Research paper

A prognostic and predictive computational pathology image signature for added benefit of adjuvant chemotherapy in early stage non-small-cell lung cancer

Xiangxue Wanga, Kaustav Beraa, Cristian Barreraa, Yu Zhoua, Cheng Lua, Pranjal Vaidya,a, Pingfu Fuba, Michael Yangc, Ralph Alexander Schmide, Sabina Berezowskaef, Humberto Choi, Vamsidhar Velchetih, Anant Madabhushi,*.†,‡

‡ Center for Computational Imaging and Personalized Diagnostics, Case Western Reserve University, OH, USA
† Department of Population and Quantitative Health Sciences, Case Western Reserve University, OH, USA
§ Division of General Thoracic Surgery, University Hospital Berne, Bern, Switzerland
© Institute of Pathology, University of Bern, Bern, Switzerland
¢ Department of Pathology-Anatomic, University Hospitals, OH, USA
£ Department of Laboratory Medicine and Pathology, Institute of Pathology, Lausanne University Hospital and Lausanne University, Lausanne, Switzerland
∞ Department of Hematology and Oncology, NYU Langone Health, NY, USA

ABSTRACT

Background: We developed and validated a prognostic and predictive computational pathology risk score (CoRiS) using H&E stained tissue images from patients with early-stage non-small cell lung cancer (ES-NSCLC).

Methods: 1330 patients with ES-NSCLC were acquired from 3 independent sources and divided into four cohorts D1-4. D1 comprised 100 surgery treated patients and was used to identify prognostic features via an elastic-net Cox model to predict overall and disease-free survival. CoRiS was constructed using the Cox model coefficients for the top features. The prognostic performance of CoRiS was evaluated on D2 (N=331), D3 (N=657) and D4 (N=242). Patients from D2 and D3 which comprised surgery + chemotherapy were used to validate CoRiS as predictive of added benefit to adjuvant chemotherapy (ACT) by comparing survival between different CoRiS defined risk groups.

Findings: CoRiS was found to be prognostic on univariable analysis, D2 (hazard ratio (HR) = 1.41, adjusted (adj.) P = .01) and D3 (HR = 1.35, adj. P < .001). Multivariable analysis showed CoRiS was independently prognostic, D2 (HR = 1.41, adj. P = .001) and D3 (HR = 1.35, adj. P < .001), after adjusting for clinico-pathologic factors. CoRiS was also able to identify high-risk patients who derived survival benefit from ACT D2 (HR = 0.42, adj. P = .006) and D3 (HR = 0.46, adj. P = .08).

Interpretation: CoRiS is a tissue non-destructive, quantitative and low-cost tool that could potentially help guide management of ES-NSCLC patients.

Funding: Data collection, analysis, and computation resources of the research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award numbers: 1U24CA199374-01, R01CA202752-01A1, R01CA208236-01A1, R01 CA216579-01A1, R01 CA220581-01A1, 11U01 CA239055-01, National Center for Research Resources under award number 1 C06 RR12463-01. VA Merit Review Award IBX004121A from the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service, the DOD Prostate Cancer Idea Development Award (W81XWH-15-1-0558), the DOD Lung Cancer Investigator Initiated Translational Research Award (W81XWH-18-1-0440), the DOD Peer Reviewed Cancer Research Program (W81XWH-16-1-0329), the Ohio Third Frontier Technology Validation Fund, the Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering and the Clinical and Translational Science Award Program (CTSA) at Case Western Reserve University.

© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

Early-stage non-small cell lung cancer (ES-NSCLC) usually comprises of stage I and II cancers and complete surgical excision is the...
standard of care treatment for these patients [1–3]. While current guidelines recommend adjuvant cisplatin based chemotherapy (ACT) for stage II patients, stage I patients continue to be treated with surgery alone. Benefit of ACT following surgical resection has been shown in multiple large clinical trials, with significant improvement in overall survival (OS) and Disease-free survival (DFS). Clinical parameters such as tumor stage, nodal status, age, and performance score have been traditionally shown to be prognostic, but presently there is a paucity of accurate and validated biomarkers based off clinicopathologic factors which can identify patients who would benefit from ACT. With the advent of digital pathology and the corresponding increase in machine learning and computerized pathology image analysis, there is the opportunity to mine and associate quantitative features relating to tumor morphology to cancer prognosis and outcome. A few groups have recently shown that machine learning based prognostic classifiers can predict disease recurrence and survival in the context of NSCLC.

### 2. Patients and methods

#### 2.1. Ethics statement

An Institutional review board (IRB) approved protocol was used for the retrospective analysis, and the informed consent requirement was waived by the IRB. The study was compliant with the Health Insurance Portability and Accountability Act (HIPPA). All data used in this study were de-identified and no protected health data was needed.

#### 2.2. Patients

Retrospective chart review of patients continuously admitted in the Cleveland Clinic Foundation (CCF) with NSCLC between 2005–15 yielded 670 patients. All resected stage I and II NSCLC were included in the study, however those tissue slides which following pathologic evaluation did not meet quality requirements such as poor stain yield 670 patients. Of these, 83 patients received ACT. 100 patients with surgery alone formed the discovery cohort (D_1) while the remaining N=331 formed the validation cohort (D_2). The Cancer Genome Archive (TCGA)-lung adenocarcinomas (ADC; N=523) and TCGA-lung squamous cell carcinomas (SCC; N=409) cohorts are publicly available datasets, assembled from different institutions [20]. After applying the inclusion and exclusion criteria, a TCGA derived independent validation cohort D_3 (N=657; ADC=378, SCC=279) was identified. In D_3, 179 patients received ACT. Additionally, a cohort of N=269 consecutive, primary resected ES-NSCLC patients from the University of Bern was used for prognostic validation and formed D_4 (N=242 after applying the inclusion and exclusion criteria). D_4 comprised only SCC cases
and all of the patients underwent surgery with 62 patients received ACT. Patient demographics and characteristics are summarized in Table 1. (Fig. 2).

### 2.3. Image acquisition

Whole slide images (WSI) obtained from routine H&E diagnostic tissue slides of the primary tumor were collected for D1, D2 and D3. D4 was collected in the form of tissue microarrays (TMA) to represent the core of the tumor. The H&E slides in D1 and D2 were scanned using a Roche-Ventana iScan HT scanner (serial #: BI15N7205) at a magnification of 20x. D3 was a publicly acquired dataset from multiple institutions and heterogeneous scanners. The different pathology labs that contributed studies to the TCGA likely used different vendors for the whole slide scanning, unfortunately the specific scanner make and model for the individual TCGA sites were not available. For our analysis, the image and feature analysis were consistently performed at 20x magnification for all datasets. D4 was in the form of TMAs scanned at 40x (down-sampled to 20x) by Panoramic Digital Slide Scanner 250 (version: 1.23.1.71684).

### 2.4. Automatic tumor detection and segmentation of cancer nuclei and perinuclear region

A U-Net based convolutional neural network was employed for segmentation with adversarial training (training rates are 0.001 and 0.01 for regular and adversarial training, respectively) [21]. Two different U-net based models were trained for tumor detection and nuclei segmentation respectively. While the tumor detector was used to generate the heat map for WSI to indicate the probability of tumor, the nuclei segmentation model was used to delineate the boundary pixels of each nuclei. The perinuclear region was segmented by taking 15 pixels at 20x magnification outward from the boundaries of nuclei. Each of WSI was sliced into 2000 by 2000 pixels consecutive tiles and only the tiles from the detected tumor regions were used to represent each patient. The ground truth set of nuclei and tumor was generated by two pathologists from University Hospital. For nuclei

---

**Table 1**

| Variable               | D1 N(%) | D2 N(%) | D3 N(%) | D4 N(%) | Total N | Adj. P |
|------------------------|---------|---------|---------|---------|---------|--------|
| Number of patients     | 100     | 331     | 657     | 242     | 1330    | <0.001 |
| Age (mean+/-std year)  | unknown | unknown | 64 +/-15| 67 +/-8  | 65 +/-13| <0.001 |
| Gender                 | Male    | unknown | unknown | 438(66.7)| 207(85.5)| 645    |
| Tumor Size (mean+/-std mm) | 201x131 | 206(80.7)| 580(88.3)| 46.61+/-23.8| 39.36+/-46.66| <0.002 |
| Smoking Status         | Previous/Current | Never | 61(18.4)| 155(23.6)| 88(26.4)| 312    |
| pN                     | 0       | 82(82)  | 230(69.5)| 491(74.7)| 152(62.8)| 955    |
| pT                     | 1       | 8(8)    | 61(18.4)| 155(23.6)| 88(26.4)| 312    |
| Overall Stage          | IA      | 45(45)  | 108(32.6)| 223(31.9)| 65(26.9)| 441    |
| Treatment              | Surgery Only | Surg. + Chemo | 3(0.9)| 5(0.8)| 1(0.4)| 9      |
| Recurrence             | Non-recurrence | Recurrence | 66(19.9)| 141(21.5)| 75(31.0)| 296    |
| Tumor types            | Adenocarcinoma | SCC | 41(12.4)| 90(13.7)| 64(25.4)| 203    |
| Abbreviations: SCC, squamous cell carcinoma; Surg. + Chemo: surgery and chemotherapy; Adj. p: adjusted p; pT: pathological tumor stage; pN: pathological nodal stage.
segmentation, 8000 nuclei were annotated from 100 digitized H&E images of breast (40) and lung cancer (60). For tumor annotation, 125 whole slide sections of lung cancer were manually annotated, 80 of them were used for training and rest 45 were used for validation. The tumor detector achieved 90.6% patch-level accuracy on a ground truth set curated by pulmonary pathologists. Meanwhile, nuclei segmentation model yielded an F-score 0.88, comparable to current state-of-the-art nuclear segmentation algorithm [22]. All U-Net based segmentations were implemented in Tensorflow 1.6 on Nvidia Titan XP GPU clusters (network details are specified in supplementary Table S3).

2.5. Quantitative feature extraction

We extracted 242 nuclear descriptors from previously segmented nuclei. These features corresponded to five categories: nuclei shape [23], orientation entropy [24], texture, local and global graph [25]. On the one hand, shape features included basic measurements of nuclear area, perimeters and mathematical descriptors of contour. Orientation entropy and texture on the other hand characterized directionality coherence and pixel intensity distribution of the nuclei [24]. While local graph measured the architecture in relation to neighborhood cells, global graph captured the arrangement in relation to the entire WSI [6]. A graph was a mathematical operation which included a set of nodes (nucleus) to capture relationships through pairwise edges formed between the nodes (Details in supplementary Table S3).

Another 35 peri-nuclear features from the adjoining cytoplasmic area were also extracted. These included quantitative measurements including area, area ratio of the peri-nuclear region to the nuclear region, pixel intensities as well as texture [26]. These features would not only characterize the space adjoining the nuclei in the cancer cell but would also highlight relationships between the nucleus and cytoplasm in the cancer cell (Details in supplementary Table S2). Finally mean, standard deviation, min and max value of patch-level features were calculated and concatenated to generate a patient level image signature. The feature extraction was implemented in MatLab 2020a.

2.6. Constructing the CoRiS risk score

A total of 277 nuclear and peri-nuclear features were extracted for each patient in D1. In order to regularize the number of features proportionate to samples size [27], the top discriminative features were selected by Elastic-Net regularization with non-zero coefficients and these features were fit into Cox Proportional Hazard model with OS
and DFS as the outcomes of interest, respectively. The CoRiS was computed by a weighted linear combination of selected features and their corresponding coefficients. The tradeoff value alpha (mixing parameter) between L1 and L2 for elastic net was evaluated from 0 to 1 with step size of 0.1. It was determined that 0.8 would be the optimal value. The optimal value of the tuning parameter in the Elastic-Net Cox (lambda) was determined by 10-fold cross validation in $D_1$.

### 2.7. Statistical analysis

OS was measured from the date of diagnosis to the date of death and censored at the date of last follow-up for survivors. DFS was calculated from the date of surgery to the date of recurrence or death whichever occurred earlier and censored at the date of last follow-up for those still alive without recurrence. The CoRiS was divided into low and high risk based on the median value of CoRiS classifying OS/DFS obtained on $D_1$. Further stratification of CoRiS was done by dividing it into four groups ($H$, $I_H$, $I_L$, $L$) based on quartile values from training CoRiS. High-risk group comprised the upper two quartiles, high ($H$) and intermediate high ($I_H$), while low risk comprised lower two quartiles, intermediate low ($I_L$) and low ($L$). Univariable analysis of CoRiS and the clinicopathologic variables (i.e. smoking history, tumor subtypes, pathological stage) were conducted. Multivariable Cox-regression models were built to assess the relationships between the various covariates and OS/DFS while adjusting for baseline factors [28]. Forest plots were constructed to show the HRs comparing OS between ACT and the surgery alone group in all the cohorts with patients being stratified based on the quartiles of CoRiS. Further sub-set analysis involved looking at survival differences in the different AJCC stages and subtypes of tumor, i.e. SCC and ADC. Kaplan-Meier survival curves were obtained to visualize the differences based on CoRiS and Hazard Ratios were computed. All p-values were adjusted based on Benjamin and Hochberg’s procedure[29,30] and a significance level of 0.05 was set to be statistically significant. Statistical analysis was implemented in R 3.6.1.

### 3. Experiments and results

#### 3.1. CoRiS predicts OS and DFS on validation sets, independent of clinicopathologic factors

Eleven most discriminative features were selected to construct CoRiS ($Figs. S1–S2$). CoRiS, ($Fig. 3$) was found to be prognostic of both OS ($D_1$, HR = 2.97, 95% CI: 1.87–4.71, adjusted (adj.) $P < .001$; $D_2$, HR = 1.33, 95% CI: 1.05–1.68, adj. $P = .067$; $D_3$, HR = 1.62, 95% CI: 1.24–2.11, adj. $P < .001$; $D_4$, HR = 1.54, 95% CI: 1.03–2.3, adj. $P = .082$, $Figs. 3$ and S5) and DFS ($D_1$, HR = 2.4, 95% CI: 1.54–3.73, adj. $P < .001$; $D_2$, HR = 1.27, 95% CI: 1.01–1.61, adj. $P = .082$; $D_3$, HR = 1.7, 95% CI: 1.24–2.33, adj. $P = .028$; $D_4$, HR = 1.54, 95% CI: 1.05–2.25, adj. $P = .082$, $Figs. 3$ and S6). On univariable analysis, only CoRiS was significantly prognostic among all test sets (Table 2). Kaplan-Meier (KM) curves for predicting OS/DFS by CoRiS are shown in $Fig. 3$ for $D_1–D_3$. On multivariable analysis with controlling covariates, CoRiS was found to be independently prognostic ($D_2$, HR = 1.24, 95% CI: 1.06–1.39, adj. $P = 0.01$ and $D_3$, HR = 1.14, 95% CI: 1.02–1.28, adj. $P = 0.01$ Table 2). In addition, CoRiS could separately predict the OS/DFS in two major NSCLC subtypes, ADC and SCC. Details of the subtype analysis are included in Figs. S5–S8.

#### 3.2. CoRiS predicts ACT benefit in two independent validation sets

CoRiS classified 38 and 45 patients who received ACT into low-risk and high-risk groups based on median CoRiS in $D_2$ ($Fig. 4a$). Similarly, in $D_3$, CoRiS classified 93 and 86 of those patients who received ACT into low-risk and high-risk groups respectively ($Fig. 4c$). Survival comparisons between the groups (low-and high risk) in patients who received ACT showed no statistically significant difference in OS for both $D_2$ ($HR = 0.83$ 95% CI: 0.52–1.32, adj. $P = .631$, $Fig. 4a$) and $D_3$ ($HR = 1.42$, 95% CI: 0.86–2.37, adj. $P = .218$, $Fig. 4c$). In contrast, for patients who underwent surgery alone without ACT, there was a statistically significant difference in OS between the low and high risk groups in $D_2$ ($HR = 1.75$, 95% CI: 1.33–2.31, adj. $P < .001$, $Fig. 4b$) and $D_3$ ($HR = 1.73$, 95% CI: 1.26–2.36, adj. $P = .004$, $Fig. 4d$). Results for DFS were similarly significant (see Supplementary Fig. S9). Granular analysis of CoRiS showed the patients with increased risk ($H$ and $I_H$) tended to have longer survival when ACT was administered. The $H$ group showed improved median OS by about 35 months (95% longer) when ACT was given. In the $I_H$ group, median OS was found to be higher by 21 months (58% longer) in $D_2$ (HR=0.42, 95% CI: 0.26–0.69, adj. $P = .006$, $Fig. 4e$) and 46 months (115% longer) in $D_3$ (HR=0.46, 95% CI: 0.24–0.87, adj. $P = .082$, $Fig. 4f$) between ACT and surgery alone patients. In the $I_L$ and $L$ groups, the ACT population showed worse survival as compared to the surgery alone group but was not statistically significant (adj. $P > .05$) in both $D_2$ and $D_3$ ($Fig. 4e, f$). Estimated survival benefit according to DFS was similar to OS, higher CoRiS showing an estimated DFS benefit in $D_2$ (HR=0.36,
X. Wang et al. / EBioMedicine 69 (2021) 103481

ACT in stage IB patients with a high risk of recurrence [11,31]. There is a need for predictive biomarkers to identify tumors at higher risk of recurrence [12,116–19]. While the American Society of Clinical Oncology guidelines do not recommend ACT in stage IB patients, the NCCN guidelines currently only recommend ACT in stage IB patients with a high risk of recurrence [11,31]. There is thus a need for predictive biomarkers to identify tumors at higher risk of recurrence and would be potential candidates for ACT. Identifying the low risk patients who will do well with surgery alone would spare them from toxicity of ACT.

Amongst the existing biomarkers in NSCLC, most are prognostic and reliant on molecular or multi-gene assays. For instance, many studies showed that class III β-tubulin expression, abnormalities in the k-ras oncogene and p53 tumor suppressor gene, and DNA methylation markers could potentially identify the high-risk patients who would benefit from ACT [32,33]. The only known molecular assay predictive of benefit to ACT was published by Zhu et al. [19] who showed that a 15-gene signature was not only prognostic but predicted improved survival after ACT in signature defined high-risk patients (HR = 0.33; 95% CI: 0.17 to 0.63; P = .0005), but not in low-risk patients (HR = 3.67; 95% CI: 1.22 to 11.06; P = .0133; interaction P < .001) [19]. However, all the mentioned biomarkers are tissue destructive, expensive and time-consuming involving RNA expression and microarray profiling analysis.

4. Discussion

Due to contradictory results from multiple clinical trials, ACT is currently not recommended in stage IA while there is controversy regarding its use in stage IB patients [1,2,11,16–19]. While the American Society of Clinical Oncology guidelines do not recommend ACT in stage IB patients, the NCCN guidelines currently only recommend ACT in stage IB patients with a high risk of recurrence [11,31]. There is thus a need for predictive biomarkers to identify tumors at higher risk of recurrence and would be potential candidates for ACT. Identifying the low risk patients who will do well with surgery alone would spare them from toxicity of ACT.

Amongst the existing biomarkers in NSCLC, most are prognostic and reliant on molecular or multi-gene assays. For instance, many studies showed that class III β-tubulin expression, abnormalities in the k-ras oncogene and p53 tumor suppressor gene, and DNA methylation markers could potentially identify the high-risk patients who would benefit from ACT [32,33]. The only known molecular assay predictive of benefit to ACT was published by Zhu et al. [19] who showed that a 15-gene signature was not only prognostic but predicted improved survival after ACT in signature defined high-risk patients (HR = 0.33; 95% CI: 0.17 to 0.63; P = .0005), but not in low-risk patients (HR = 3.67; 95% CI: 1.22 to 11.06; P = .0133; interaction P < .001) [19]. However, all the mentioned biomarkers are tissue destructive, expensive and time-consuming involving RNA expression and microarray profiling analysis.

In this work we presented CoRiS, the first of its kind digital pathology based companion diagnostic test, which is not only prognostic but also predictive of added benefit of ACT in ES-NSCLC. The CoRiS comprises 11 features relating to nuclei and peri-nuclear histomorphometric attributes obtained from digitized H&E tissue images. We used a group of resected ES-NSCLC without ACT to train CoRiS as prognostic model. CoRiS was further independently validated on multiple sets (independent of clinical factors such as tumor stage and smoking history; see Table 2).

For predicting benefit to ACT, the two top CoRiS groups (H, Ia) showed statistically significant survival benefit for validation set D2. While CoRiS did not yield the same significant survival difference after p-value adjustment (adj. P > .01) on D3, there was a clear trend that patients who received ACT had a longer median survival time. In fact, the highest risk CoRiS group (H) showed >90% median OS improvement (Fig. 4e, f) for the ACT as compared to the surgery alone patients. Interestingly, the low CoRiS groups (Ia, L) across the validation cohorts showed no statistically significant differences in HR between the surgery alone and ACT groups, and in some cases showed detrimental effects of ACT (HR > 1; Fig. 4e, f). This seems to suggest that patients in the CoRiS low group (L) would do equally well with surgery alone and can be spared the deleterious effects of ACT.

A subgroup analysis on stage IB patients showed that CoRiS divided D2-4 into low risk (Fig. S3d) and high risk (Fig. S3e) groups. This illustrates that the patients who received ACT had reduced hazard of dying in high risk group but in the low risk group, no survival difference between two cohorts of patients (with or without ACT) was identified with HR = 0.96 and adj. P = .899. In addition, the CoRiS defined high-risk group had significantly improved OS with ACT versus surgery alone (Fig. 4e). However, the low risk CoRiS group had no additional benefit with ACT (Fig. 4e). While the CALGB9663 and the LACE meta-analysis showed a small but non-significant statistical benefit to HR in stage IB with ACT, the IALT and the JBR10 trials did not show OS differences in stage IB [2,4,5]. These results thus seem to
suggest a basis for the non-significant benefit to ACT in the completed clinical trials in stage IB patients. The combination of two distinct risk groups within a homogenous clinically defined stage could be a possible reason for the low benefit to ACT seen in published studies.

Meanwhile, in stage II patients where the present recommendation is ACT following resection, the CoRiS signature identified a low-risk group that did not have a significantly different HR when compared to the surgery alone group (Fig. S3f), thus potentially identifying and unveiling a group with relatively good survival that might be spared the toxicity of ACT.

Machine learning approaches have been applied to digital pathology images for different cancer types to prognosticate patients’ outcome[21,22]. To the best of our knowledge, CoRiS, is different from previous works[19,34] in that it is not just prognostic of risk of recurrence but also predictive of added benefit of ACT for early stage NSCLC. Wang et al. demonstrated that nuclear shape and texture features based off H&E biopsy TMAs could identify patients who would recur following surgery in early-stage NSCLC[6]. Meanwhile, Corredor et al. showed that the spatial architecture of tumor infiltrating lymphocytes (TILs) was prognostic of recurrence free survival in several independent validation datasets, pathologist based TIL estimation by comparison was not prognostic for those datasets [4]. The selected features for CoRiS includes descriptors characterizing nuclei arrangement and abundance. For example, the average area of the Voronoi diagram obtained by connecting nuclei could allow for capture of the number and proximity of nuclei within the tumor. In addition, the texture feature (Haralick) characterizing the peri-nuclei region might be reflective of the coherence of extracellular staining, in other word, less aggressive tumor might present with more homogenous cytoplasm formation. This work was significantly different from previous related publications [4,6,23] by (a) the features comprising CoRiS included not just the morphology and spatial arrangement (i.e. Voronoi and cell cluster graph) of nuclei but also a set of innovative peri-nuclear features (texture features from peri-nuclei region); (2) CoRiS has shown to be not only prognostic to patients’ outcome but also predictive to benefits of ACT; and (3) CoRiS was validated on over 1000 patients from multiple different institutions.

The study did have its limitations. Firstly, CoRiS was developed and validated using retrospective data from different institutions, which means the pathological staging criteria applied might have varied at the time of tissue examination [35], additionally at least a few demographic related parameters were not available for some of the datasets. Secondly, for predicting benefit to ACT, the surgery only and the surgery+ACT groups used in the analysis were not strictly and homogeneously controlled (including ACT protocol), it is likely that the assignment and protocol of ACT might have differed across
the institutions considered in this study. Recently, transfer learning based approaches have been applied in tumor detection and classification [36,37]. An avenue for future investigation might involve the use of transfer learning, potentially leveraging other data streams like quantitative immunofluorescence, for the problems of cancer prognosis and response prediction. While the difference of median OS between CoRiS defined low and high risk group is over 90% in D2, the survival benefit between two groups is not significant after p-value adjustment. However, the effect sizes (HRs) of CoRiS across the validation datasets (D2 and D3) are similar (see Table 2). Multiple test correction using FDR approach was done only for each task separately rather than doing considering all related survival prediction tasks together. As a result, the family-wise error rate was not controlled at 0.05 level for all classification tasks. In addition, we focused on features relating to nuclei from within cancer-identified regions on H&E images without differentiating cell types (i.e. tumor and lymphocyte cells). Computerized discrimination of the different cell type categories (i.e. lymphocytes, cancer nuclei, fibroblasts, macrophages) can be challenging on H&E images alone, and we were cautious to include another possible confounding variable to our predictor. Manually checking the fidelity of the detected lymphocytes on over 1,000 H&E WSIs from multiple institutions was clearly not feasible. Additionally, we did not have access to immunohistochemistry (IHC) or quantitative multiplex immunofluorescence (qmIF) images that would have allowed us to better define and employ features from different immune cell subtypes (e.g. CD4, CD3, and CD20). Clearly an avenue for future investigation will be the possible combination of features from H&E images along with corresponding features from IHC and qmIF images. Deep learning has shown better performance in different tumor segmentation tasks compared to hand-crafted based approaches [38,39]. However, in detecting tumor regions directly from whole slide images (WSI), our approach based on U-Net and adversarial training achieves comparable results (Table S4) to recently published deep learning methods in terms of both accuracy and computational efficiency [40]. For clinical utility and deployment, CoRiS needs to be prospectively validated, and needs to be applied on clinical trials with randomly assigned patients to surgery and surgery +ACT to truly validate its utility in predicting benefit to ACT. In summary, we developed and validated an 11-feature prognostic and predictive signature for ACT benefit in patients with ES-NSSCL. With additional validation, possibly in the context of clinical trials like JBR10 and IALT[4,7], CoRiS could be validated as an inexpensive, tissue non-destructive, prognostic and predictive companion diagnostic for ES-NSSCL that could possibly have global impact.

Data sharing statement

The CoRiS codes and related data during the current study were available at:
- Nuclei segmentation codes: https://github.com/maberyick/ nucleiSegmentationHEDL
- Nuclei data: https://cwru.app.box.com/s/3pf10foxpzgzn-fwztviif9b5uxdr4
- Tumor detection codes: https://github.com/maberyick/ TumorSegmentationHUNET
- Tumor data: https://cwru.app.box.com/s/vq1q01x6cifjbjb8v-56tqbn1nb340dp
- Quantitative feature codes: https://github.com/maberyick/ periNuclearHE

Declaration of Competing Interest

X.W., K.B., C.B., Y. Z., C.L., P.V., P. F., M.Y., H.C. disclosed no conflicts of interest. S.B. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: institution has grants or grants pending with Roche and Basilea; institution has received payment for conference organization from Roche, MSD, AstraZeneca, Agilent and Biosystems; is on the speakers bureaus of MSD, AZ and Roche; has received payment for the development of educational presentations from Roche. Other relationships: disclosed no relevant relationships. V.V. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: is a consultant for BMS, Genentech, AstraZeneca, Celgene, Foundation Medicine, Taekeda, Merck, Alkermes, and Nektar Therapeutics; institution has grants or grants pending with AstraZeneca, Merck, BMS, Genentech, and Alkermes; is on the speakers bureaus of Novartis, BMS, Celgene, and Foundation Medicine; has received payment for the development of educational presentations from BMS and Foundation Medicine. Other relationships: disclosed no relevant relationships. A.M. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed no relevant relationships.

Contributors

X.W., C.B., Y. Z., C.L., P.V., P. F., A.M. designed and conducted the experiments. X.W., K.B., C.B., A.M. wrote the manuscripts. M.Y., R.S., S.B., H.C., V.V. acquired and provided the image and clinical data and also the medical writing guidance. All authors read and approved the final version of the manuscript and have verified the underlying data.

Funding

Data collection, analysis, and computation resources of the research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award numbers: 1U24CA199374-01, R01CA202752-01A1, R01CA208236-01A1, R01 CA216579-01A1, R01 CA220581-01A1, 1U01 CA239055-01. National Center for Research Resources under award number 1 C06 RR12463-01. VA Merit Review Award IBX004121A from the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service, the DOD Prostate Cancer Idea Development Award (W81XWH-15-1-0558), the DOD Lung Cancer Investigator-Initiated Translational Research Award (W81XWH-18-1-0440), the DOD Peer Reviewed Cancer Research Program (W81XWH-16-1-0329), the Ohio Third Frontier Technology Validation Fund, the Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering and the Clinical and Translational Science Award Program (CTSA) at Case Western Reserve University.

Acknowledgements

Research reported in this publication was supported by the National Cancer Institute under award numbers 1U24CA199374-01, R01CA249992-01A1, R01CA202752-01A1, R01CA208236-01A1, R01CA216579-01A1, R01CA220581-01A1, R01CA257612-01A1, 1U01CA239055-01, 1U01CA248226-01, 1U54CA254566-01, National Heart, Lung and Blood Institute 1R01HL151277-01A1, National Institute of Biomedical Imaging and Bioengineering 1R43EB028736-01, National Center for Research Resources under award number 1 C06.
RR12463-01. VA Merit Review Award IBX004121A from the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service, the Office of the Assistant Secretary of Defense for Health Affairs, through the Breast Cancer Research Program (W81XWH-19-1-0668), the Prostate Cancer Research Program (W81XWH-15-1-0558, W81XWH-20-1-0851), the Lung Cancer Research Program (W81XWH-18-1-0440, W81XWH-20-1-0595), the Peer Reviewed Cancer Research Program (W81XWH-18-1-0404), the Kidney Precision Medicine Project (KMPM) Glue Grant, the Ohio Third Frontier Technology Validation Fund, the Clinical and Translational Science Collaborative of Cleveland (UL1TR0002548) from the National Center for Advancing Translational Sciences (NCATS) component of the National Institutes of Health and NIH roadmap for Medical Research, The Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering at Case Western Reserve University. Sponsored research agreements from Bristol Myers-Squibb, Boehringer-Ingelheim, and Astrazeneca.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the U.S. Department of Veterans Affairs, the Department of Defense, or the United States Government.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi: 10.1016/j.ebiom.2021.103481.

References

[1] Arriagada R, Dunant A, Pignon JP, et al. Long-term results of the international adjuvant lung cancer trial evaluating adjuvant Cisplatin-based chemotherapy in resected lung cancer. J Clin Oncol 2010;28:35–42.
[2] Arriagada R, Bergman B, Dunant A, et al. Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. N Engl J Med 2004;350:351–60.
[3] Winton T, Livingston R, Johnson D, et al. Vinorelbine plus cisplatin vs. observation in resected non-small-cell lung cancer. N Engl J Med 2005;352:2589–97.
[4] Corredor G, Wang X, Zhou Y, et al. Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non-small-cell lung cancer. Clin Cancer Res 2019;25:1526–34.
[5] Lu C, Wang X, Prasanna P, et al. Feature driven local cell graph (FeDeG): predicting overall survival in early stage lung cancer. In: Proceedings of the 21st international conference medical image computing and computer assisted intervention. Springer Verlag; 2018. p. 407–16.
[6] Wang X, Janowczyk A, Zhou Y, et al. Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images. Sci Rep 2017;7. doi: 10.1038/s41598-017-13773-7.
[7] Yu RH, Zhang C, Berry GJ, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun 2016;7:12474.
[8] Pignon JP, Tribodet H, Scagliotti GV, et al. Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE collaborative group. J Clin Oncol 2008;26:3552–9.
[9] Scagliotti GV, Fossati R, Torri V, et al. Randomized study of adjuvant chemother-apy for completely resected stage I, II, or IIIA non-small-cell lung cancer. J Natl Cancer Inst 2003;95:1453–61.
[10] Waller D, Poake MD, Stephens RJ, et al. Chemotherapy for patients with non-small cell lung cancer: the surgical setting of the big lung trial. Eur J Cardiothorac Surg 2004;26:173–82.
[11] Ettinger DS, Asinier DL, Wood DE, et al. NCCN guidelines insights: non-small cell lung cancer. Version 3.2018. J Natl Compr Canc Netw 2018;16:807–21.
[12] Kelsey CR, Marks LB, Hollis D, et al. Local recurrence after surgery for early stage lung cancer: an 11-year experience with 975 patients. Cancer 2009;115:5218–27.
[13] Neshbor JC, Putnam JB, Walsh GL, Roth JA, Mountain CF. Survival in early-stage non-small cell lung cancer. Ann Thorac Surg 1995;60:466–72.
[14] Wisnivesky JP, Henschke C, McGinn T, Iannuzzi MC. Prognosis of stage II non-small cell lung cancer according to tumor and nodal status at diagnosis. Lung Cancer 2005;49:181–6.
[15] Bueno R, Hughes E, Wagner S, et al. Validation of a molecular and pathological model of five-year mortality risk in patients with early stage adenocarcinoma. J Thorac Oncol 2015;10:67–73.
[16] Custodio AB, Gonzalez-Larriba JL, Bobokova J, et al. Prognostic and predictive markers of benefit from adjuvant chemotherapy in early-stage non-small cell lung cancer. J Thorac Oncol 2009;4:891–910.
[17] Kim MH, Kim YK, Shin DH, et al. Yes associated protein is a poor prognostic factor in well-differentiated lung adenocarcinoma. Int J Clin Exp Pathol 2015;8:15933–9.
[18] Roepman P, Jassim J, Smitt EF, et al. An immune response enriched 72-gene prognostic profile for early-stage non-small-cell lung cancer. Clin Cancer Res 2009;15:284–90.
[19] Zhu CQ, Ding R, Strumf D, et al. Prognostic and predictive gene signature for adjuvant chemotherapy in resected non-small-cell lung cancer. J Clin Oncol 2010;28:4417–24.
[20] Chang JTH, Lee YM, Huang RS. The impact of the cancer genome atlas on lung cancer. Transl Res 2015;166:568–85.
[21] Canin Y, Ustunova E, Ajakan H, et al. Domain-adversarial training of neural net-works. In: Curcuga K, editor. Domain adaptation in computer vision applications. Cham: Springer International Publishing; 2017. p. 189–209.
[22] Janowczyk A, Doyle S, Gilmore H, Madabhushi A. A resolution adaptive deep hierarchical (RADHistGAL) learning scheme applied to nuclear segmentation of digital pathology images. Comput Methods Biomech Biomed Eng Imaging Vis 2016;1:7–.
[23] Lu C, Romo-Bucheli D, Wang X, et al. Nuclear shape and orientation features from H&E images predict survival in early-stage estrogen receptor-positive breast cancers. Lab Investig 2018 published online June 29. doi: 10.1038/s41374-018-0095-7.
[24] Lee C, Ab S, Veltri R, Epstein J, Christadhas C, Madabhushi A. Cell orientation entropy (COrE): predicting biochemical recurrence from prostate cancer tissue microarrays. Med Image Comput Comput Assist Interv 2013:16:396–403.
[25] Leo P, Lee G, Shih NNC, Elliott R, Feldman MD, Madabhushi A. Evaluating stability of histomorphometric features across scanner and staining variations: prostate cancer diagnosis from whole slide images. J Med Imaging (Bellingham) 2016;3:047502.
[26] Haralick RM. Statistical and structural approaches to texture. Proc IEEE 1979;67:798–804.
[27] Wu Y. Elastic net for Cox’s proportional hazards model with a solution path algo-rithm. Stat Sin 2012;22:279–294.
[28] Chen CH, George SL. The bootstrap and identification of prognostic factors via Cox’s proportional hazards regression model. Stat Med 1985;4:39–46.
[29] Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med 1990;9:811–8.
[30] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 1995;57:289–300.
[31] Pisters KMW, Evans WK, Azzoli CG, et al. Cancer care ontario and american soci-ety of clinical oncology adjuvant chemotherapy and adjuvant radiation therapy for stages I-IIIA resectable non-small-cell lung cancer guideline. J Clin Oncol 2007;25:5006–18.
[32] Schneider PM, Praeuer HW, Stoecklin G, et al. Multiple molecular marker testing (p53, C-erbB-2) improves estimation of prognosis in potentially curative resected non-small cell lung cancer. Br J Cancer 2000;83:473–9.
[33] Endoh H, Tomida S, Yatabe Y, et al. Prognostic model of pulmonary adenocarcinoma by expression profiling of eight genes as determined by quantita-tive real-time reverse transcription polymerase chain reaction. J Clin Oncol 2004;22:811–9.
[34] Lu C, Koyuncu C, Corredor G, et al. Feature-driven local cell graph (FlocK): new computational pathology-based descriptors for prognosis of lung cancer and HPV status of oropharyngeal cancers. Med Image Anal 2021;68:101903.
[35] Dettberracher FC. The eighth edition TNM stage classification for lung cancer: What does it mean on main street? J Thorac Cardiovasc Surg 2018;155:356–9.
[36] Kim Y-G, Kim S, Cho CE, et al. Effectiveness of transfer learning for enhancing tumor classification with a convolutional neural network on frozen sections. Sci Rep 2020;10:21899.
[37] Mehrotra R, Ansari MA, Agrawal R, Anand RS. A transfer learning approach for AI-based classification of brain tumors. Mach Learn Appl 2020;2:100003.
[38] Qaiser T, Tsang Y-W, Taniyama D, et al. Fast and accurate tumor segmentation of histology images using persistent homology and deep convolutional features. Med Image Anal 2019;55:1–14.
[39] Roy M, Kong J, Kashyap S, et al. Convolutional autoencoder based model HistCAE for segmentation of viable tumor regions in liver whole-slide images. Sci Rep 2021;11:139.
[40] Wang X, Chen H, Gan C, et al. Weakly supervised deep learning for whole slide lung cancer image analysis. IEEE Trans Cybern 2020;50:3950–62.