**POSTER PRESENTATIONS 3**

**3B: DIGITAL PATHOLOGY**

**Presenter’s Name:** Coats, Jennifer

**Additional Author(s):** Dammak S, Lin SXJ, Ward A, Cecchini MJ

**Abstract Title:** Analysis of tumour cellularity variance in simulated core needle biopsy specimens across resected lung cancer cases.

**Abstract:**

**Introduction:** Core needle biopsies (CNB) are routinely used for the diagnosis of lung cancer. Imaging modalities, such as magnetic resonance imaging (MRI), ultrasound, and computed tomography (CT) scans, are used to guide the CNB into the target lesion. Currently, CNB are frequently targeted at the center of the tumour unless there is a necrotic core. We hypothesize that the center of the tumour is not always the most cellular given the observed variance in tumours seen in clinical practice. With increasing numbers of molecular tests performed on these small biopsies, it is critical that the tumour cellularity is optimized to ensure there is sufficient material for molecular testing.

**Methods:** Digital slides were randomly selected from the TCGA LUAD (lung adenocarcinoma) n=100 and LUSC (lung squamous cell carcinoma) n=50 datasets. The area of tumour was annotated and a total cell detection was performed in QuPath. An object classifier was then introduced based on annotations (reviewed by a thoracic pathologist) to distinguish tumour cells from surrounding inflammatory and stromal cells. We then simulated ideal core needle biopsies (measuring 0.25 mm x 2.5 mm) and tiled these across the entire slide. The number of tumour cells, total cells and tumour cellularity was recorded within each simulated core. The percent tumour cellularity and the number of tumour cells was then mapped back across the tumour and visualized with density maps.

**Results:** In the squamous cell carcinoma cases, the percent tumour cellularity was highest in the central region in 25% of the cases, the intermediate region in 48% of the cases, and the peripheral region in 27% of the cases. Similarly, in the adenocarcinoma cases, the percent tumour cellularity was highest in the central region in 19% of the cases, the intermediate region in 51% of the cases, and the peripheral region in 30% of the cases.

**Discussion:** The majority of the time the most cellular cases were in the intermediate area between the center and the periphery of the tumour. However, there was a spectrum of findings with some cases having the highest cellularity at the periphery and lower cellularity in other areas. This highlights the need for future studies to correlate these findings with imaging data to develop more sophisticated algorithms to better guide the targeting of core needle biopsy sampling.

---

**POSTER PRESENTATIONS 3**

**3B: DIGITAL PATHOLOGY**

**Presenter’s Name:** Dammak, Salma

**Additional Author(s):** Cecchini MJ, Ward AD

**Abstract Title:** Predicting Tumor Mutational Burden from H&E Slides of Lung Squamous Cell Carcinoma

**Abstract:**

**Introduction:** The PD-L1 score is used to guide treatment decisions, but it does not accurately predict response to immunotherapy in all cases. While adding tumor mutational burden (TMB) to PD-L1 improves response prediction, it is costly and typically requires high tumor cellularity. Previous studies have demonstrated that genomic information such as driver mutations, are encoded in the morphologic appearance of cancer cells1. However, there are not currently any established means to visually distinguish them. In this study, we hypothesized that a neural network could be trained to distinguish the two based on digitized standard-of-care H&E slides.

**Methods:** We utilized digital slides of tumour resections from the Cancer Genome Atlas lung squamous cell carcinoma. The dataset had 50 patients across 30 centers which we split into 30 training and 20 testing slides, each from a unique set of centers. We calculated the TMB with a 10 mutations/Mb threshold utilized to separate TMB-High and Low cases. We explored various different model parameters using the training set only, then fixed the model and tested it on the test set.

**Results:** The selected model is VGG162, trained using the technique of transfer learning, and had an area under the receiver operating characteristic curve of 0.65, accuracy of 65%, sensitivity of 77% and specificity of 43%.

**Discussion:** This study suggests that complex genetic features of a tumor are encoded in the morphologic appearance on H&E slides, and this is the first study that shows that TMB specifically can be detected in the morphology of squamous cell carcinoma on H&E slides across multiple centers. Further, it shows that even with a small training set, it is possible for a neural-network-based model to detect this relationship. This motivates additional work in this direction to build a system that can be used in the future to help physicians decide which patients with squamous lung carcinoma would benefit from immunotherapy.

Abstract: How tall are you? Using digital pathology to automatically quantify and characterize the tall cell variant of papillary thyroid carcinoma.

Methods: Cases of papillary carcinoma (n=20) were randomly selected from the Cancer Genome Atlas (TCGA) PTC database, including TCV (n=10) and non-TCV (n=10). All cases were reviewed by a Head and Neck pathologist (MJC). Cell detection was performed using QuPath with parameters optimizing cell expansion and nuclear fragmentation. Representative tumour and non-tumour areas were used as training data for an object classifier. Calliper values were taken from detected cell parameters to estimate and stratify cell measurements by length-to-width ratio. Calliper-generated ratios were then compared to manual cell measurements to determine the efficacy of the automated approach.

Results: The object classifier accurately identified 80.2% of cells that had a length-to-width ratio of 2:1 or greater within a 5% margin of error. There was a 14.7% difference between the mean ratios of the manually classified cells and the automatically detected cells. When the automated system was applied to 10 TCV and 10 non-TCV PTC cases, there was a trend towards an increased number of tall cells calculated in cases of TCV PTC.

Discussion: We were able to demonstrate that TCV could be distinguished from non-TCV with moderate accuracy using automated cell detection software. However, the object classifier was only tested against a single representative slide in each case of PTC. Future work remains to be completed with validation against all representative tissue in each case, and further refinement of cell border detection before we can validate the established WHO-defined diagnostic criteria.
Abstract:

Introduction: Lung cancer is staged based on the size of the tumor and involvement of other structures. This staging may be a surrogate measure for the number of cells present in the tumor. The recently updated grading system for lung adenocarcinoma assesses the presence of high risk architectural patterns, which tend to have more complex cellular growth. Counting individual tumor cells is impractical for a pathologist using a conventional light microscope. Image analysis tools applied to digital slides can be utilized to automate the quantification of lung adenocarcinoma. We hypothesize that tumor cellularity can be used as a novel prognostic tool in lung cancer that integrates quantification of high risk architectural patterns.

Methods: Digital slides (n=102) from the Cancer Genome Atlas (TCGA) lung adenocarcinoma (LUAD) dataset were obtained and analyzed in QuPath. Representative areas of tumor were annotated and reviewed by a thoracic pathologist, the annotations were used as training data for a random trees based object classifier that utilized detected cell features to identify and quantify tumor cells across entire slides. This was normalized with the surface area of the tumor present on the slide to provide a measure of tumor density. The overall total cellularity was calculated by combining the size of the grossly measured tumor with the tumor density. Major histologic patterns in representative panels were determined by a thoracic pathologist and were compared with the tumor density of the tile. The overall and progression free survival was compared between groups of high and low tumor cellularity.

Results: High-grade histologic patterns had a significantly greater tumor density compared with other patterns of lung adenocarcinoma. A trend between survival and cellularity was seen a correlated jump, with an estimated 21-23% increase in workload. Already, CRC gross examination is considered more labour-intensive and time-consuming than other specimens. For proper Tumour-Node-Metastasis (TNM) classification, it is vital to identify all lymph nodes (LNs) in a specimen, which directly affect cancer treatment and prognosis. Identification of LNs during gross assessment is crucial but challenging, as LNs can be small and difficult to find via manual palpation and dissection. In this project, we have developed a device that utilizes ultrasound to detect LNs in resected tissues.

Discussion: Tumor cellularity represents a novel prognostic tool in lung cancer that takes into account both the size and composition of the tumor. Use of advanced image analysis tools allows for the automation of this task in a simplified and efficient manner. Future work will seek to validate these findings in additional larger datasets to refine the classification of tumors by cellularity.
POSTER PRESENTATIONS 3
3B: DIGITAL PATHOLOGY

Presenter’s Name: Pierce, Kevin


Abstract Title: Pathology Annotated Tile-based High-throughput Classification Application (PATHCA) for Lung Cancer Classification

Abstract:

Introduction: In pathology and other fields, one of the greatest barriers to developing machine learning models is the need for large numbers of labelled examples for training. Traditionally, generating a collection of tumour annotations on H&E slides requires highly skilled pathologists and is an extremely time-consuming task. We have built an online collaboration platform designed to collect labelled tilesets to support the training of a supervised machine learning model. Using this approach, the task of labelling slides can be reduced to a simple image recognition task, requiring minimal supervision to generate labelled training data.

Methods: 11 undergraduate, 1 graduate student and 1 medical student without formal pathology training labelled tiles in the application. 320 tiles of dimension 570x570 μm² were fragmented from contours of adenocarcinoma cases from the TCGA diagnostic lung adenocarcinoma dataset. Fragmented tiles were isolated in identical batches of 16 and displayed in a 4x4 grid within the application; participants were instructed to identify tiles that were positive for tumour. Each user’s selection or non-selection was used to generate a group consensus on whether or not the tile contained tumour. Each user’s accuracy was then evaluated with respect to the consensus. The same 320 tiles were given to a thoracic pathologist, and both the consensus and each individual were scored against the pathologist’s selections.

Results: Tiles with high agreement showed the presence or absence of tumour cell nests with varying degrees of eosinophilic cytoplasm and nuclear pleomorphism whereas tiles with split consensus potentiate scenarios of normal goblet cells, single isolated tumour cells, or presence of artefacts, all showing tumour-like characteristics. The cases with high consensus showed good accuracy (84.4%) with respect to the pathologist. Cases of discrepancy between the consensus typically involved focal involvement of the tile by tumour. This is a highly efficient means for labelling with participants able to quickly identify tumour (range: 1.7 - 32.3; average 9.1 seconds per tile).

Discussion: We demonstrate that a clinically-oriented collaborative tool can simplify the identification of tumours into an image recognition task performed by non-pathologist observers to generate labelled training data. We anticipate this application can remove barriers in generating labelled tilesets for machine learning models in pathology.

---

POSTER PRESENTATIONS 3
3B: DIGITAL PATHOLOGY

Presenter’s Name: Skeba, Danielle

Additional Author(s): Asfaha S, Shin AE, Cecchini MJ

Abstract Title: Digital Characterization of Bowel Damage in Chemical- and Bacterial-Induced Experimental Colitis in Mice

Abstract:

Crohn’s disease and ulcerative colitis comprise inflammatory bowel disease (IBD), an idiopathic disease with genetic and environmental influences. IBD is becoming increasingly prevalent in the global population. To facilitate mechanistic studies of IBD, the colitis phenotype is often induced in mice using orally administered dextran sodium sulfate (DSS). Other models of colitis are achieved with the use of oxazolone, TNBS, Citrobacter rodentium and doxorubicin. Currently, researchers working with histologic colitis samples from these models manually identify distinguishable features of each model and quantify areas of disease activity. Manual analysis introduces inter- and intra-observer variability, as well as considerable time invested. More powerful, reliable, and refined analytical capability may be achieved using digital pathology tools. This study utilizes QuPath, an open-source whole-slide image analysis program, in conjunction with CytoMAP, a built-in MatLab tool for tissue spatial analysis, to compare and characterize the different models of colitis in mice. This method, leveraging both the cell detection and classification features of QuPath, as well as the streamlined clustering analysis pipeline of CytoMap, will extract features from each of the models of colitis. The findings of this study may contribute to the creation of a more efficient and reliable method for identifying colitis in mice. In addition, this study will contribute to research into the emerging field of digital pathology as a tool in research and potentially the clinical assessment of IBD.
**Abstract:**

Shades of Hematoxylin and Eosin: Digital Image Analysis of Stain Variability

**Methods:** Tissue control blocks containing hepatic, renal cortex, small and large intestinal tissue sections were serial sections and stained utilizing a routine hematoxylin and eosin staining procedure. A total of 40 slides were included in this study, consecutively scanned on the Aperio AT. QuPath was utilized for digital analysis by colour deconvolution and optical density. Prism graphing software was utilized for statistical analyses and visualization of the data obtained.

**Results:** Data represented graphically allowed for dramatic coefficients of variation to be better understood when contrasted with variation occurring in other tissue types. Data obtained for the optical density of tissue sections followed expected patterns with hematoxylin having higher mean density compared to eosin staining. All nonnuclear deconvoluted hematoxylin results were excluded. Deconvoluted eosin staining results showed to be consistently problematic across all cell types analyzed.

**Discussion:** Images of minimum and maximum optical density sections were included for side-by-side comparison and visualization of the span of variation that is occurring. Through analysis of both deconvoluted hematoxylin and eosin, over-staining with eosin as observed in the maximum optical density sections proved to be extremely problematic for interpretation. Further investigation of the staining procedures and the effects of the reagent change schedule would help eliminate occurrences of over-staining and over differentiation.