DISTRIBUTION AGREEMENT

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Jordan S. Goldstein Date

Utilizing Digital Pathology Informatics Algorithms for Diffuse Large B-cell Lymphoma Subtyping

By

Jordan S. Goldstein

Master of Science

Clinical Research

_____

Christopher R. Flowers, MD, MS
Advisor

_____

Mitchel Klein, PhD
Committee Member

_____

Muna Qayed, MD, MSc
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____

Date

Utilizing Digital Pathology Informatics Algorithms for Diffuse Large B-cell Lymphoma Subtyping

By

Jordan S. Goldstein

BS, Cornell University, 2013

Advisor: Christopher R. Flowers, MD, MS

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Clinical Research
2018

ABSTRACT

Introduction:
Through gene expression profiling (GEP), two biologically distinct molecular subgroups of diffuse large B-cell lymphoma (DLBCL) have been identified – germinal center B-cell (GCB) and activated B-cell – with disparate prognostic implications. Because GEP is not easily implemented in clinical practice, immunohistochemical (IHC) algorithms have been developed and validated as surrogates for DLBCL subtype. The most commonly employed algorithm uses CD10, BCL6 and MUM1 to distinguish GCB and non-GCB subtypes, but has been shown to misclassify up to 20% of cases, partially due to the significant variability that exists in IHC interpretation. Advancements in digital pathology technology have led to the development of image analysis algorithms that allow for the extraction of quantitative descriptions of pathology features such as level of IHC staining. We explore the feasibility of developing precise, objective image analysis-based scoring algorithm for DLBCL digital pathology whole slide images that can effectively distinguish between GCB and non-GCB subtype.

Methods:
Tissue slides of the immunostains CD10, BCL6, and MUM1 for 40 DLBCL patients were digitized at 40x objective resolution. We developed and trained an image analysis algorithm on the IHC whole slide images that returns the percentage of positive regions over the DLBCL tissue area. Receiver operating characteristics (ROC) curves were calculated to assess ability of image analysis-based measurements of region positivity to predict pathologist classifications as positive or negative for CD10, BCL6, and MUM1. Using thresholds for percent positive regions established by the ROC curves, the patients were classified into GCB/non-GCB by the Hans algorithm using sequential application of the identified thresholds, and concordance with pathologist classification was calculated.

Results:
Area under the ROC curve (AUC) for predicting pathologist classifications of positivity from image analysis measurements for CD10, BCL6, and MUM1 were 0.92, 1.0, and 0.95 respectively. Thresholds from the ROC curves for CD10, BCL6, and MUM1 were 13%, 15%, and 20% respectively. Using these thresholds, classification by image analysis algorithm was concordant with pathologist classification in 82.5% ($\kappa = 0.65$) of cases.

Conclusion:
The image analysis algorithm could provide an effective support tool for pathologists, improving the IHC classification of DLBCL subtype.

Utilizing Digital Pathology Informatics Algorithms for Diffuse Large B-cell Lymphoma Subtyping

By

Jordan S. Goldstein

BS, Cornell University, 2013

Advisor: Christopher R. Flowers, MD, MS

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Clinical Research
2018

ACKNOWLEDGEMENTS

I'd like to thank the following people for their tremendous help with the preparation of this thesis:

Christopher R. Flowers, MD, MS

Lee A.D. Cooper, PhD

Sanghoon Lee, PhD

David L. Jaye, MD

Jacob Jordan

TABLE OF CONTENTS

INTRODUCTION

Diffuse large B-cell lymphoma (DLBCL), the most commonly occurring lymphoid malignancy in the United States, is an aggressive disease characterized by heterogenous outcomes. Through gene expression profiling (GEP), two biologically distinct molecular subgroups of DLBCL have been identified – germinal center B-cell (GCB) and activated B-cell – with disparate prognostic implications.(1) Due to its cost and resource intensiveness, GEP has not yet been implemented in clinical practice. As such, immunohistochemical (IHC) classification systems have been developed and validated as GEP surrogates to segment DLBCL into its distinct cell-of-origin subtypes.(2-4) The most commonly employed classification system uses CD10, BCL6 and MUM1 to distinguish GCB and non-GCB subtypes, but has been shown to misclassify up to 20% of cases.(2) Studies have shown significant variability exists in the interpretation of IHC stains in DLBCL, likely accounting for some of the misclassification.(5)

Image analysis in digital pathology presents the opportunity to decrease the variance in IHC interpretation and reduce subtype misclassification in DLBCL. Advancements in digital pathology technology have led to the development of image analysis algorithms that allow for the extraction of quantitative descriptions of pathology features such as level of IHC staining across whole slide images (WSI). Recent studies in other malignancies have shown that image analysis of IHC digital pathology slides can be accurate, be efficient, and reduce variability.(6-8)

We sought to develop an objective and precise IHC image analysis algorithm using DLBCL digital pathology WSI that can accurately distinguish between GCB and non-GCB subtype using the IHC stains in the Hans classification system (CD10, BCL6, MUM1).

BACKGROUND

Non-Hodgkin lymphoma (NHL) is the most prevalent hematologic malignancy in the United States, and account for 4% of all malignancies in incidence and death. DLBCL is the most common form of NHL, comprising about 30-35% of cases, with an estimated 27,650 new diagnoses in 2016.(9) DLBCL is an aggressive disease with an untreated survival on the order of just months.(10) Fortunately, chemotherapeutic regimens including cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) have allowed for a cure in about 50% of patients.(11) Since the CHOP regimen came out in 1976, only the addition of the immunotherapy rituximab to the CHOP regimen (R-CHOP) has been shown to improve DLBCL survival.(12-15) Despite the effectiveness of the modern chemoimmunotherapy regimen, outcomes for DLBCL remain heterogenous, with a sizable portion of DLBCL patient who progressing early or experiencing relapse, and about 40% still dying from their disease.(12, 16-18)

Given the variability in outcomes, research has focused on studying the underlying biological diversity of DLBCL. In 2000, gene expression profiling (GEP) was used to identify two molecularly distinct subgroups of DLBCL based on their cell-of-origin: one that clusters with normal germinal center B-cells (GCB) and one that clusters with activated B-cell (ABC).(1) Despite their identical histologic appearance, the ABC subtype of DLBCL was shown to have a significantly worse than the GCB subtype. Although GEP is widely used in bench research, it has not been readily adopted into clinical practice, and only recently into clinical trials, as it is expensive and requires abundant fresh-frozen paraffin embedded tissue (FFPET). A recent study on early-stage breast cancer found a

substantially higher cost-effectiveness ratio for GEP use in community practice relative to ideal conditions.(19)

To molecularly stratify the cell-of-origin subtypes for DLBCL in the clinical setting, immunohistochemical (IHC) classification systems have been developed as surrogates for GEP. The first such classification system was developed in 2004 by Hans et al. and used the differential expression of the proteins CD10, BCL6, and MUM1 to separate GCB and non-GCB phenotypes.(2) CD10 is a membrane-associated, neutral endopeptidase that has restricted expression in activated germinal center cells; BCL6 is a zinc-finger protein that acts as transcription repressor and is expressed in germinal center B cells; and MUM1 is a transcription factor expressed in a subset of germinal center cells. When sequentially combined in a classification system (See Figure 1), these IHC stains are able to stratify GCB and non-GCB DLBCLs with a sensitivity of 70% and 87% and positive predictive value of 84% and 75% for the GCB and non-GCB groups, when compared to GEP classification, respectively.(2) The Hans classification system resulted in an 86% concordance with GEP, and therefore misclassifies 14% of patients.(2) Eight other IHC classification systems have subsequently been developed that reported higher concordance with molecular-based classification of DLBCL, though none have been as readily used in clinical practice as the Hans classification system likely due to its simplicity and historical use. (3, 4, 20-24)

Although IHC staining have been a mainstay in the determination of oncologic diagnosis and prognosis, pathologist IHC interpretation is fraught with limitations. In IHC staining, the “brown” diaminobenzidine (DAB) chromogenic stain represent positive protein expression and provides a contrast to the “purple” hematoxylin counterstain. One

issue is that the brown DAB stain has a narrow dynamic range relative to the purple hematoxylin stain, meaning it is difficult for the human eye to distinguish the brown "signal" from the purple "noise".(25) Another limitation is the subjective interpretation of protein expression levels by pathologists, generally resulting in ordinal outputs and leaving the potential for a pathologist artifact.(26) Further, interpretation of IHC staining by pathologists has been shown to be subject to significant interlaboratory and interobserver variability.(5, 27, 28) This has particularly been the case when pathologists are asked to estimate percentages of stained areas (28), which is what is required in DLBCL.

Studies examining interlaboratory and interobserver variation in interpretation of the various IHC biomarker stains for DLBCL have produced uninspiring results. De Jong et al. found interrater agreement among 24 pathologists examining DLBCL tissue microarray for CD10, BCL6, and MUM1 to be just 65% ($\kappa = 0.39$), 34% ($\kappa = 0.17$), and 34% ($\kappa = 0.16$) respectively.(5) This resulted in an interrater agreement for subtype by the Hans classification system of 57% ($\kappa = 0.36$). Pathologist reliability improved somewhat when optimizing staining technique, controlling for interlaboratory variability, to 70% ($\kappa = 0.72$), 53% ($\kappa = 0.42$), and 54% ($\kappa = 0.41$) for CD10, BCL6, and MUM1 respectively and to 77% ($\kappa = 0.62$) for subtype classification.(5) Based on this data, there remains a need for the reduction of IHC interobserver variability in practice, improving accuracy and utility of the IHC classification systems.

Significant advancements in digital pathology technology, such as scanning and storage, have allowed for the full digitalization of the stained tissue sections. The digital whole slide images (WSI) are able to capture histopathological details in high resolution, providing a resource rich in data, in addition to the opportunities for telepathologic

diagnosis and education. Computational imaging of WSI allow for a quantitative description of disease pathology and create compelling opportunities for precision medicine. Histology image analysis algorithms have received significant attention and been developed to predict metastasis (29), survival (30-34), grade (35), and classification (36-38), as well as linked to genomic and biologic data (39-41).

Several studies looking at computer-aided diagnosis have focused on using image analysis algorithms for interpretation of IHC stains. Estrogen receptor- (ER), progesterone receptor- (PR), and her2 receptor-positive status in breast cancer can be accurately determined by quantitative image analysis algorithm.(8, 42-45) When IHC automated measurements are provided to pathologists as a diagnostic aid, inter- and intra-observer agreement improves substantially.(6, 7) Meanwhile, automated her2 IHC scores can provide more reliable data, closely estimating consensus visual scores by multiple expert pathologists.(46) In 2016, the FDA approved the use of quantitative IHC interpretation in breast cancer for determination of ER/ PR status. Automated IHC analysis algorithms have been shown to be highly effective in quantifying Ki67 prognostic marker in breast cancer (47), S100A in ovarian cancer (48), and epidermal growth factor receptor in colon cancer (49).

In April 2017, the US FDA approved the use of digital pathology for primary diagnosis of surgical pathology slides. The FDA reviewed data on 2,000 surgical pathology cases and found that the diagnoses generated from the digital pathology images was noninferior to optical microscopy. This signifies a major step in the adoption of digital pathology scanners nationwide. As the use of digital pathology scanners continues to

expand, computer-aided diagnosis through image analysis algorithms will likely develop into an essential resource in clinical oncologic practice.

The promise of quantitative IHC interpretation of digital pathology images may offer a solution to the high rate of DLBCL subtype misclassification, which is partially a result of limitations in IHC stain interpretation. We set out to develop an objective and precise quantitative IHC image analysis algorithm using digital pathology WSI that can accurately predict pathologist classification as positive or negative for each IHC stain included in the Hans classification system (CD10, BCL6, MUM1), and can effectively distinguish between GCB and non-GCB subtype. Such an algorithm would reduce inter-rater reliability and establish a pathway toward improving correlations between IHC classification and GEP.

METHODS

Aims

Aim 1. Develop and apply IHC image analysis algorithm to DLBCL digital pathology WSI for IHC stains CD10, BCL6, and MUM1.

Aim 2. Assess the ability of the image analysis algorithm to discriminate between pathologist classification of positive or negative for each IHC stains CD10, BCL6, and MUM1.

Aim 3. Estimate concordance of image analysis algorithm-generated DLBCL subtype classification with pathologist generated DLBCL subtype classification.

Data Source

DLBCL patients were retrospectively identified from a database of Emory University DLBCL cases who had available whole slide images for each of the stains included in the Hans classification system (CD10, BCL6, MUM1). Each patient had been seen by a physician at the Winship Cancer Institute at Emory University and had their slides read by subspecialized hematopathologists from Emory University Hospital. Pathology classification by the hematopathologists for each IHC stain CD10, BCL6 and MUM1 as positive or negative and for DLBCL subtype was recorded from the pathology reports. The whole slides were de-identified and digitized at 40x objective resolution (specimen-level pixel size 0.226 x 0.226) using 2 different scanners: Hamamatsu Nanozoomer 2.0 HT or Aperio AT2.

Image Analysis Algorithm

An image analysis algorithm was developed and trained on the 40 IHC whole slide images (WSI) for each IHC stain: CD10, BCL6, and MUM1 (See Figure 2A for example

of WSI). The image analysis pipeline used consists of three steps: color normalization, segmentation, and feature extraction.

First, each of 40 IHC whole slide images were independently normalized to a standard IHC stained image with desired Hematoxylin and DAB characteristics using Reinhard normalization (Figure 2B). This normalization process uses a linear transform in a perceptual color-space to match the means and standard deviation of the two color channels (brown and purple) between the whole slide image and the standard image, improving consistency of the subsequent segmentation and feature extraction.

Second, the tissue pixels in the normalized images were masked from the background using linear discriminant analysis to ensure only cellular tissue is included. The masked tissue area was then segmented into regions through superpixelation, which creates compact regions (60 micron squares) based on image content and location, so that each region contains similar amount cellular material while remaining roughly equivalent in size (Figure 2C). The normalized regions have a color deconvolution algorithm applied, separating the purple hematoxylin stain from the brown DAB stain based on pixel thresholds (Figure 2D).

Last, the features describing the distribution of both the Hematoxylin and DAB pixel stains were extracted using a convolutional autoencoder, which is a form of unsupervised learning that identifies visual features through dimensionality reduction. A 28x28 image patch was encoded into a set of 64 features for each superpixel. The encoded features for the 40 IHC whole slide images were trained and classified using a random forest algorithm on active machine learning.

Active machine learning is an iterative process that first learns an unknown decision boundary that classifies regions as positive or negative based on the region stain features (Figure 2E). Regions of low confidence, those that are difficult to classify as positive/negative, are presented to the user for labeling. The algorithm is then retrained with the new labels, correcting prediction errors of low confidence regions and improving classification, and applied to the entire dataset. The active machine learning process was performed over several iterations. The algorithm classifies all the tissue regions of the WSI as positive or negative and outputs a percentage of positive regions. An example heatmap showing the positive class density determined by the image analysis algorithm of a MUM1 WSI is shown in Figure 3.

The image analysis algorithm uses tools from the HistomicsTK Python library for histologic analysis (http://github.com/DigitalSlideArchive/HistomicsTK) and HistomicsML pipeline for Active machine learning.

Statistical Analysis

The image analysis algorithm was applied to each of the WSI for the 40 DLBCL cases returning the percentage of positive regions across the tissue for each IHC stain. Receiver Operating Characteristic (ROC) curves were generated for each IHC stain to assess the ability of percent positive regions to discriminate between pathologist classification as positive or negative for CD10, BCL6, and MUM1. Area under the curves (AUC) and ideal classification thresholds for percent positive regions were calculated. Using these thresholds, DLBCL cases were classified into GCB/non-GCB by the Hans classification system using sequential application of the identified thresholds on the percentage of positive regions determined by the image analysis algorithm. Concordance

of image analysis subtype classification with pathologist subtype classification was calculated by agreement rate and Cohen's kappa statistic. The Cohen's kappa statistic is a measure of interrater reliability that accounts for chance agreement. A random forest classifier was trained to use the percentage of positive regions output by the image analysis algorithm, as the features, to classify the DLBCL subtypes, using pathologist subtype classification as the training labels. Concordance of the random forest classifier using the image analysis algorithm outputs with pathologist subtype classification was calculated. Statistical analyses were performed using Python libraries SciPy and scikit-learn.

RESULTS

Forty patients with DLBCL had staining for all IHC stains in the Hans classification system (CD10, BCL6, and MUM1). All patients were diagnosed between 2006 and 2016 at a variety of hospitals around the Southeastern United States, where their histology slides were prepared. Pathologists classified CD10, BCL6, and MUM1 as positive in 27.5%, 90%, and 84.9% respectively. Pathologists classified 15 patients as GCB and 25 patients as non-GCB. Patient characteristics are shown in Table 1. No differences in patient characteristics were found between GCB and non-GCB DLBCL subtypes.

The image analysis algorithm was applied to each the 40 DLBCL cases, returning the percentage of positive regions across the tissue for each IHC stain. When examining the ability of the percent positive regions generated by the image analysis algorithms to discriminate between pathologist classification as positive or negative, the area under the ROC curves for CD10, BCL6 and MUM1 were 0.92, 1.0, and 0.95 respectively (see Figure 4, 5, and 6 respectively). The optimal decision threshold calculated from the ROC curves, or the percentage of positive regions at which sensitivity and specificity for distinguishing pathologist classification as positive or negative are maximized, for CD10, BCL6, and MUM1 were 13%, 15%, and 20% respectively. Using these decision thresholds, sensitivity and specificity for CD10, BCL6, and MUM1 were 82% and 90%, 100% and 100%, and 94% and 83% respectively.

Sequential application of these thresholds in the Hans classification system was used to generate subtype classification by image analysis algorithm is shown in Figure 7. Of the 40 DLBCL cases, 13 had percent positive regions for CD10 greater than the 13% optimal threshold and were therefore classified as GCB by the image analysis algorithm.

Of these 13 cases identified as GCB by the image analysis algorithm, 11 (85%) were consistent with pathologist classification as GCB. Of the 27 remaining DLBCL cases (those with CD10 percent positive regions < 13%), 4 had percent positive regions for BCL6 less than the 15% threshold and were classified as non-GCB by the image analysis algorithm. All 4 (100%) of these cases were consistent with pathologist classification as non-GCB. Of the 23 remaining DLBCL cases (those with percent positive regions for CD10 < 13% and for BCL6 > 15%), 17 had percent positive regions for MUM1 greater than the 20% threshold and were diagnosed as non-GCB. Of these 17 cases, 16 (94%) were consistent with pathologist classification as non-GCB. The remaining 6 cases with percent positive regions for CD10 < 13%, BCL6 > 15% and MUM1 < 20% were classified as GCB by the image analysis algorithm. Of these 6, 3 (50%) were consistent with pathologist classification as GCB. Given the image analysis algorithm and pathologist subtype agreed on 34 of the 40 cases, concordance between pathologist subtype classification and subtype classification by image analysis algorithm was 85% ($\kappa = 0.70$).

After training a random forest classifier using the percent positive regions from the image analysis algorithm for CD10, BCL6, and MUM1, the subtype from the image analysis algorithm agreed with the pathologist subtype on 33/40 DLBCL cases. Therefore, concordance between pathologist classification and subtype classification from the random forest classifier was 82.5% ($\kappa = 0.63$).

DISCUSSION

Historically, DLBCL patients have been stratified into low-, intermediate-, and high-risk based on the international prognostic index that incorporates clinical factors.(50) While the IPI remains useful clinically, cell-of-origin subtyping has come to prominence as the most clinically useful prognostic stratification as the ABC subtype has worse outcomes regardless of IPI risk level.(1, 2, 4, 51) Utilizing the biologically-based cell-of-origin classification system provides the possibility of more targeted treatment. The ABC subtype of DLBCL has shown a significantly increased activation of NF-kB, prompting a variety of clinical trials in the ABC subtype focusing on therapies affecting this pathway, such as bortezomib (52), lenalidomide (53, 54), and ibrutinib (55). Although cell of origin differentiation by gene expression profiling has prognostic significance, the IHC surrogate classification systems have yielded conflicting results on its ability to predict outcomes.(56-61) A systematic review and meta-analysis found that IHC classification systems can be useful in predicting subtype, but GEP remains the preferred method for predicting prognosis.(62) Further, there is poor concordance in predicting cell-of-origin subtype even between the various IHC classification systems.(24) This could be a result of the significant pathologist variation in IHC stain interpretation, in addition to the imperfect concordance of the classification systems with GEP. An objective, efficient cell-of-origin classification system utilizing IHC staining is needed for current prognostic risk stratification, effective subtyping for clinical trials, and personalization of treatment strategies moving forward.

The image analysis algorithm that we developed represents a novel, objective, quantitative method for classification of DLBCL subtype. The algorithm makes use of

whole slide images and assesses the percentage of positive regions throughout the DLBCL tissue. The algorithm output of percent positive regions is highly predictive of pathologist classification as positive or negative for each IHC stain in the Hans classification system with high AUC values. To our knowledge, this is the first image analysis algorithm showing the ability to effectively discriminate between pathologist classification as positive or negative for IHC stained DLBCL whole slide images. The high agreement rate and kappa-statistic of 0.7 suggest that the image analysis algorithm has high, though imperfect, concordance with pathologist classification as GCB or non-GCB by the Hans classification system. The high concordance with pathologist classification suggests potential as a decision-support tool for pathologists in clinical practice and for efficient use in subtyping for clinical trials. The image analysis algorithm will improve inter-rater reliability. In addition, the algorithm was trained and tested on slides prepared from a variety of laboratories, allowing it to account for inter-laboratory variation. Our future research will focus on investigating the correlation of the algorithm with GEP. It may have an improved correlation, or could be better trained to subtype DLBCL, when compared against the gold-standard GEP.

In current clinical practice, pathologists examine IHC-stained DLBCL whole slides by examining specific representative regions in a high-power field and estimating if >30% of tumor cells are positive. This process introduces significant variation in selection of representative regions and estimation, suffers from a binary classification, and may be unable to include tumor heterogeneity. Our image analysis algorithm's use of the whole tissue and quantitative region positivity output help to address these issues. The quantitative IHC stain positivity renders the possibility of generating new, more specific

positivity cutoff values. In addition, efficient subtype classification of the considerable number of patients included in clinical trials currently relies on the use of tissue microarrays, single slides for a given IHC stain that contain up to 100 DLBCL tissue cores, small regions of DLBCL selected for their perceived representativeness of the tumor. This allows the pathologist to read the many stains efficiently without having to regularly switch slides, which would take significant time, effort, and storage. However, tissue microarrays can be difficult to make, require FFPET, introduce variation in core selection, and fail to capture tumor heterogeneity.(63) In addition, tissue microarrays have produced inconsistent results with regard to subtype prediction based on IHC classification systems.(64) Our image analysis algorithm has the potential to automate this process, using digitized whole slides, and improve classification efficiency and consistency in DLBCL clinical trials.

We compared two different methods to predict DLBCL subtype using the image analysis algorithm. The first uses sequential application of the Hans classification system using the ROC-generated thresholds for the three IHC stains CD10, BCL6, and MUM1. This method (shown in Figure 7) mimics the Hans classification system used in clinical practice. The second method utilized a random forest classification. The random forest classifier is a supervised ensemble machine learning method that generates a multitude of decision trees, much resembling the Hans classification system (Figure 1), and outputs the mode classification. In our case, the decision trees are constructed from the 3 IHC stain image analysis outputs; each decision tree recommends a classification of either GCB or non-GCB; and the mode classification represents the subtype determined by the random forest classifier. Whereas sequential application of the Hans classification system is a rigid

decision tree and requires region positivity thresholds generated from the ROC curves, the random forest classifier provides flexibility in the classification system. Random forest classification does not require set thresholds for IHC region positivity and can easily incorporate other IHC stains into its classification process, such as those for the Choi or Tally classification system, as well as others. This provides a classification system that can be easily adapted for future research purposes. The random forest subtype classification provides a similar, though slightly less, concordance with pathologist classification relative to that of the sequential application of the Hans algorithm. We would expect that the random forest classifier's concordance with pathologist subtype would improve with increased sample size, but our results suggests that the Hans algorithm is the most efficient decision tree.

Independent of the Hans algorithm, positive IHC staining for BCL2 and MYC have been shown to be strong predictors of worse outcomes regardless of cell-of-origin and IPI.(65-68) BCL2 is an anti-apoptotic gene that plays is essential for normal B-cell development and differentiation, while MYC plays important roles in cell-cycle promotion and apoptosis. Those with high expression of both BCL2 and MYC, referred to as "double-hit" lymphomas, have a particularly aggressive disease course, with a 9x increased risk of death relative to that of those with low expression of the stains.(69, 70) Interpretation of the BCL2 IHC stain has been shown to be subject to variability with an agreement rate of just 47% ($\kappa = 0.23$).(5) In addition, the scoring of a positive BCL2 IHC stain has not been universally agreed upon. A recent study focused on developing a new scoring system for BCL2 hoping to improve consistency.(71) Meanwhile, studies looking at interrater reliability of pathologists on scoring whole slides for MYC positivity yielded almost 40%

of discordant cases.(72) This suggests a major opportunity for improving the variability of MYC scoring. The Lymphoma Epidemiology of Outcomes cohort study, a large multi-institution study accumulating in-depth biologic and clinical data on lymphoma patients, currently has a prospective project underway further detailing interrater reliability between 9 expert hematopathologists for both BCL2 and MYC in DLBCL. Our image analysis algorithm will be adjusted to run on these IHC stains, and we will compare the classification by our algorithm with that of the pathologists. These stains have the potential for incorporation into our random forest classification of subtype that uses the image analysis algorithm outputs. Further, including the quantitative output of the image analysis algorithm for BCL2 and MYC will allow for a more personalized prognostic prediction, especially when incorporated with BCL6, CD10, and MUM1.

Follicular lymphoma (FL) is the second most common type of NHL and has similarly suffered from issues with interobserver variability in its grading, with agreement between pathologists ranging from 61% to 73%.(73) However, computer-aided image analysis has sought to address this issue. The grading of FL by pathologists relies upon the number of centroblast cells within a follicle center in high-power fields. However, the selection of high-power fields by pathologists was a major source of variability. The automated selection of high-power fields, through image analysis selection of representative tumor regions, was able to reduce lack of consensus from 41% to just 6%.(74) Image analysis algorithms were able to effectively perform follicle detection(73, 75, 76) and differentiate centroblast cells within the follicles (77-79). Using these informatics tools, Fauzi et al created an automated FL grading system, called FLAGS, that was used as a decision support tool for pathologists.(80) Exposure to FLAGS improved

grading accuracy both expert and resident hematopathologists, and reduced sampling and reader bias.(80) This iterative development of informatics tools for FL grading represents a direct pipeline for the creation of effective decision support tools for pathologist classification of lymphoid malignancies. Our algorithm represents an important first step in improving the subtype classification of DLBCL. Future research will need to focus on improving our algorithm for DLBCL prognostic prediction and adapting the algorithm for use as a decision support tool in the clinical setting.

Other research has gone into adapting DLBCL cell-of-origin subtyping to the bedside. The Lymphoma/Leukemia Molecular Profiling Project developed and validated a digital gene expression-based assay, called Lymph2Cx, that can determine cell-of-origin using fresh frozen paraffin embedded tissue biopsies with accuracy and consistency.(81) The Lymph2Cx assay also maintained the prognostic significance of the ABC and GCB subtypes.(82) However, a cost-effectiveness analysis of the use of this assay has not been performed, and it has not gained any traction for use in clinical practice. Of note, cost-effectiveness studies on breast cancer did not yield promising results for routine use in clinical practice.(19) The need remains for improvement in IHC-based cell-of-origin subtype determination in clinical practice, and digital pathology presents a promising tool to fill the current void.

Our study suffered from a few limitations that present opportunities for future research. Due to the small sample size, no validation set was used to generate the ROC curves assessing the ability of the image analysis algorithm to discriminate between pathologist classification as positive and negative. The Lymphoma Epidemiology of Outcomes cohort study has been accumulating DLBCL tissue slides from several

institutions around the United States with staining for the Hans algorithm. These will be used for a future validation set. Meanwhile, GEP remains the gold standard for classifying DLBCL subtype rather than pathologist classification. While we've shown high concordance of the image analysis algorithm with pathologist subtype classification, it remains necessary to compare subtype classification generated by the image analysis algorithm to that of the gold standard GEP cell-of-origin subtype. The imperfect concordance between the image analysis algorithm and pathologist subtype classification leaves the possibility of a higher concordance of the algorithm with GEP. We have begun the process of obtaining GEP data for DLBCL tissue that also has the proper staining for the Hans classification system to perform this study. In addition, although the slides all were classified by expert hematopathologists, some variability in classification was introduced by using reports from multiple pathologists. Finally, we were unable to assess the relationship between image analysis algorithm classification and survival in this retrospective cohort as many patients entered the cohort after 2015 and did not have adequate follow-up time for analysis. Future studies must examine the prognostic ability of the image analysis-determined subtype. The Lymphoma Epidemiology of Outcomes cohort study also can provide a large prospective dataset for analysis of the prognostic ability of the IHC image analysis algorithm. Our future research will build on this work by expanding upon the image analysis algorithm presented here: developing quantitative IHC profiling of digital pathology images, creating IHC scoring frameworks to classify patients utilizing IHC expression profiles, and integrating quantitative IHC staining with genomic and clinical features into personalized prognostic models for lymphoid malignancies.

The image analysis algorithm we've developed represents an important first step in developing IHC informatics tools for adaptation to both clinical practice and research. These tools present the potential to address important limitations in IHC interpretation, improve efficiency, reduce the rate of subtype misclassification, and provide more personalized prognoses in DLBCL.

REFERENCES

1. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403(6769):503-11.
2. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* 2004;103(1):275-82.
3. Meyer PN, Fu K, Greiner TC, et al. Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. *J Clin Oncol* 2011;29(2):200-7.
4. Choi WW, Weisenburger DD, Greiner TC, et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clin Cancer Res* 2009;15(17):5494-502.
5. de Jong D, Rosenwald A, Chhanabhai M, et al. Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications--a study from the Lunenburg Lymphoma Biomarker Consortium. *J Clin Oncol* 2007;25(7):805-12.
6. Gavrielides MA, Gallas BD, Lenz P, et al. Observer variability in the interpretation of HER2/neu immunohistochemical expression with unaided and computer-aided digital microscopy. *Arch Pathol Lab Med* 2011;135(2):233-42.
7. Bloom K, Harrington D. Enhanced accuracy and reliability of HER-2/neu immunohistochemical scoring using digital microscopy. *Am J Clin Pathol* 2004;121(5):620-30.

8. Diaz LK, Sahin A, Sneige N. Interobserver agreement for estrogen receptor immunohistochemical analysis in breast cancer: a comparison of manual and computer-assisted scoring methods. *Ann Diagn Pathol* 2004;8(1):23-7.

9. Teras LR, DeSantis CE, Cerhan JR, et al. 2016 US lymphoid malignancy statistics by World Health Organization subtypes. *CA Cancer J Clin* 2016.

10. Flowers CR, Sinha R, Vose JM. Improving outcomes for patients with diffuse large B-cell lymphoma. *CA Cancer J Clin* 2010;60(6):393-408.

11. McKelvey EM, Gottlieb JA, Wilson HE, et al. Hydroxyldaunomycin (Adriamycin) combination chemotherapy in malignant lymphoma. *Cancer* 1976;38(4):1484-93.

12. Coiffier B, Lepage E, Briere J, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med* 2002;346(4):235-42.

13. Pfreundschuh M, Trumper L, Osterborg A, et al. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. *Lancet Oncol* 2006;7(5):379-91.

14. Sehn LH, Berry B, Chhanabhai M, et al. The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. *Blood* 2007;109(5):1857-61.

15. Habermann TM, Weller EA, Morrison VA, et al. Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. *J Clin Oncol* 2006;24(19):3121-7.

16. Maurer MJ, Ghesquieres H, Jais JP, et al. Event-free survival at 24 months is a robust end point for disease-related outcome in diffuse large B-cell lymphoma treated with immunochemotherapy. *J Clin Oncol* 2014;32(10):1066-73.

17. Larouche JF, Berger F, Chassagne-Clement C, et al. Lymphoma recurrence 5 years or later following diffuse large B-cell lymphoma: clinical characteristics and outcome. *J Clin Oncol* 2010;28(12):2094-100.

18. Friedberg JW, Fisher RI. Diffuse large B-cell lymphoma. *Hematol Oncol Clin North Am* 2008;22(5):941-52, ix.

19. Chandler Y, Schechter CB, Jayasekera J, et al. Cost Effectiveness of Gene Expression Profile Testing in Community Practice. *J Clin Oncol* 2018;36(6):554-62.

20. Muris JJ, Meijer CJ, Vos W, et al. Immunohistochemical profiling based on Bcl-2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. *J Pathol* 2006;208(5):714-23.

21. Natkunam Y, Farinha P, Hsi ED, et al. LMO2 protein expression predicts survival in patients with diffuse large B-cell lymphoma treated with anthracycline-based chemotherapy with and without rituximab. *J Clin Oncol* 2008;26(3):447-54.

22. Nyman H, Jerkeman M, Karjalainen-Lindsberg ML, et al. Prognostic impact of activated B-cell focused classification in diffuse large B-cell lymphoma patients treated with R-CHOP. *Mod Pathol* 2009;22(8):1094-101.

23. Visco C, Li Y, Xu-Monette ZY, et al. Comprehensive gene expression profiling and immunohistochemical studies support application of immunophenotypic algorithm for molecular subtype classification in diffuse large B-cell lymphoma: a

report from the International DLBCL Rituximab-CHOP Consortium Program Study. *Leukemia* 2012;26(9):2103-13.

24. Coutinho R, Clear AJ, Owen A, et al. Poor concordance among nine immunohistochemistry classifiers of cell-of-origin for diffuse large B-cell lymphoma: implications for therapeutic strategies. *Clin Cancer Res* 2013;19(24):6686-95.

25. Rimm DL. What brown cannot do for you. *Nat Biotechnol* 2006;24(8):914-6.

26. Rimm DL, Giltnane JM, Moeder C, et al. Bimodal population or pathologist artifact? *J Clin Oncol* 2007;25(17):2487-8.

27. Parker RL, Huntsman DG, Lesack DW, et al. Assessment of interlaboratory variation in the immunohistochemical determination of estrogen receptor status using a breast cancer tissue microarray. *Am J Clin Pathol* 2002;117(5):723-8.

28. Jaraj SJ, Camparo P, Boyle H, et al. Intra- and interobserver reproducibility of interpretation of immunohistochemical stains of prostate cancer. *Virchows Arch* 2009;455(4):375-81.

29. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* 2017;318(22):2199-210.

30. Beck AH, Sangoi AR, Leung S, et al. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci Transl Med* 2011;3(108):108ra13.

31. Kong J, Cooper LA, Wang F, et al. Machine-based morphologic analysis of glioblastoma using whole-slide pathology images uncovers clinically relevant molecular correlates. *PLoS One* 2013;8(11):e81049.

32. Chang H, Han J, Borowsky A, et al. Invariant delineation of nuclear architecture in glioblastoma multiforme for clinical and molecular association. *IEEE Trans Med Imaging* 2013;32(4):670-82.

33. Madabhushi A, Agner S, Basavanhally A, et al. Computer-aided prognosis: predicting patient and disease outcome via quantitative fusion of multi-scale, multi-modal data. *Comput Med Imaging Graph* 2011;35(7-8):506-14.

34. Sertel O, Kong J, Shimada H, et al. Computer-aided Prognosis of Neuroblastoma on Whole-slide Images: Classification of Stromal Development. *Pattern Recognit* 2009;42(6):1093-103.

35. Kong J, Sertel O, Boyer KL, et al. Computer-assisted grading of neuroblastic differentiation. *Arch Pathol Lab Med* 2008;132(6):903-4; author reply 4.

36. Litjens G, Sanchez CI, Timofeeva N, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep* 2016;6:26286.

37. Hou L, Samaras D, Kurc TM, et al. Patch-based Convolutional Neural Network for Whole Slide Tissue Image Classification. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2016;2016:2424-33.

38. Kothari S, Phan JH, Young AN, et al. Histological image classification using biologically interpretable shape-based features. *BMC Med Imaging* 2013;13:9.

39. Yuan Y, Failmezger H, Rueda OM, et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci Transl Med* 2012;4(157):157ra43.

40. Rutledge WC, Kong J, Gao J, et al. Tumor-infiltrating lymphocytes in glioblastoma are associated with specific genomic alterations and related to transcriptional class. *Clin Cancer Res* 2013;19(18):4951-60.

41. Cooper LA, Gutman DA, Chisolm C, et al. The tumor microenvironment strongly impacts master transcriptional regulators and gene expression class of glioblastoma. *Am J Pathol* 2012;180(5):2108-19.

42. Mouelhi A, Sayadi M, Fnaiech F, et al. A new automatic image analysis method for assessing estrogen receptors' status in breast tissue specimens. *Comput Biol Med* 2013;43(12):2263-77.

43. Rexhepaj E, Brennan DJ, Holloway P, et al. Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of oestrogen and progesterone receptor levels in breast cancer. *Breast Cancer Res* 2008;10(5):R89.

44. Minot DM, Voss J, Rademacher S, et al. Image analysis of HER2 immunohistochemical staining. Reproducibility and concordance with fluorescence in situ hybridization of a laboratory-validated scoring technique. *Am J Clin Pathol* 2012;137(2):270-6.

45. Ayad E, Mansy M, Elwi D, et al. Comparative study between quantitative digital image analysis and fluorescence in situ hybridization of breast cancer equivocal

human epidermal growth factor receptors 2 score 2(+) cases. *J Pathol Inform* 2015;6:31.

46. Skaland I, Ovestad I, Janssen EA, et al. Digital image analysis improves the quality of subjective HER-2 expression scoring in breast cancer. *Appl Immunohistochem Mol Morphol* 2008;16(2):185-90.

47. Abubakar M, Orr N, Daley F, et al. Prognostic value of automated KI67 scoring in breast cancer: a centralised evaluation of 8088 patients from 10 study groups. *Breast Cancer Res* 2016;18(1):104.

48. Rizzardi AE, Johnson AT, Vogel RI, et al. Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. *Diagn Pathol* 2012;7:42.

49. Messersmith W, Oppenheimer D, Peralba J, et al. Assessment of Epidermal Growth Factor Receptor (EGFR) signaling in paired colorectal cancer and normal colon tissue samples using computer-aided immunohistochemical analysis. *Cancer Biol Ther* 2005;4(12):1381-6.

50. International Non-Hodgkin's Lymphoma Prognostic Factors P. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med* 1993;329(14):987-94.

51. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* 2002;346(25):1937-47.

52. Dunleavy K, Pittaluga S, Czuczman MS, et al. Differential efficacy of bortezomib plus chemotherapy within molecular subtypes of diffuse large B-cell lymphoma. *Blood* 2009;113(24):6069-76.

53. Yang Y, Shaffer AL, 3rd, Emre NC, et al. Exploiting synthetic lethality for the therapy of ABC diffuse large B cell lymphoma. *Cancer Cell* 2012;21(6):723-37.

54. Nowakowski GS, Chiappella A, Witzig TE, et al. ROBUST: Lenalidomide-R-CHOP versus placebo-R-CHOP in previously untreated ABC-type diffuse large B-cell lymphoma. *Future Oncol* 2016;12(13):1553-63.

55. Wilson WH, Young RM, Schmitz R, et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med* 2015;21(8):922-6.

56. Moskowitz CH, Zelenetz AD, Kewalramani T, et al. Cell of origin, germinal center versus nongerminal center, determined by immunohistochemistry on tissue microarray, does not correlate with outcome in patients with relapsed and refractory DLBCL. *Blood* 2005;106(10):3383-5.

57. Ott G, Ziepert M, Klapper W, et al. Immunoblastic morphology but not the immunohistochemical GCB/nonGCB classifier predicts outcome in diffuse large B-cell lymphoma in the RICOVER-60 trial of the DSHNHL. *Blood* 2010;116(23):4916-25.

58. Ott MM, Horn H, Kaufmann M, et al. The Hans classificator does not predict outcome in diffuse large B cell lymphoma in a large multicenter retrospective analysis of R-CHOP treated patients. *Leuk Res* 2012;36(5):544-5.

59. Gutierrez-Garcia G, Cardesa-Salzmann T, Climent F, et al. Gene-expression profiling and not immunophenotypic algorithms predicts prognosis in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. *Blood* 2011;117(18):4836-43.

60. Fu K, Weisenburger DD, Choi WW, et al. Addition of rituximab to standard chemotherapy improves the survival of both the germinal center B-cell-like and non-germinal center B-cell-like subtypes of diffuse large B-cell lymphoma. *J Clin Oncol* 2008;26(28):4587-94.

61. Berglund M, Thunberg U, Amini RM, et al. Evaluation of immunophenotype in diffuse large B-cell lymphoma and its impact on prognosis. *Mod Pathol* 2005;18(8):1113-20.

62. Read JA, Koff JL, Nastoupil LJ, et al. Evaluating cell-of-origin subtype methods for predicting diffuse large B-cell lymphoma survival: a meta-analysis of gene expression profiling and immunohistochemistry algorithms. *Clin Lymphoma Myeloma Leuk* 2014;14(6):460-7 e2.

63. Zu Y, Steinberg SM, Campo E, et al. Validation of tissue microarray immunohistochemistry staining and interpretation in diffuse large B-cell lymphoma. *Leuk Lymphoma* 2005;46(5):693-701.

64. Laszlo D, Pruneri G, Andreola G, et al. Tissue microarrays in diffuse large B-cell lymphomas: are they really able to identify distinct prognostic groups in lymphomas of both nodal and extranodal origin? *Int J Surg Pathol* 2011;19(4):417-24.

65. Barrans S, Crouch S, Smith A, et al. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. *J Clin Oncol* 2010;28(20):3360-5.

66. Savage KJ, Johnson NA, Ben-Neriah S, et al. MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. *Blood* 2009;114(17):3533-7.

67. Iqbal J, Neppalli VT, Wright G, et al. BCL2 expression is a prognostic marker for the activated B-cell-like type of diffuse large B-cell lymphoma. *J Clin Oncol* 2006;24(6):961-8.

68. Iqbal J, Meyer PN, Smith LM, et al. BCL2 predicts survival in germinal center B-cell-like diffuse large B-cell lymphoma treated with CHOP-like therapy and rituximab. *Clin Cancer Res* 2011;17(24):7785-95.

69. Johnson NA, Slack GW, Savage KJ, et al. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J Clin Oncol* 2012;30(28):3452-9.

70. Green TM, Young KH, Visco C, et al. Immunohistochemical double-hit score is a strong predictor of outcome in patients with diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J Clin Oncol* 2012;30(28):3460-7.

71. Tsuyama N, Sakata S, Baba S, et al. BCL2 expression in DLBCL: reappraisal of immunohistochemistry with new criteria for therapeutic biomarker evaluation. *Blood* 2017;130(4):489-500.

72. Mahmoud AZ, George TI, Czuchlewski DR, et al. Scoring of MYC protein expression in diffuse large B-cell lymphomas: concordance rate among hematopathologists. *Mod Pathol* 2015;28(4):545-51.

73. Samsi S, Krishnamurthy AK, Gurcan MN. An Efficient Computational Framework for the Analysis of Whole Slide Images: Application to Follicular Lymphoma Immunohistochemistry. *J Comput Sci* 2012;3(5):269-79.

74. Lozanski G, Pennell M, Shana'ah A, et al. Inter-reader variability in follicular lymphoma grading: Conventional and digital reading. *J Pathol Inform* 2013;4:30.

75. Belkacem-Boussaid K, Samsi S, Lozanski G, et al. Automatic detection of follicular regions in H&E images using iterative shape index. *Comput Med Imaging Graph* 2011;35(7-8):592-602.

76. Samsi S, Lozanski G, Shana'ah A, et al. Detection of follicles from IHC-stained slides of follicular lymphoma using iterative watershed. *IEEE Trans Biomed Eng* 2010;57(10):2609-12.

77. Belkacem-Boussaid K, Pennell M, Lozanski G, et al. Computer-aided classification of centroblast cells in follicular lymphoma. *Anal Quant Cytol Histol* 2010;32(5):254-60.

78. Sertel O, Lozanski G, Shana'ah A, et al. Computer-aided detection of centroblasts for follicular lymphoma grading using adaptive likelihood-based cell segmentation. *IEEE Trans Biomed Eng* 2010;57(10):2613-6.

79. Kong H, Gurcan M, Belkacem-Boussaid K. Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting. *IEEE Trans Med Imaging* 2011;30(9):1661-77.

80. Fauzi MF, Pennell M, Sahiner B, et al. Classification of follicular lymphoma: the effect of computer aid on pathologists grading. *BMC Med Inform Decis Mak* 2015;15:115.

81. Scott DW, Wright GW, Williams PM, et al. Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. *Blood* 2014;123(8):1214-7.

82. Scott DW, Mottok A, Ennishi D, et al. Prognostic Significance of Diffuse Large B-Cell Lymphoma Cell of Origin Determined by Digital Gene Expression in Formalin-Fixed Paraffin-Embedded Tissue Biopsies. *J Clin Oncol* 2015;33(26):2848-56.

TABLES/FIGURES

Figure 1. Hans Classification System.

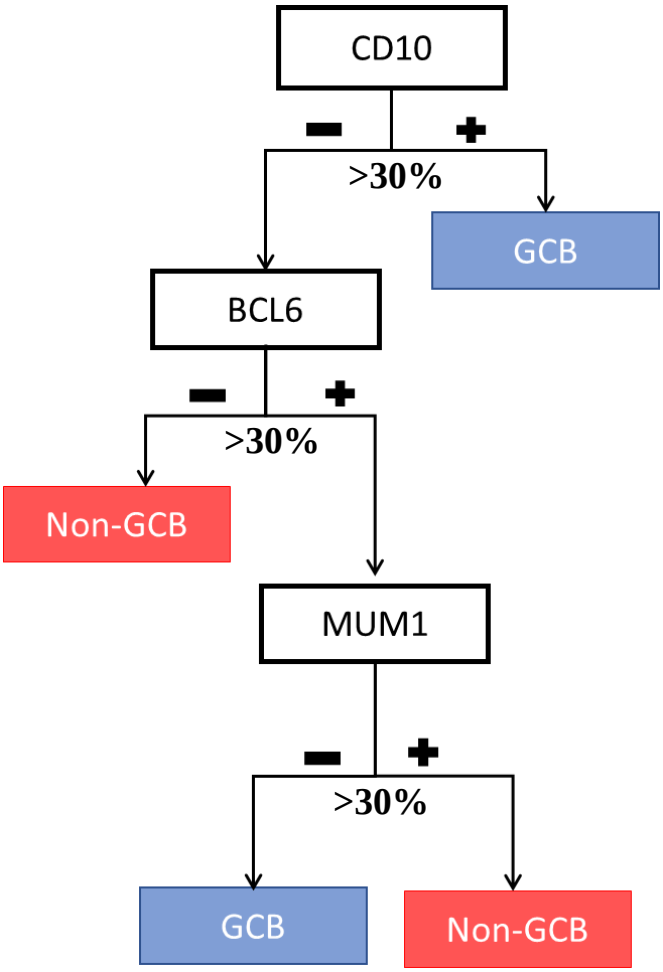

Figure 2. Image Analysis Algorithm.

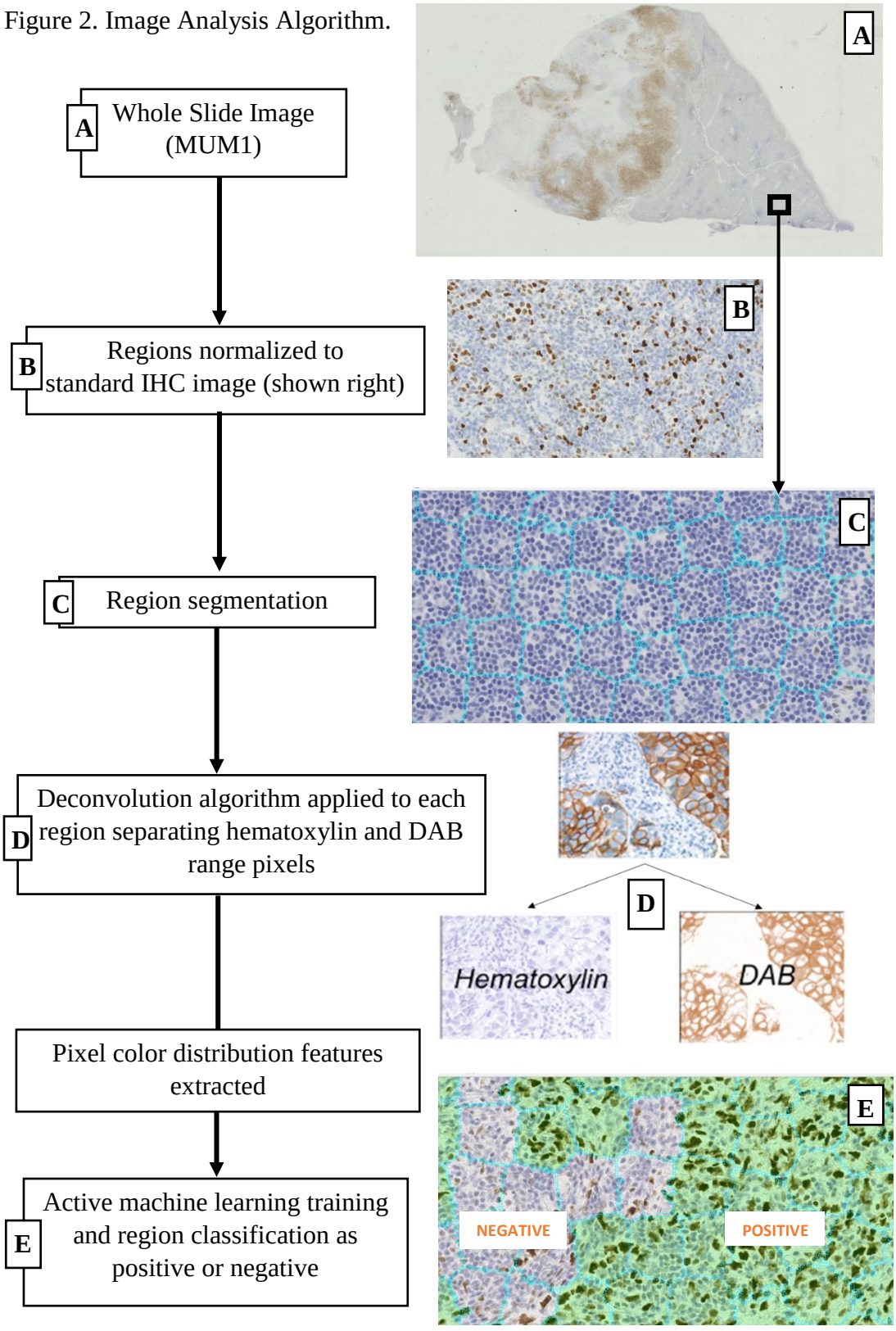

Figure 3. Heatmap of positive region density (right) for MUM1 immunohistochemical Whole Slide Image. The red represents high positive region density, while blue represents low positive region density.

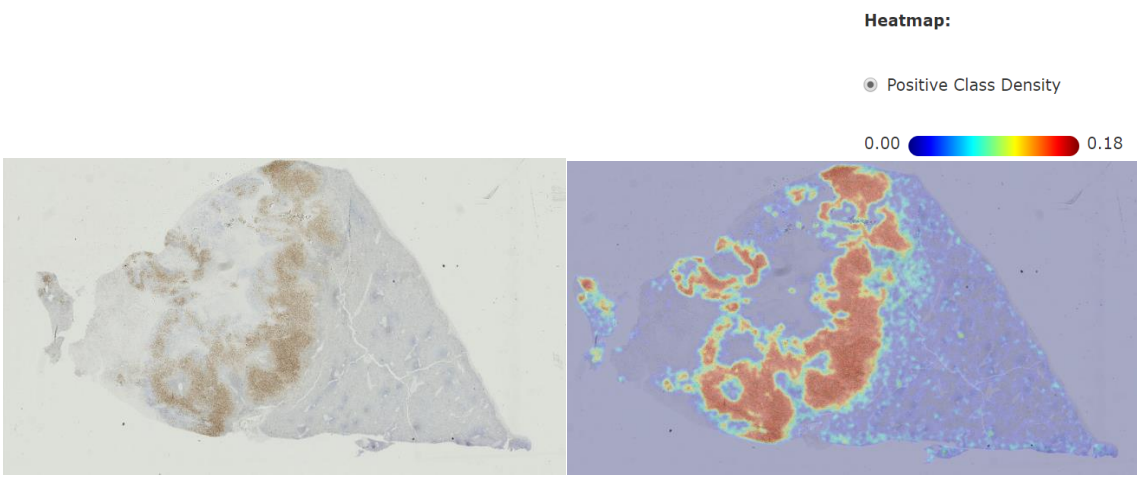

Figure 4. Receiver operator curve comparing percent positive regions from image analysis algorithm to pathologist classification for CD10.

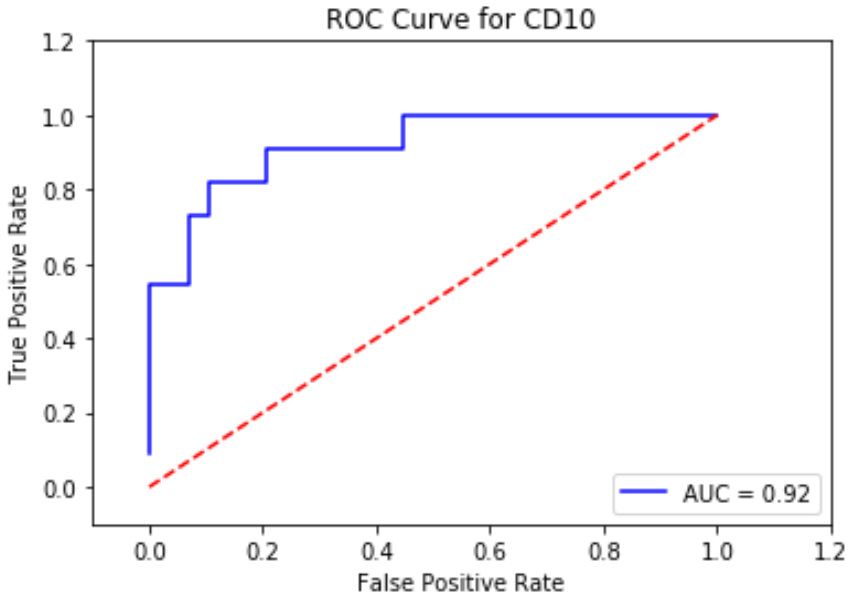

Figure 5. Receiver operator curve comparing percent positive regions from image analysis algorithm to pathologist classification for BCL6.

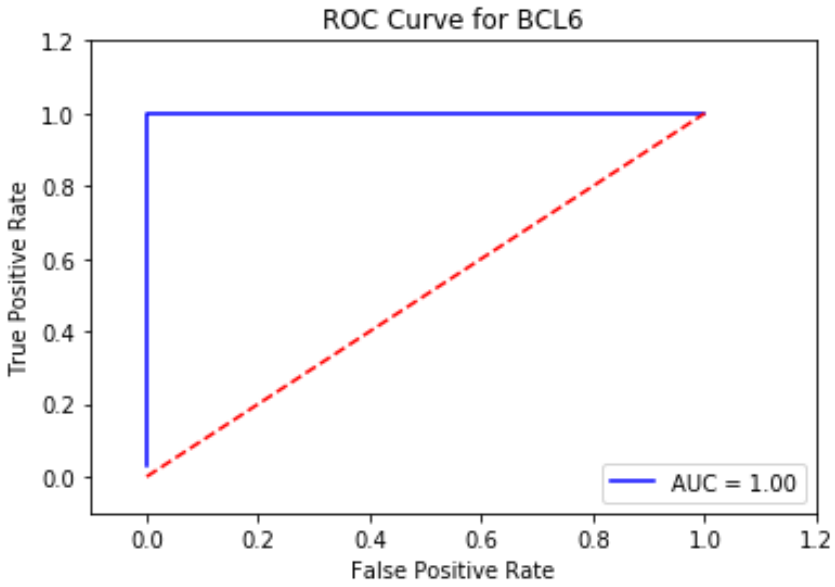

Figure 6. Receiver operator curve comparing percent positive regions from image analysis algorithm to pathologist classification for MUM1.

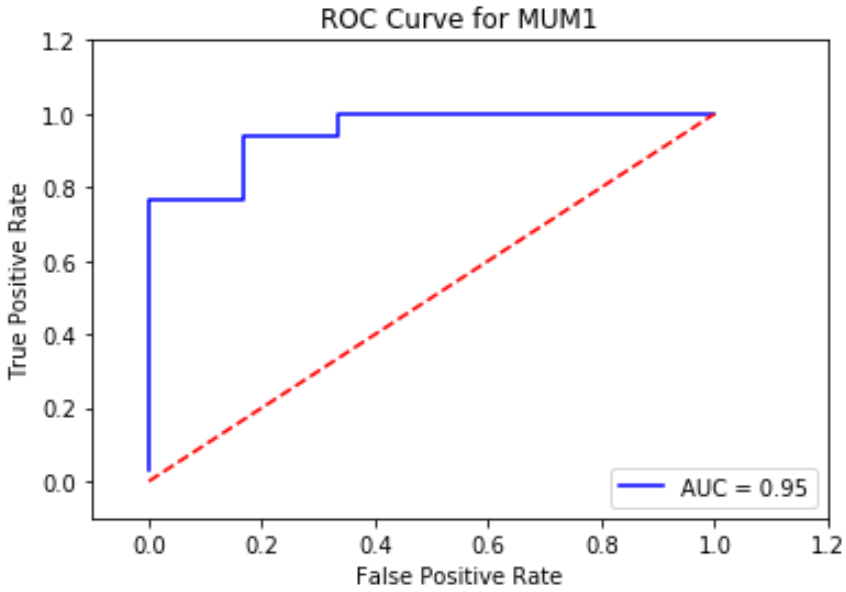

Figure 7. Sequential application of the Hans classification system using image analysis algorithm output to predict subtype. The x-axis for each of the charts on the right represents the percent of positive regions for each IHC stain. The line color represents pathologist-determined subtype. The dotted lines represent the region positivity thresholds determined by the ROC curves. The blue and red shaded regions represent GCB and non-GCB subtype classification by image analysis algorithm respectively.

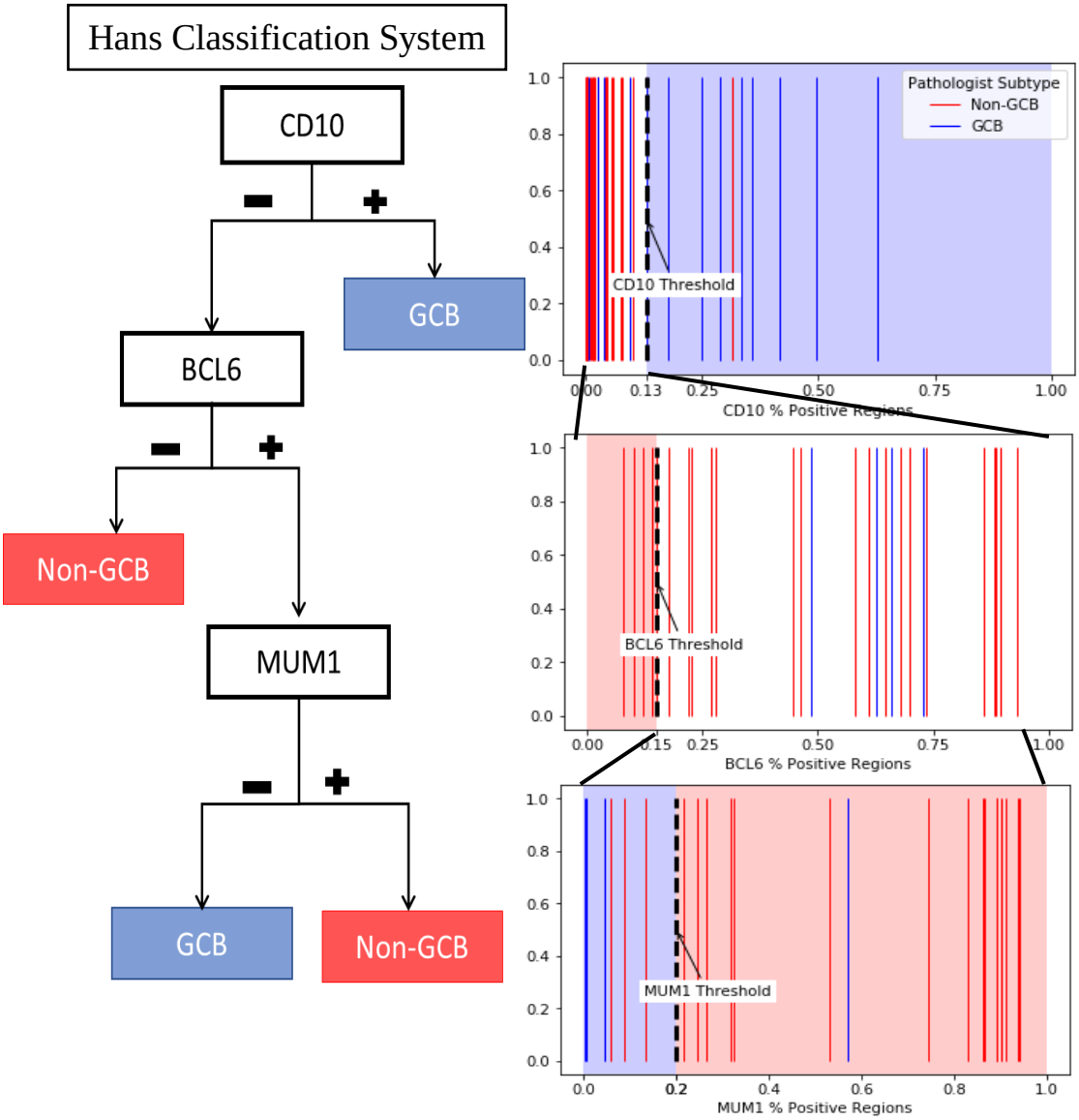

Table 1. Clinical Characteristics of DLBCL patients by Hans classification subtype.

Characteristic	GCB	NonGCB	All	p-value*
Total Number of Patients (%)	15	25	40	
Age, years				0.86
Mean	65	60	61.5	
IQR	48-76	46-68	46-76	
≤60	5	9	14	
>60	10	16	26	
Male gender	6	11	17	0.69
ECOG performance status				0.43
0-1	7	14	21	
≥2	6	5	11	
Unknown	2	6	8	
Disease stage				0.39
I/II	6	6	12	
III/IV	6	15	21	
Unknown	3	4	7	
LDH level				0.46
Normal	8	9	17	
>ULN	7	16	23	
B-symptoms				0.56
Absent	9	11	20	
Present	3	8	11	
Unknown	3	6	9	
No. of extranodal sites				0.41
0-1	7	6	13	
≥2	7	14	21	
Unknown	1	5	6	
International Prognostic Index				0.22
0-1	5	1	6	
2	2	5	7	
3	2	4	6	
4-5	2	7	9	
Unknown	4	8	12	
First line of treatment				0.90
R-CHOP	8	18	26	
R+/- Other	4	6	10	
Other	3	1	4	