on conventional cytology using cytomorphology alone, we hypothesized that ML may more clearly define an algorithm on the basis of the known cytologic features segregating LUAD and LUSC that could be easily explained and taught. To meet this end, a survey was conducted among expert cytopathologists to subtype NSCLC cases on cytology. ML was applied to identify individual cytomorphologic features associated with accurate and successful subtyping.

Materials and Methods

Cytology cases of NSCLC with four to five representative high-quality microphotographs were collected from members of the IASLC cytology working group. For the purpose of uniformity, only non–cell block, ethanol-fixed and Papanicolaou or hematoxylin and eosin–stained preparations were used for this study. The gold standard diagnosis was determined by histology with the aid of IC on concurrent or subsequent biopsy or resection specimens whenever needed or by cytomorphology with unequivocal IC results (p40/TTF-1) on the basis of the WHO classification of lung tumors.[2] The LUSC cases were enriched for the nonkeratinizing type. The images were randomly displayed in an online gallery using the publicly available “Google forms” survey administration software. The participants, who were blinded to the final diagnosis, attempted to assign the cases into the following three categories: LUAD, LUSC, and NSCC-NOS, immunostaining (or mucin stains) required, according to their diagnostic practice.

A list of 23 predefined morphologic parameters was displayed below each case (Table 1), and participants were required to choose any number of morphologic features they considered adequate to describe the particular case. The experts first used their own semi-objective criteria of daily practice to assign the cytologic subtype and were then asked to mark these features independent of their diagnosis in an objective manner. The results were compiled in an anonymized fashion.

Statistical Analysis

Data extracted from the Google forms gallery platform contained the subtype chosen by each of the experts per case, a list of predefined morphologic criteria used by each expert for the evaluation of each slide, and the gold standard diagnosis of each image. Gold standard LUAD subtyped as LUAD on cytology and gold standard LUSC subtyped as LUSC on cytology were classified as “concordant.” Gold standard LUAD subtyped as LUSC and gold standard LUSC subtyped as LUAD were classified as “misclassified.” Gold standard LUAD or LUSC subtyped as NSCC-NOS on cytology was noted as “failure to classify.” We restructured the data by transposing the experts’ evaluations and creating a column for each morphologic

Table 1. Predefined Morphologic Criteria That Were Displayed to the Participants Asked to Choose the Following Subtypes: LUAD, LUSC, NSCC-NOS, Immunostains (or Mucin Stain) Required

LUAD Cytologic Features	LUSC Cytologic Features
Intracellular mucin	Keratinization
Extracellular mucin	Intercellular bridges
Acinus formation	Dense (opaque) cytoplasm
Rounded 3D cell clusters	Tadpole or spindle cell shape in the absence of keratinization
Papillary formations with fibrovascular cores	Cell cannibalism
Cylindrical cell shape	Hyperchromatic nuclei
Translucent (delicate) cytoplasm	Heterochromatic nuclei
Vacuolated cytoplasm	Pyknotic chromatin in the absence of keratinization
Eccentric nuclei	Irregular nuclear membranes
Vesicular nuclei with open chromatin	Necrotic debris
Prominent nucleoli	
Nuclear grooves	
Nuclear pseudoinclusions	

3D, three-dimensional; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; NSCC-NOS, non-small cell carcinoma—not otherwise specified.

characteristic to create a matrix. Each characteristic was indicated as "used" or "not used" by the single expert for the evaluation of each case. Random forest method from ML (package random Forest version 4.6-14)[21] and regression tree on the basis of the "rpart package"[22] were applied for data analysis. Both methods require splitting the overall data into training (approximately 60% of all data randomly selected), validation (approximately 20% of the data), and test (the remaining 20% of the data) subsets to carefully construct the algorithm, validate and optimize it on the validation subset, and finally test for its accuracy on the test subset of data. Data preparation and analysis were performed using the Statistical Package Software R (version 4.0.2, (2020-06-22), www.r-project.org/).

Results

Case Selection and Participant Information

A total of 119 cases were included in the survey to be completed without a time limit. The cases consisted of ethanol-fixed, non–cell block cytology specimens, including smears and cytospins, from bronchial secretions, bronchial brushes, bronchoalveolar lavages, or endobronchial ultrasound–guided transbronchial fine-needle aspirates. Most had been stained according to Papanicolaou and a small number ($n = 10$) with hematoxylin and eosin.

The gold standard diagnoses in the 119 cases were 80 LUAD (67%) and 39 LUSC (33%). The gold standard diagnosis was achieved on histomorphology in corresponding histologic specimens in 32 cases, histologic specimens with additional IC in 73 cases, and on cytology with IC in 14 cases (Table 2).

A total of 13 participants took part in the survey. One participant could not complete 73 cases owing to impaired internet connection. Overall, 1474 responses/answers were recorded (Fig. 1*A*).

Categorization of Cytology Subtyping Responses

As depicted in Figure 1*B* and Table 3, concordant subtyping as LUAD and LUSC was achieved in 792 responses (792 of 1474, 53.7% overall concordance) with concordant LUAD subtyping achieved in 58.5% of the responses (580 of 992) and concordant LUSC subtyping in 44.0% of the responses (212 of 482).

Overall misclassification rate on cytology was 9.5% (140 of 1474), with gold standard LUSC being more frequently misclassified (85 of 482, 17.6% misclassifications) as compared with gold standard LUAD (55 of 992, 5.5% misclassifications, $p < 0.0001$).

Failure of classification in cytology was found in 36.8% of overall responses (542 of 1474) with gold standard LUAD or LUSC subtyped as NSCC-NOS on cytology. Classification failure rates were comparable between gold standard LUAD (36%) and LUSC (38.4%).

Correlation of Individual Cytologic Features With Subtyping

On average, most participants marked three to six cytologic criteria (mean = 3.9, range: 1–10 of number of criteria used), and there was no association between accuracy of the diagnosis and the total number of criteria used by individual participants (Supplementary Fig. 1).

The most common cytologic features observed by experts in "concordant," "misclassified," and "failure to classify" responses are summarized in Table 3 and detailed in Supplementary Figures 2*A* and *B*, 3*A* and *B*, and 4*A* and *B*.

Although vesicular nuclei with open chromatin and prominent nucleoli were the most common features observed in accurately classified LUADs (Fig. 2*A*–*D*), the same features predominated in the misclassified LUSCs driving the wrong classification. Similarly, hyperchromatic nuclei and dense opaque cytoplasm were the

Table 2. Method of Determination of Gold Standard Diagnosis for the Selected 119 Cases

	Method of Gold Standard Diagnosis		
Gold Standard Diagnosis (Number)	Histology Alone (%)	Histology + p40/TTF-1 IC (%)	Cytology + p40/TTF-1 IC (%)
LUAD (80)	18 (23)	53 (66)	9 (11)
LUSC (39)	14 (36)	20 (51)	5 (13)
Total (119)	32 (27)	73 (61)	14 (12)

IC, immunohistochemistry; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.

most common features observed in correctly classified LUSCs (Fig. 3*A–D*) and in the misclassified LUADs. In gold standard LUSC, keratinization was mentioned in 160 of 482 responses and in 14 of the 39 cases by at least six experts.

A total of 10 cases (seven LUSC and three LUAD) failed to be classified or were misclassified by all experts, details of which are tabulated in Table 4 and further detailed in Supplementary Figure 5*A* and *B*. The 86 LUSC responses included nearly equal number of misclassifications (40%) and classification failures (60%), whereas most (81%) of the 37 LUAD responses were classification failures.

The misclassifications in these cases were driven by the same cytologic features as noted previously in overall misclassified responses, which are as follows: vesicular nucleus with open chromatin, rounded 3D clusters, prominent nucleoli, and translucent (delicate) cytoplasm being observed in gold standard LUSC, and hyperchromatic nuclei, dense opaque cytoplasm, and tadpole-shaped cells without keratinization observed in gold standard LUAD (Table 5, Fig. 4*A–D*, and Supplementary Fig. 5*B*). Keratinization was falsely assigned to 11 of the 80 LUAD cases by at least one expert and a maximum of six experts leading to 23 observations of keratinization in LUAD images and 15 misclassified responses (LUSC instead of LUAD; the eight remaining responses were "failed to classify"). Review of LUAD cases with observed keratinization revealed that hypereosinophilia by technical staining variability, eosinophilic intra- or extracellular mucin, admixed benign squamous cells, or non-specific hypereosinophilic structures had most likely been misinterpreted as keratinization (Fig. 4*E* and *F*).

Application of ML-Based Random Forest Method to Develop Predictive Algorithms

We created a matrix of the individual responses with 1474 rows (13 experts evaluated 46 images [32 LUAD and 14 LUSC] and 12 the remaining 73 images [48 LUAD and 25 LUSC]) and 23 columns corresponding to the predefined morphologic criteria found in Table 1. Furthermore, we created additional columns indicating whether the experts had performed a "concordant" diagnosis corresponding to the gold standard, a "misclassified" diagnosis in case of LUAD instead of LUSC and vice versa, or a "failure to classify" in cases of gold standard LUAD or LUSC typed as NSCC-NOS, immune stains (mucin stains) required.

Algorithm 1 Developed Based on "Concordant" Responses

First, to identify the most informative cytologic features for concordant subtyping, the rows with "concordant" subtype classification ($n = 792$) were selected and

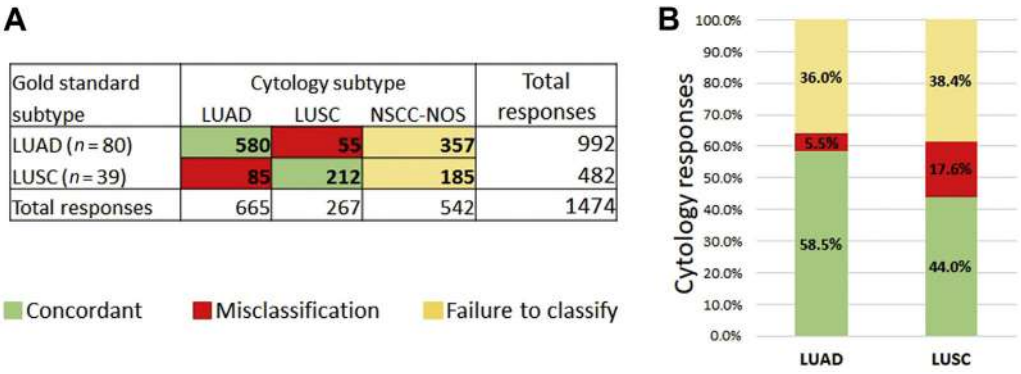

Figure 1. Comparison of cytology subtyping results with gold standard diagnosis. (*A*) Tabulation of the cytology subtyping responses. The numbers in bold-face represent the categorisation of responses on cytology. (*B*) Categorization of the cytology subtyping responses on gold standard LUAD and LUSC cases. LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; NSCC-NOS, non-small cell carcinoma—not otherwise specified.

Table 3. Common Cytologic Features Observed in "Concordant," "Misclassified," and "Failure to Classify" Responses in Gold Standard LUAD and LUSC

Categorization of Responses ($n = 1487$)	All	Gold Standard LUAD		Gold Standard LUSC	
		Category	Common cytologic features	Category	Common cytologic features
Concordant	53.7% (792/1474)	58.5% (580/992)	Vesicular nuclei with open chromatin (57.6%) Prominent nucleoli (55.7%) Delicate/translucent cytoplasm (50.0%)	44% (212/482)	Hyperchromatic nuclei (79.2%) Dense opaque cytoplasm (75.5%) Keratinization (70.3%)
Misclassification	9.5% (140/1474)	5.5% (55/992)	Dense (opaque) cytoplasm (87.3%) Hyperchromatic nuclei (70.9%) Tadpole or spindle cell morphology without keratinization (56.4%)	17.6% (85/482)	Vesicular nuclei with open chromatin (50.6%) Rounded 3D cell clusters (42.4%) Prominent nucleoli (38.8%) Acinus formation (38.8%)
Failure to classify	36.8% (542/1474)	36.0% (357/992)	Prominent nucleoli (46.5%) Translucent cytoplasm (31.4%) Hyperchromatic nuclei (30.5%)	38.4% (185/485)	Hyperchromatic nuclei (42.7%) Dense cytoplasm (40.0%) Prominent nuclei (34.1%)

3D, three-dimensional; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.

randomly subdivided into the following three subsets: training ($n = 474$, ~60%), test ($n = 167$, ~20%), and validation ($n = 151$, ~20%) data sets. By means of the ML random forest method on the training and test subsets, a predictor for LUAD or LUSC was developed. This ML random forest predictor used 28 trees and the following 12 cytomorphologic characteristics: keratinization, dense (opaque) cytoplasm, hyperchromatic nuclei, pyknotic chromatin in the absence of keratinization, tadpole or spindle cell shape in the absence of keratinization, necrotic debris, vesicular nuclei with open chromatin, prominent nucleoli, translucent (delicate) cytoplasm, vacuolated cytoplasm, eccentric nuclei, and acinus formation. In the training

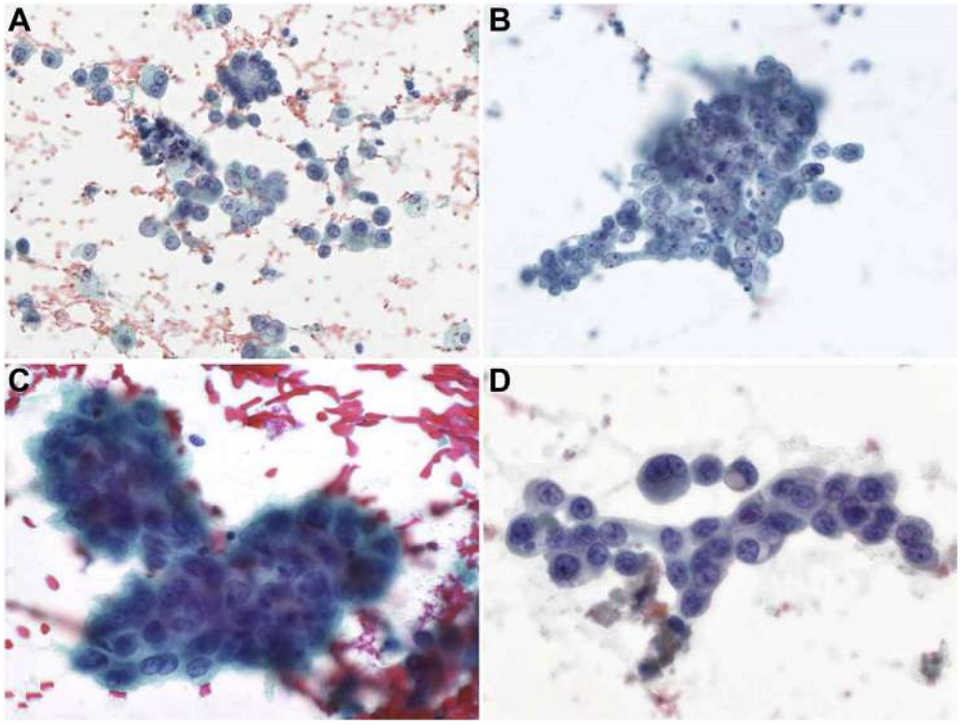

Figure 2. Characteristic cytologic features of LUAD (*A–D*, four examples, Papanicolaou). Single cells or rounded clusters of tumor cells with delicate/translucent cytoplasm. Vesicular nuclei with open chromatin and sometimes prominent nucleoli. One case with intracytoplasmic mucus containing a vacuole (*D*). LUAD, lung adenocarcinoma.

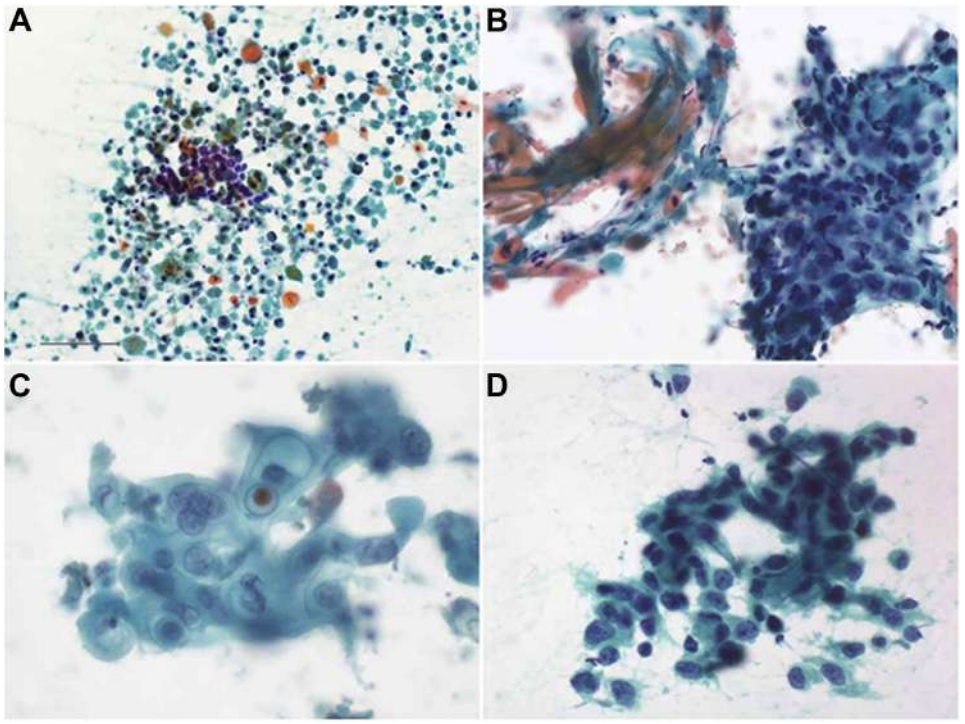

Figure 3. Characteristic cytologic features of LUSC (*A*–*D*, four examples, Papanicolaou). Tumor cells with hyperchromatic, dark nuclei, at least focal keratinization (orange, *A*–*C*), necrotic background (*A*), and dense/opaque cytoplasm (prominent features in *C* and *D*). LUSC, lung squamous cell carcinoma.

cohort, the developed predictor worked with an accuracy of 0.99 (1 of 125 misclassified LUSC; all concordant LUADs). In the validation subset, the accuracy of 0.99 was achieved (1 of 37 misclassified LUSC; all concordant LUADs). Testing the obtained algorithm on the test subset, a similar ($n = 167$; 3 of 50 misclassified LUSC and all concordant LUADs) accuracy of 0.98 was reached (Supplementary Table 1*A*).

As already mentioned in the Methods section, ML random forest is a robust predictor but a kind of black box; therefore, mainly for visualization purpose, we further analyzed the effect of the 12 cytopathologic features on the overall "coherent" classified LUAD and LUSC cases ($n = 792$) by means of a single regression tree. The resulted algorithm 1 is depicted in Figure 5.

The most important cytologic feature was keratinization, which when present, correctly classified 149 of 212 LUSCs and was absent in all 580 LUADs. Subsequently, the absence of dense (opaque) cytoplasm followed by absence of spindle/tadpole-shaped cells led to almost exclusive LUAD subtype, sequentially identifying 568 of 580 LUADs and 564 of 568 LUADs, respectively. In contrast, the presence of dense/opaque cytoplasm in the absence of keratinization identified 53 of the remaining 63 LUSCs, and subsequently, presence of spindle/tadpole cells in the absence of dense cytoplasm and keratinization identified eight of the remaining 10 LUSCs. Thus, the use of only three of the above-mentioned cytomorphologic features applied in a hierarchical pattern allows an accurate (0.97) classification, with only 2.7% (16 of 580) of LUADs and less than 1% (2 of 212) of LUSCs being misclassified. In other words, such an algorithm visualizes how the experts took their evaluation decisions when correctly evaluating the cases on the basis of the gold standard, including histology and/or IC.

Nevertheless, when this algorithm was tested on the overall collective (all 1474 responses), the accuracy

Table 4. Details of the 10 Cases That Were Misclassified or Failed Classification by All Experts

Gold Standard Diagnosis (No. of cases[a])	Failure to Classify (NSCC-NOS) (Number/Responses)	Misclassification (Number/Responses)
LUSC ($n = 7$)	52/86 (60%)	34/86 (40%)
LUAD ($n = 3$)	30/37 (81%)	7/37 (19%)

[a]Three (two LUSC and one LUAD) of the 10 cases were evaluated by 13 experts, the remaining seven by 12 experts.
LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; No., number; NSCC-NOS, non-small cell carcinoma—not otherwise specified.

Table 5. Common Cytologic Features Observed in the 10 Cases "Misclassified" by All Experts

Gold Standard LUAD ($n = 3$)		Gold Standard LUSC ($n = 7$)	
No. of responses	Common cytologic features	No. of responses	Common cytologic features
$n = 7$	Dense opaque cytoplasm ($n = 7$) Hyperchromatic nuclei ($n = 5$) Tadpole or spindle shaped cell without keratinization ($n = 5$)	$n = 34$	Vesicular nuclei with open chromatin ($n = 17$) Rounded 3D cell clusters ($n = 17$) Prominent nucleoli ($n = 16$) Translucent (delicate) cytoplasm ($n = 16$)

3D, three-dimensional; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; No. number.

reduced to 0.78 misclassifying 119 of 992 LUADs and 204 of 482 LUSCs (Supplementary Table 1*B*). It dropped further to 0.05 when tested only on the misclassified LUAD/LUSC ($n = 140$). On applying on the "failure to classify" LUAD responses ($n = 357$), the accuracy was 0.81 (289 of 357); however, when applied on the "failure to classify" LUSC responses ($n = 185$), the accuracy reduced drastically to 0.37 (68 of 185).

Here, it becomes clear that although algorithm 1 works perfectly within the "concordant" cases, it did not

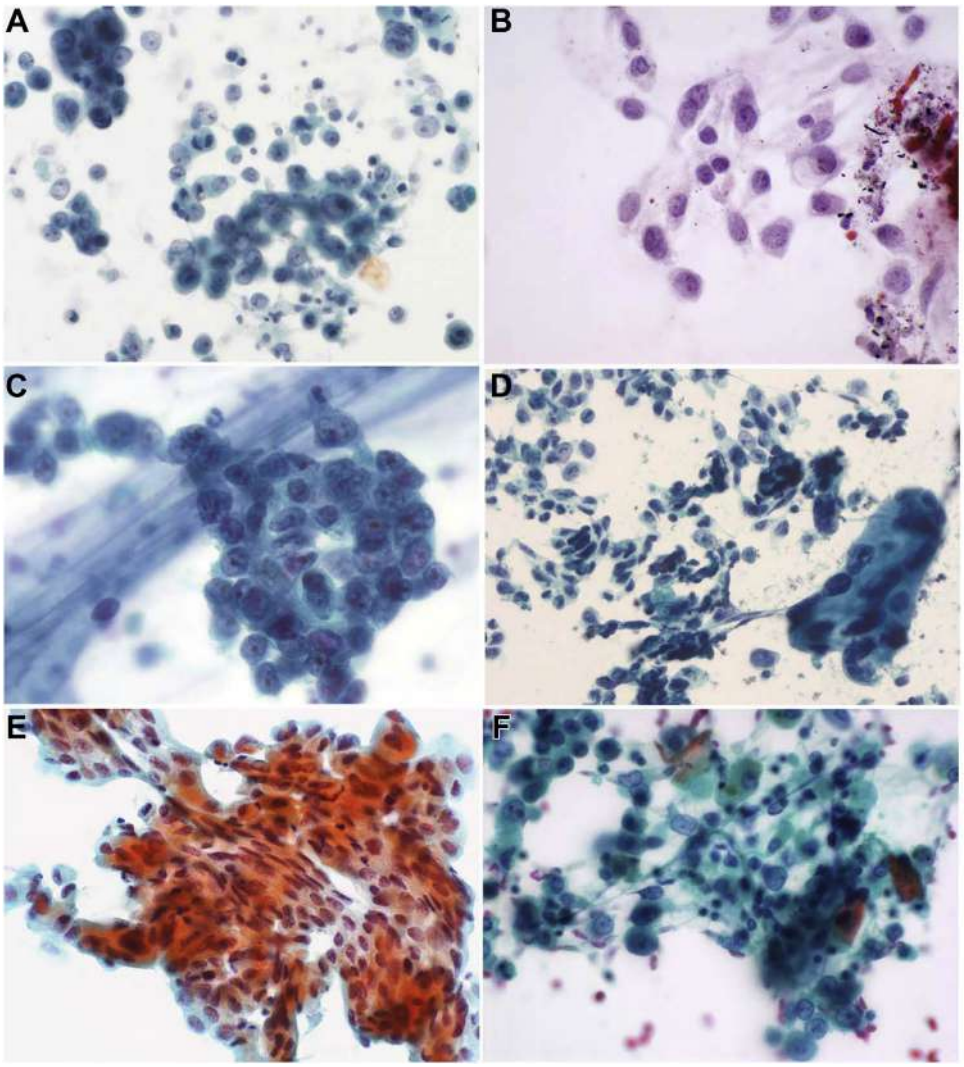

Figure 4. Cytologic features of frequently misclassified NSCLC (four examples, Papanicolaou). LUSC misclassified as LUAD (*A*-*C*) and LUAD misclassified as LUSC (*D*-*F*). (*A*) LUSC (case 92) classified as LUAD by six experts and as NOS by six. (*B*) LUSC (case 122) classified as LUAD by three experts and as NOS by nine. (*C*) LUSC (case 55) classified as LUAD and NOS by six experts, each. Four of the experts had misinterpreted the small round and empty space within the cell cluster as acinus lumen. (*D*) LUAD (case 36) classified as LUSC by four experts and as NOS by nine. (*E*) LUAD (case 91) classified as LUSC by three experts misinterpreting hypereosinophilia by technical variability as keratinization. (*F*) LUAD (case 60) classified as LUSC by five experts owing to focal, not otherwise specified hypereosinophilic structures mistaken for keratinization. LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; NOS, not otherwise specified.

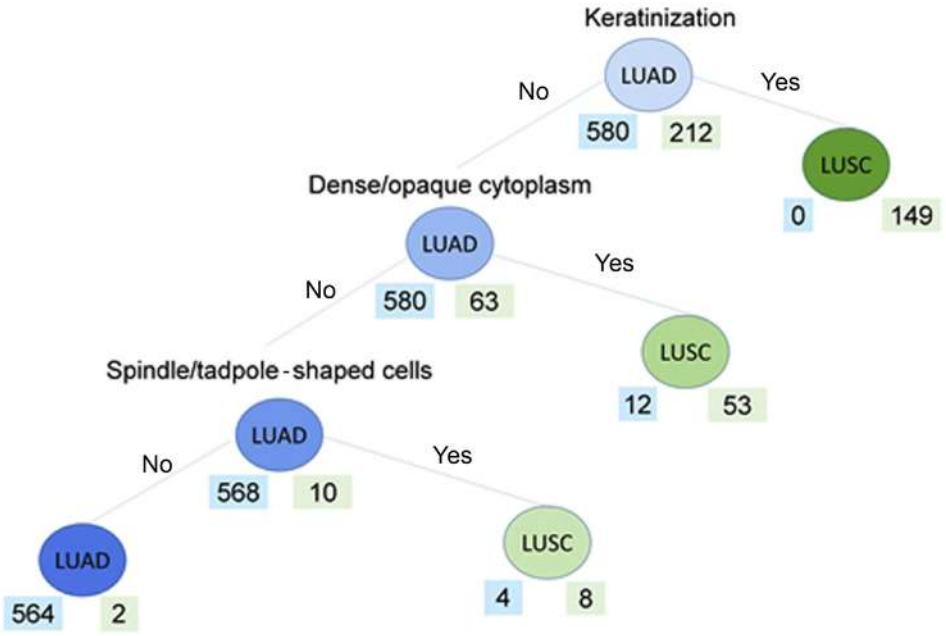

Figure 5. Regression tree algorithm for the classification of LUAD and LUSC on the basis of "concordant" responses (algorithm 1). Algorithm 1 indicates that if keratinization, dense, or opaque cytoplasm and tadpole or spindle cell shape without keratinization are absent, 97% of all LUADs were correctly classified, else, if present 99% of all LUSCs were correctly classified. The numbers below the nodes indicate the number of gold standard LUAD (on the left) and LUSC (on the right) cases. LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.

improve misclassification or failure rates in the misclassified or failure to classify cases.

Algorithm 2 Developed Based on "Concordant" and "Failure to Classify" Responses

A second random Forest predictor was calculated on the basis of the 792 "concordant" and 543 "failure to classify" responses ($n = 1334$; 580 concordant and 357 failed LUADs, 212 concordant and 185 failed LUSCs) with 30 trees and the same 11 cytomorphologic characteristics except acinus formation as described previously. Acinus formation was dropped by the algorithm because it did not help improving the classification of the algorithm built when evaluating the experts' chosen cytomorphologic characteristics of the "concordant" and "failure to classify" cases. The obtained accuracy in the three randomly selected subsets was 0.89 in the training ($n = 807$), 0.86 in the test ($n = 270$), and 0.83 in the validation ($n = 256$) subset, respectively (Supplementary Table 2*A*).

For visualization purposes, a single regression tree analysis resulted in algorithm 2 depicted in Figure 6, which when applied on all "concordant" and "failure to classify" responses returned an accuracy of 0.87 (Supplementary Table 2*B*).

Absence of keratinization without dense opaque cytoplasm or hyperchromatic nuclei accurately classified 78% of gold standard LUADs (730 of 937). In the presence of hyperchromatic nuclei observed in 90 gold standard LUADs, absence of necrotic debris correctly classified 86 LUADs (86 of 937). Presence of dense opaque cytoplasm without pyknotic chromatin but with vacuolated cytoplasm (27 of remaining LUAD classified) or eccentric nuclei (12 of remaining LUAD classified) identified an additional 4% of LUADs with overall accuracy of 91.2% for LUADs (855 of 937) with 8.7% misclassification rate (82 of 937). Keratinization (160 of 397 LUSCs) and, if nonkeratinized, dense opaque cytoplasm and pyknotic chromatin (42 of remaining 237) accurately classified LUSCs. Furthermore, 74 of remaining 195 LUSCs were correctly classified on the basis of the absence of pyknotic chromatin, accompanied by the absence of vacuolated cytoplasm and of eccentric nuclei. Absence of dense opaque cytoplasm but with hyperchromatic nuclei and necrotic debris (12 of remaining LUSCs classified) identified an additional 3% of LUSC with an overall accuracy of 72% for LUSC (288 of 397) but misclassification rate of 27.5% (109 of 397).

Although algorithm 2 eliminated cytology NSCC-NOS, misclassification rates were higher than those of the experts both in gold standard LUSC (28% by the algorithm versus 20% by the experts; $p = 0.057$) and in gold standard LUAD (8%

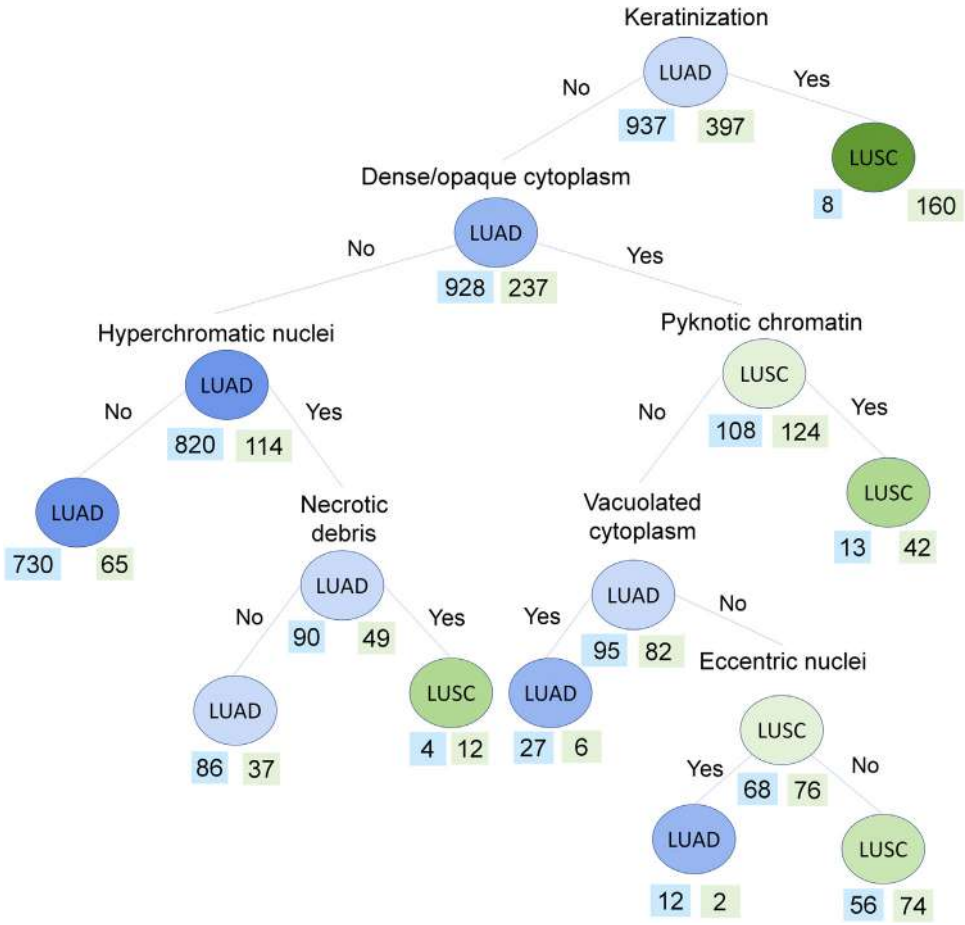

Figure 6. Regression tree algorithm for the classification of LUAD and LUSC on the basis of "concordant" and "failure to classify" responses (algorithm 2). Algorithm 2 correctly classified LUAD and LUSC with an accuracy of 0.86 but with a misclassification rate of 27.5% for LUSC and 8.7% for LUAD. The numbers below the nodes indicate the number of gold standard LUAD (on the left) and LUSC (on the right) cases. LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.

by the algorithm versus 5.5% by the experts; $p = 0.021$).

Discussion

In the present survey of 13 experts, we observed 53% accuracy in NSCLC subtyping by cytomorphology alone with individual accuracy rates of 58% for gold standard LUAD and 44% for gold standard LUSC. These are lower than previously reported rates of 66% to 99% and 53% to 96% for LUAD and LUSC, respectively.[7–12,14,15] Misclassifications constituted 10% of all responses largely comparable with published literature (1%–11%)[5,8–12,14] with LUSC being nearly four times more likely to be misclassified as compared with LUAD.[5,8–11] Classification failures were found in 37% of the instances, comparable with previous observations of 2% to 36%[5,7–12,14–16] with failure rates marginally higher among LUSC (38.4%) than LUAD (36%). The lower rates of accuracy, especially for LUSC, and the higher rates of classification failure in the present study can be explained by the intentional inclusion of difficult LUSC cases lacking overt keratinization, which is the subset that is, likely to be diagnostically challenging during real-time reporting. Furthermore, the survey was undertaken on evaluation of static images rather than traditional review of glass slides, and this could have had an effect on the accuracy.

It is important to note that nearly 95% of gold standard LUADs were classified as LUAD or NSCC-NOS on cytology by the experts in the present study. Because both of these categories are managed in the same manner in the clinical context, one may consider that cytomorphologic subtyping is adequate for LUAD diagnosis and management. Nevertheless, LUSC subtyping on cytology was a challenge with 18% being misclassified as LUAD. Nearly half of such misclassified LUSC had acini formation, rounded 3D clusters, and vesicular nuclei with prominent nucleoli, all features traditionally associated with LUAD. Furthermore, 38% of

LUSC failed to be subtyped (i.e., typed as NSCC-NOS). Notably, one-third of all NSCC-NOS responses in the present survey were made on gold standard LUSC diagnoses. This could result in a high number of LUSC cases that may be subject to clinically irrelevant molecular analysis if only morphology-based subtyping was done.

The uniqueness of the present study was the use of ML algorithms. Random forest and regression tree analyses were applied to understand the exact cytomorphologic features that aid in making an accurate subtype identification. Although 23 criteria were displayed with each case, most experts use only three to six (mean = 3.9) features for classification, and we found no statistical correlation with the number of criteria used with accuracy of subtyping. Performing a series of random forest analyses to develop ML predictors, we found that accuracy could improve to 0.87. Regression tree analysis allowed to visualize a hierarchy of cytologic features associated with the correct diagnosis of LUAD and LUSC. The final algorithm (algorithm 2 in Figure 6 and Supplementary Table 2) incorporated the most informative cytologic criteria from the "concordant" and "failed to classify" responses. Keratinization was the foremost helpful feature in subtyping, accurately identifying LUSC when present. In the absence of keratinization, subtyping becomes more challenging with hierarchy of other features being considered sequentially by algorithm 2 beginning with dense opaque cytoplasm. Dense opaque cytoplasm, traditionally considered as a feature of LUSC, was observed in a proportion of LUAD in the present study and drove most of the LUAD misclassifications as LUSC on cytology by the experts. On the basis of algorithm 2, the absence of dense opaque cytoplasm and hyperchromatic nuclei without necrotic debris predicted LUAD accurately, whereas presence of dense opaque cytoplasm with pyknotic nuclear chromatin in the absence of vacuolated cytoplasm and eccentric nuclei predicted LUSC accurately. Although the algorithm successfully classified all NSCC-NOS with high accuracy (>0.8), it increased the misclassification rates of LUSC and LUAD relative to the experts. These findings strongly suggest that suboptimal recognition of LUSC in the absence of keratinization is the major hurdle in improving cytology subtyping of NSCLC, and this was not improved by the ML algorithm that was primarily aimed at reducing NSCC-NOS rates in the present study. Thus, NSCC-NOS seems to be an inevitable morphologic diagnosis on conventional cytology. Misclassifications are also problematic; in particular, pseudokeratinized cells in LUAD and pseudoacinar formations in LUSC are potential morphologic pitfalls even for experts as observed in the present study. This re-emphasizes the need of ancillary IC for ensuring the accuracy of NSCLC subtyping in conventional cytology preparations. Although IC for TTF1/Napsin A and p40 are widely used in cell blocks by tissue-based protocols, more efforts are needed to adapt IC to ethanol-fixed cytologic specimens in a standardized manner.[23,24]

Surprisingly, cytologic features traditionally associated with LUAD such as intracellular mucin, extracellular mucin, nuclear pseudoinclusions, nuclear membrane irregularities, papillary formations with fibrovascular cores, and columnar cell shape did not emerge as decisive parameters in any of the ML algorithms in the present study. Intracellular mucin, which is a reliable criterion for LUAD,[2] could not be further explored in this study owing to low numbers of LUAD cases with mucin included in the gallery. Similar studies using ML methods with more cases (possibly consecutive cases without selection bias) increased number of participants, and with additional IC findings, will not only confirm the diagnostic value of the key differentiating cytologic features observed in the present study but also provide valuable insights for developing cytomorphologic criteria supplemented by IC. Application of automated computational/digital analysis using deep learning such as convoluted neural networks may identify additional "visual" or "subvisual" features not listed as criteria by expert cytopathologists may enhance classification on the basis of morphology alone.

In conclusion, except for keratinization, a single criterion is not sufficient to discriminate between LUSC and LUAD on cytology. ML identified additional cytomorphologic features, such as dense opaque cytoplasm, hyperchromatic nuclei, pyknotic nuclei, vacuolated cytoplasm, and eccentric nuclei, which were relevant in the accurate subtyping of NSCLC on cytology; however, some cases had superposition of these features limiting their usefulness. LUSC, in particular, was underdiagnosed in the absence of keratinization and contributed to a larger proportion of misclassifications in cytology. Thus, NSCC-NOS remains an inevitable diagnosis on cytomorphology alone and the algorithm developed in this study was able to reduce NSCC-NOS rates only at the cost of increasing misclassification rates. Any improvement in subtyping accuracy may only be achieved by ancillary IC supplementing cytomorphology. This reinforces the importance of using IC, which is not considered in many laboratories, and the need for validating IC protocols on cytology preparations.

CRediT Authorship Contribution Statement

Deepali Jain, Lukas Bubendorf: Conceptualization, Data curation, Formal analysis, Investigation,

Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Roles/Writing - original draft; Writing - review & editing.

Aruna Nambirajan: Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing.

Kim Geisinger, Kenzo Hiroshima, Lester Layfield, Yuko Minami, Andre L. Moreira, Noriko Motoi, Mauro Papotti, Natasha Rekhtman, Yasushi Yatabe: Data curation, Investigation, Methodology, Writing - review & editing.

Prudence A. Russell, Spasenija Savic Prince, Fernando Schmitt: Data curation, Investigation, Methodology, Writing - review & editing.

Serenella Eppenberger-Castori: Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - review & editing.

Acknowledgments

The authors thank Lenard Bubendorf for setting up the online survey and the image gallery. The views expressed in this manuscript are those of the authors and do not reflect the official policy of the Department of Defense or the U.S. government.

Supplementary Data

Note: To access the supplementary material accompanying this article, visit the online version of the *Journal of Thoracic Oncology* at www.jto.org/ and at https://doi.org/10.1016/j.jtho.2022.02.013.

References

1. Lindeman NI, Cagle PT, Aisner DL, et al. Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: guideline from the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. *Arch Pathol Lab Med*. 2018;142:321–346.
2. Travis WD, Scaliotti GV, Al-Dayel FH, Bubendorf L, Chung JH, Rekhtman N. WHO classification of tumors: small diagnostic samples. International Agency for Research on Cancer. https://tumourclassification.iarc.who.int/chapters/35. Accessed June 2, 2021.
3. Yatabe Y, Dacic S, Borczuk AC, et al. Best practices recommendations for diagnostic immunohistochemistry in lung cancer. *J Thorac Oncol*. 2019;14:377–407.
4. National Comprehensive Cancer Network. NCCN guidelines for non-small cell lung cancer, version 1.2018. https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf. Accessed February 3, 2022.
5. Sigel CS, Moreira AL, Travis WD, et al. Subtyping of non-small cell lung carcinoma: a comparison of small biopsy and cytology specimens. *J Thorac Oncol*. 2011;6:1849–1856.
6. Layfield LJ, Baloch Z, Elsheikh T, et al. Standardized terminology and nomenclature for respiratory cytology: the Papanicolaou Society of Cytopathology guidelines. *Diagn Cytopathol*. 2016;44:399–409.
7. Khayyata S, Yun S, Pasha T, et al. Value of P63 and CK5/6 in distinguishing squamous cell carcinoma from adenocarcinoma in lung fine-needle aspiration specimens. *Diagn Cytopathol*. 2009;37:178–183.
8. Rekhtman N, Brandt SM, Sigel CS, et al. Suitability of thoracic cytology for new therapeutic paradigms in non-small cell lung carcinoma: high accuracy of tumor subtyping and feasibility of EGFR and KRAS molecular testing. *J Thorac Oncol*. 2011;6:451–458.
9. Nizzoli R, Tiseo M, Gelsomino F, et al. Accuracy of fine needle aspiration cytology in the pathological typing of non-small cell lung cancer. *J Thorac Oncol*. 2011;6:489–493.
10. Manucha V, Wang C, Huang Y. Non-small-cell lung carcinoma subtyping on cytology without the use of immunohistochemistry - can we meet the challenge? *Acta Cytol*. 2012;56:413–418.
11. da Cunha Santos G, Lai SW, Saieg MA, et al. Cyto-histologic agreement in pathologic subtyping of non small cell lung carcinoma: review of 602 fine needle aspirates with follow-up surgical specimens over a nine year period and analysis of factors underlying failure to subtype. *Lung Cancer*. 2012;77:501–506.
12. Ebrahimi M, Auger M, Jung S, Fraser RS. Diagnostic concordance of non-small cell lung carcinoma subtypes between biopsy and cytology specimens obtained during the same procedure. *Cancer Cytopathol*. 2016;124:737–743.
13. Zakowski MF, Rekhtman N, Auger M, et al. Morphologic accuracy in differentiating primary lung adenocarcinoma from squamous cell carcinoma in cytology specimens. *Arch Pathol Lab Med*. 2016;140:1116–1120.
14. Patel TS, Shah MG, Gandhi JS, Patel P. Accuracy of cytology in sub typing non small cell lung carcinomas. *Diagn Cytopathol*. 2017;45:598–603.
15. Celik B, Bulut T, Khoor A. Subtyping of non-small cell lung cancer by cytology specimens: a proposal for resource-poor hospitals. *CytoJournal*. 2019;16:8.
16. Righi L, Graziano P, Fornari A, et al. Immunohistochemical subtyping of nonsmall cell lung cancer not otherwise specified in fine-needle aspiration cytology: a retrospective study of 103 cases with surgical correlation. *Cancer*. 2011;117:3416–3423.
17. Yildiz-Aktas IZ, Sturgis CD, Barkan GA, et al. Primary pulmonary non-small cell carcinomas: the College of American Pathologists Interlaboratory Comparison Program confirms a significant trend toward subcategorization based upon fine-needle aspiration cytomorphology alone. *Arch Pathol Lab Med*. 2014;138:65–70.
18. Nicholson AG, Gonzalez D, Shah P, et al. Refining the diagnosis and EGFR status of non-small cell lung carcinoma in biopsy and cytologic material, using a panel of mucin staining, TTF-1, cytokeratin 5/6, and P63, and

EGFR mutation analysis. *J Thorac Oncol*. 2010;5:436–441.

19. Pelosi G, Rossi G, Bianchi F, et al. Immunohistochemistry by means of widely agreed-upon markers (cytokeratins 5/6 and 7, p63, thyroid transcription factor-1, and vimentin) on small biopsies of non-small cell lung cancer effectively parallels the corresponding profiling and eventual diagnoses on surgical specimens. *J Thorac Oncol*. 2011;6:1039–1049.
20. Bannach-Brown A, Przybyła P, Thomas J, et al. Machine learning algorithms for systematic review: reducing workload in a preclinical review of animal studies and reducing human screening error. *Syst Rev*. 2019;8:23.
21. Liaw A, Wiener M. *Package 'randomForest': Breiman and Cutler's Random Forests for Classification and Regression*. Vol. 4. R Development Core Team; 2014:6–10.
22. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees. Wadsworth Statistics/Probability Series*. Belmont, CA: Wadsworth Advanced Books and Software; 1984.
23. Jain D, Nambirajan A, Borczuk A, et al. Immunocytochemistry for predictive biomarker testing in lung cancer cytology. *Cancer Cytopathol*. 2019;127:325–339.
24. Nambirajan A, Jain D. Cell blocks in cytopathology: an update. *Cytopathology*. 2018;29:505–524.