

# Innovative Tools for Quantitative Analysis of NAFLD Histology

Samer Gawrieh, MD

Associate Professor of Clinical Medicine

Indiana University School of Medicine

Why Innovate?

Genetic Modifiers

Pathogenesis-  
Animal Studies

Therapeutic Trials

NAFLD Histological  
Phenotypes

NAFL

Borderline

NASH

Steatosis

Lobular  
Inflammation

Ballooning

Portal  
Inflammation

Fibrosis



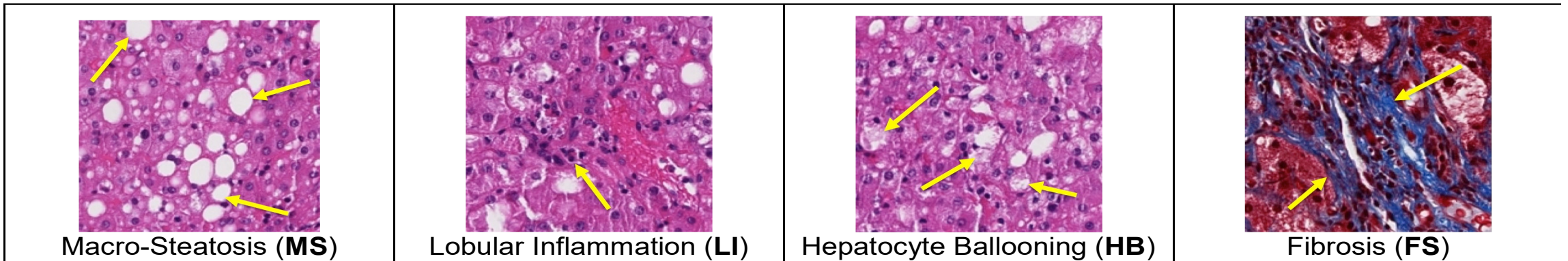


# Biopsy sample analysis

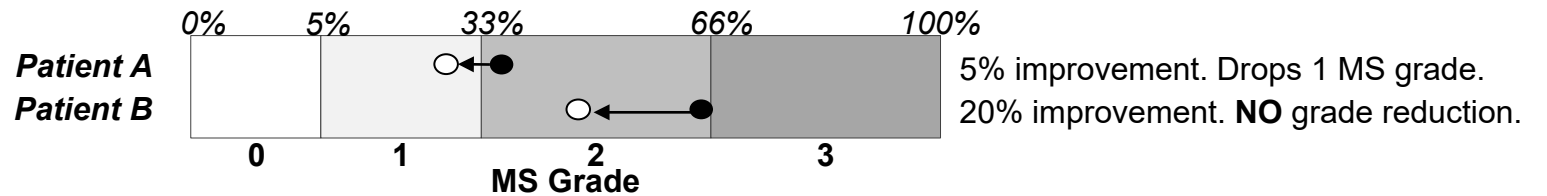


# Current State:

## Manual and Semi-quantitative Grading of NAFLD Lesions

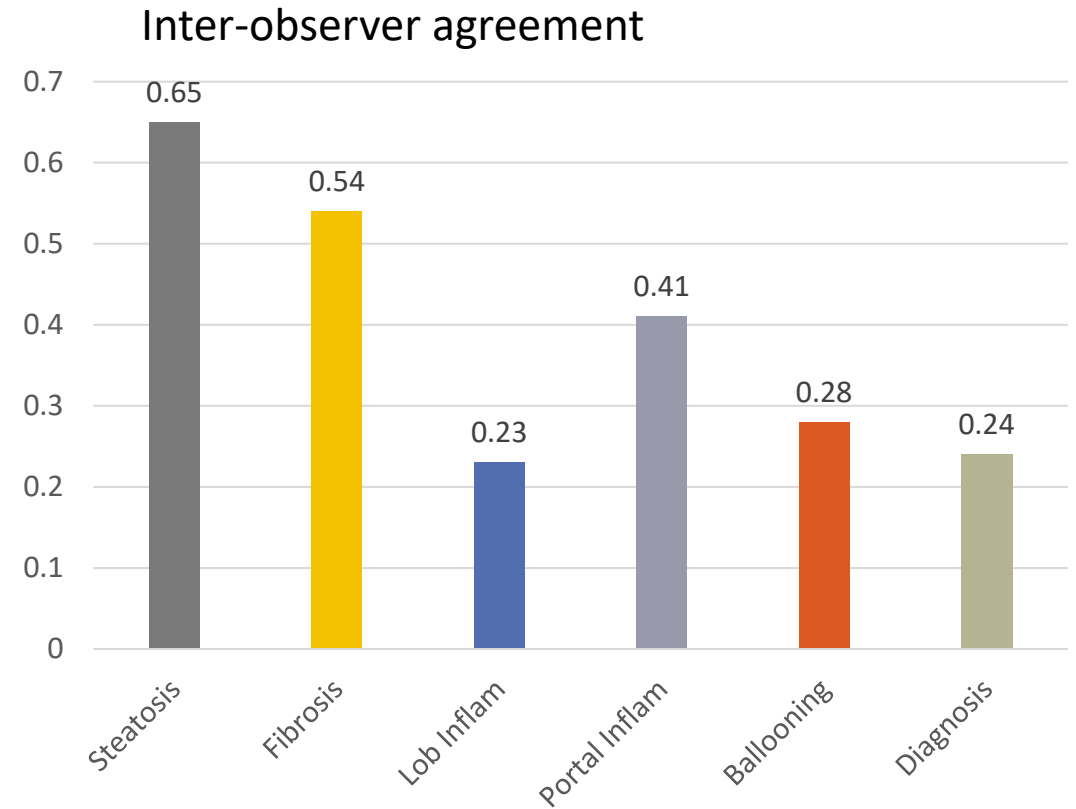
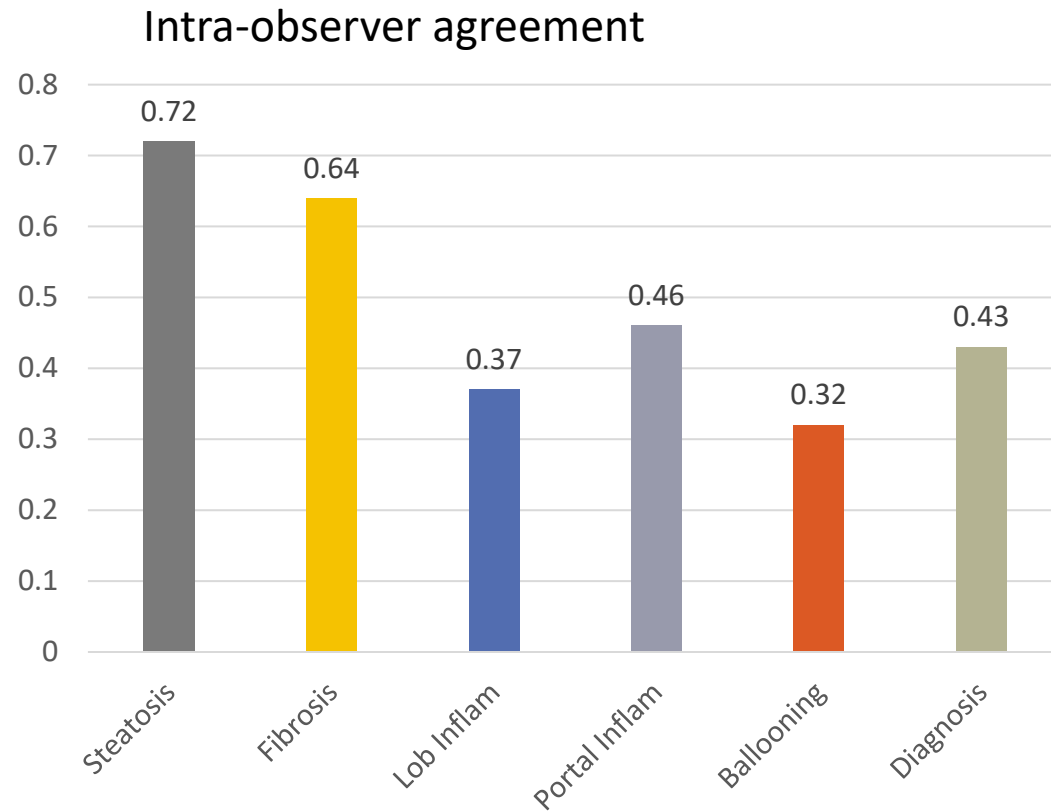


- NAFLD Activity Score (NAS) = MS + LI + HB
  - 0-3 for MS and LI
  - 0-2 for HB
  - 0-8 for NAS
  - 0-4 for FS

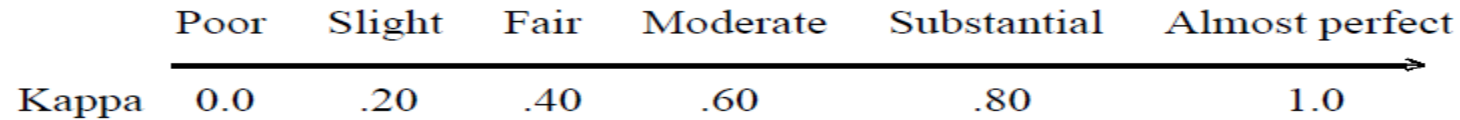


# Observer Agreement (k) on Reading NAFLD Histological Features

(65 biopsies, 2 pathologists, 260 readings)



# Reported Inter-observer Agreement ( $\kappa$ ) on Interpretation of NAFLD Histology



Feature	Inter-Observer Agreement					
	Younossi 1998	Fukusato 2005	Kleiner 2005	Juluri 2011	Gawrieh 2011	Davison 2020
Cases (n)	53	8	32	48	65	339
Observers (n)	4	21	9	2	2	3
SG	0.64	0.53	0.79	0.62	0.65-0.74	0.61
FS	0.60	0.55	0.84	0.35	0.54-0.56	0.48
LI	0.21	0.09	0.45	0.44	0.20-0.23	0.33
PI	0.18	0.19	0.45		0.32-0.41	
HB	0.50	0.14	0.56	0.25	0.18-0.28	0.52
Diagnosis	0.50	0.21	0.61	0.46	0.27-0.39	0.40

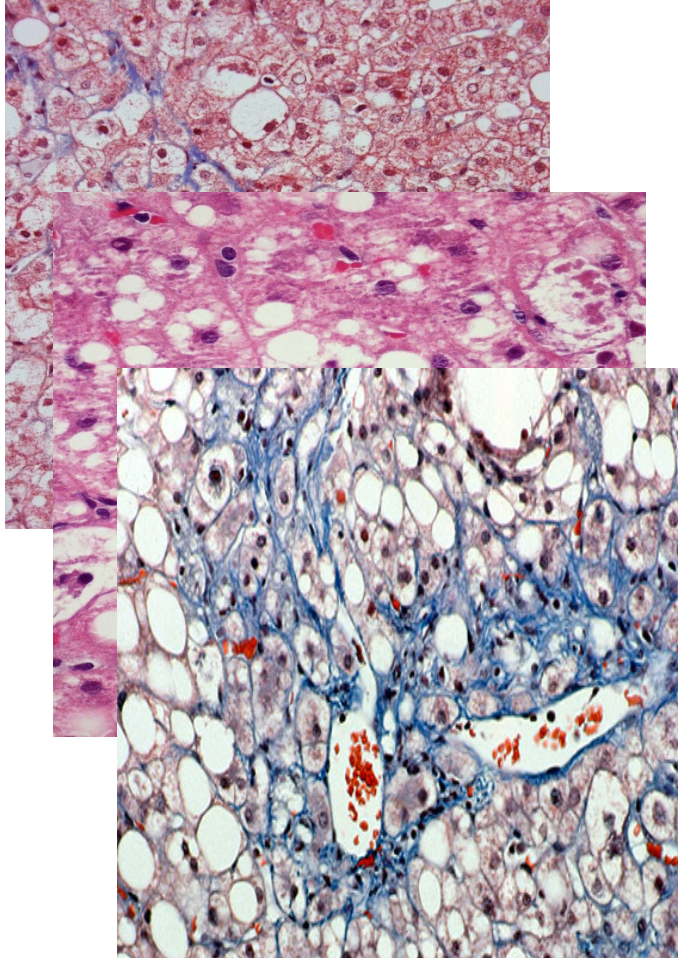


# Interventions

## Education / training

Slide review

Discussion of diagnostic criteria



## Scoring sheet

Simplified written  
criteria for diagnosis

### Fatty liver disease scoring sheet

Photo case \_\_\_\_\_

Patient- \_\_\_\_\_ S# \_\_\_\_\_ MRN \_\_\_\_\_

#### Steatosis grade \_\_\_\_\_

0 ≤ 5%  
1=6-33%  
2=34-66%  
3=>67%

#### Steatosis predominant distribution \_\_\_\_\_

0=Zone 3  
1=Zone 1  
2=Azonal  
3=Panacinar

#### Microvesicular steatosis \_\_\_\_\_

0= Not present  
1=Present

#### Fibrosis stage \_\_\_\_\_

0=none  
1=Perisinusoidal or periportal  
1A=Mild, zone 3, perisinusoidal  
1B=Moderate, zone 3, perisinusoidal  
1C=Portal/periportal,  
2=Perisinusoidal and portal/periportal  
3=bridging fibrosis  
4=cirrhosis

#### Lobular inflammation \_\_\_\_\_

0= no foci  
1= 2 foci/200 x field  
2= 2-4 foci/200 x field  
3=>4

#### Portal inflammation \_\_\_\_\_

0=none  
1=mild  
2=moderate  
3=severe

#### Zone 3 Hepatocyte ballooning \_\_\_\_\_

0=none  
1=few balloon cells  
2=many balloon cells

#### Mallory's hyaline \_\_\_\_\_

0=None to rare  
1=Many

#### Diagnosis \_\_\_\_\_

1. Fatty liver
2. Possible/borderline SH
3. Steatohepatitis (SH)
4. Normal

#### Additional or other diagnosis \_\_\_\_\_

#### Definitions

\* **FATTY LIVER** diagnosed when **ONLY STEATOSIS (> 5%)** is present

\* **Steatohepatitis (SH) is steatosis + 2 of the following zone 3, centric features**

1. Hepatocellular ballooning
2. Inflammatory infiltrate
3. Pericellular/Perisinusoidal fibrosis

\***Possible/borderline SH is steatosis + 1 of the above zone 3, centric features**

4<sup>th</sup> edition 1-30-07



# Pre- and Post-Intervention Observer Agreement (k) on NAFLD Histological Features and Phenotype

(65 biopsies, 2 pathologists, 520 readings)

Feature	Intra-observer		Inter-observer	
	Pre	Post	Pre	Post
Steatosis grade	0.72	0.75	0.65	0.74
Fibrosis stage	0.64	0.75	0.54	0.56
Lobular inflammation	0.37	0.48	0.23	0.20
Portal inflammation	0.46	0.58	0.41	0.32
Hepatocellular ballooning	0.32	0.56*	0.28	0.18
Diagnostic classification	0.51	0.54	0.27	0.39

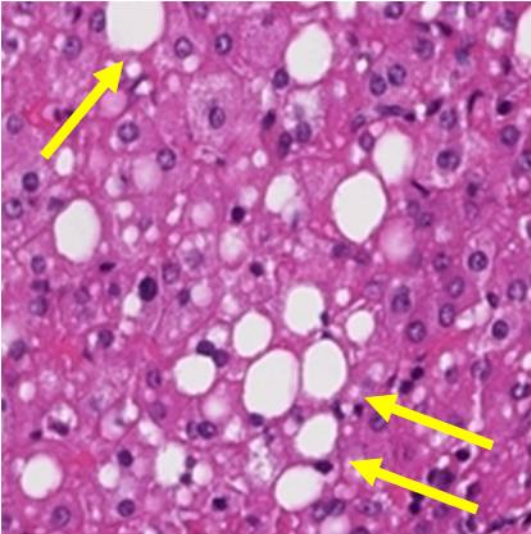
\*  $p$  for pre- versus post comparisons was significant only for intra-observer k on HB (0.009)



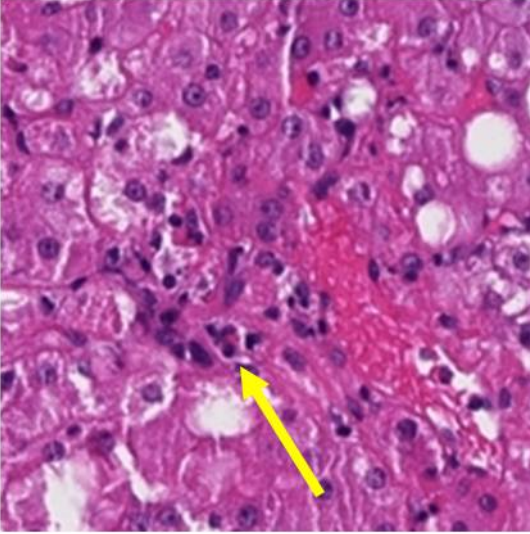
NAFLD epidemic is on the rise, new approaches to decrease interobserver variability in interpretation of NAFLD histology are urgently needed. Refined histopathologic criteria and training on assessment of HB and LI would likely have the largest impact on reproducibility of NAFLD phenotyping and staging. Automation of assessment of NAFLD histologic features to improve accuracy and reproducibility of the interpretation is another consideration. This can be envi-



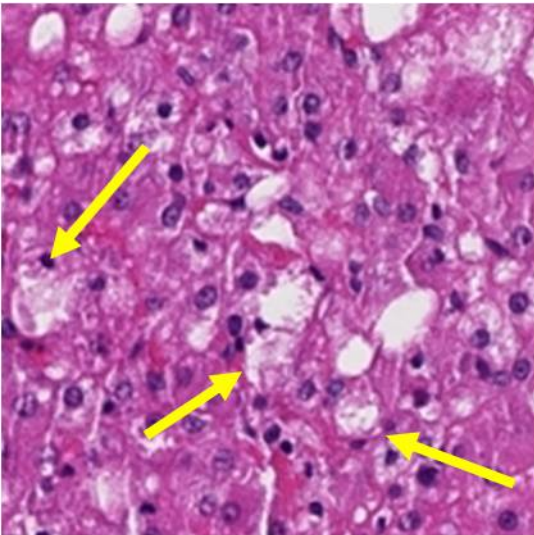
# What to Detect and Quantify: Cardinal NAFLD Lesions



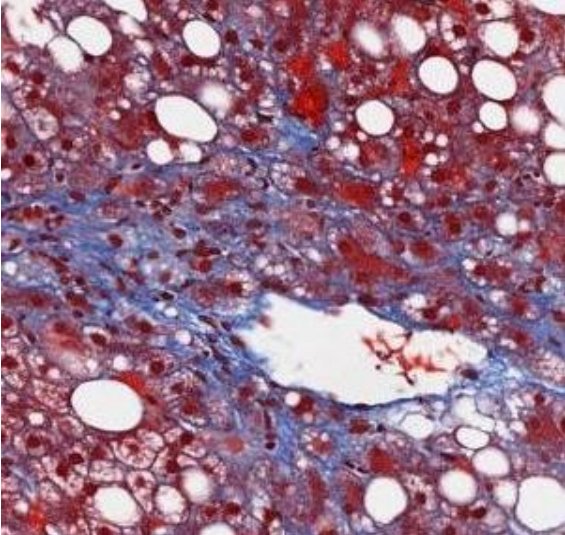
Macrosteatosis



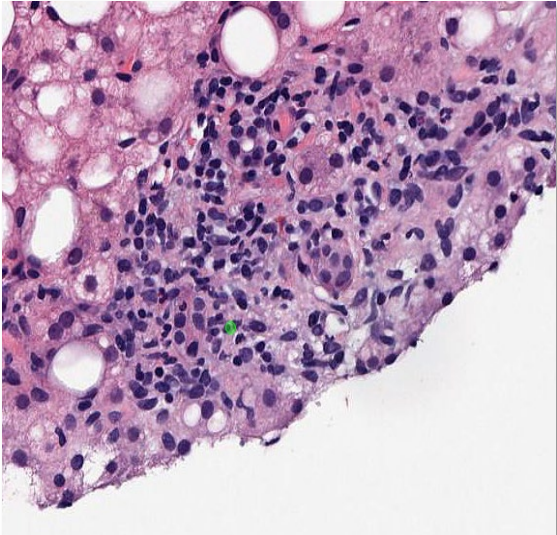
Lobular Inflammation



Hepatocyte Ballooning

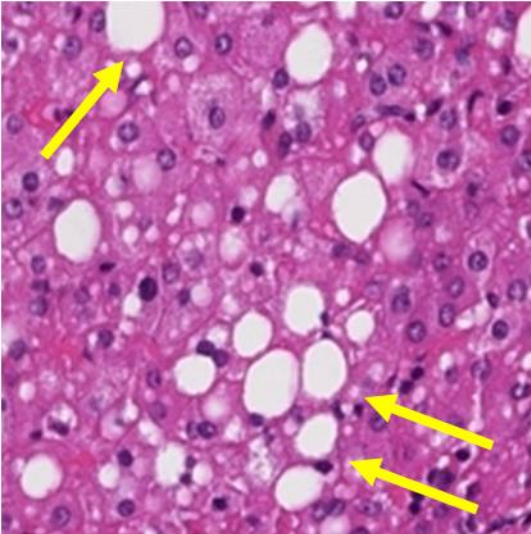


Fibrosis

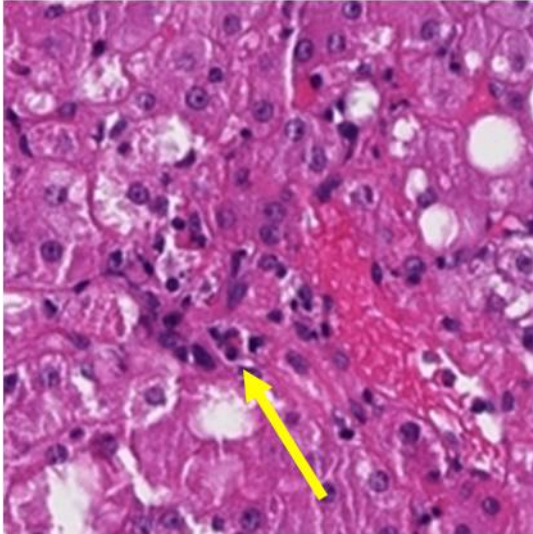


Portal Inflammation

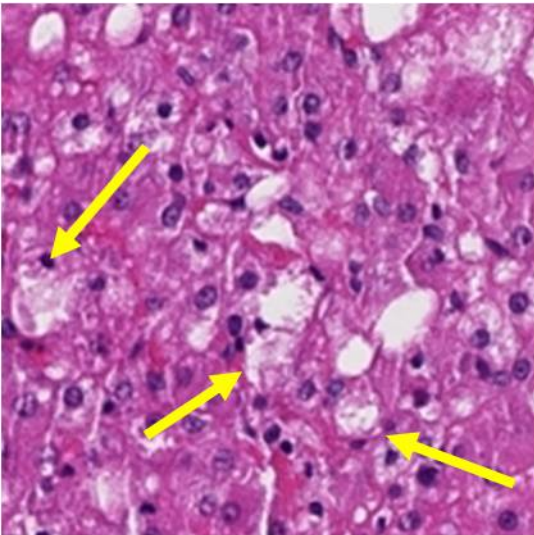
# What to Detect and Quantify: Cardinal NAFLD Lesions



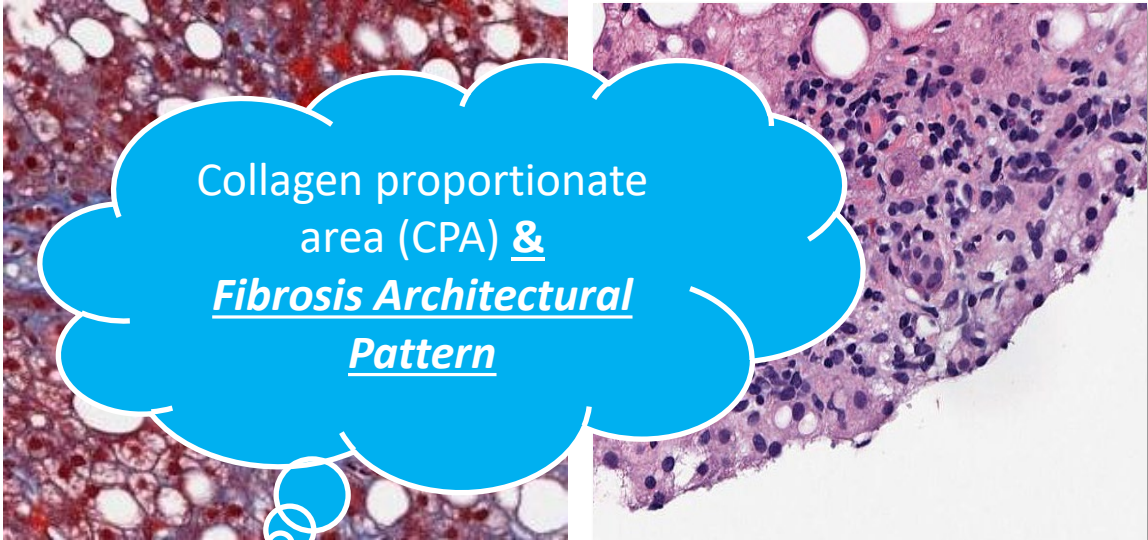
Macrosteatosis



Lobular Inflammation



Hepatocyte Ballooning



Fibrosis

Portal Inflammation

# How to Quantify: Automation Approaches and Requirements

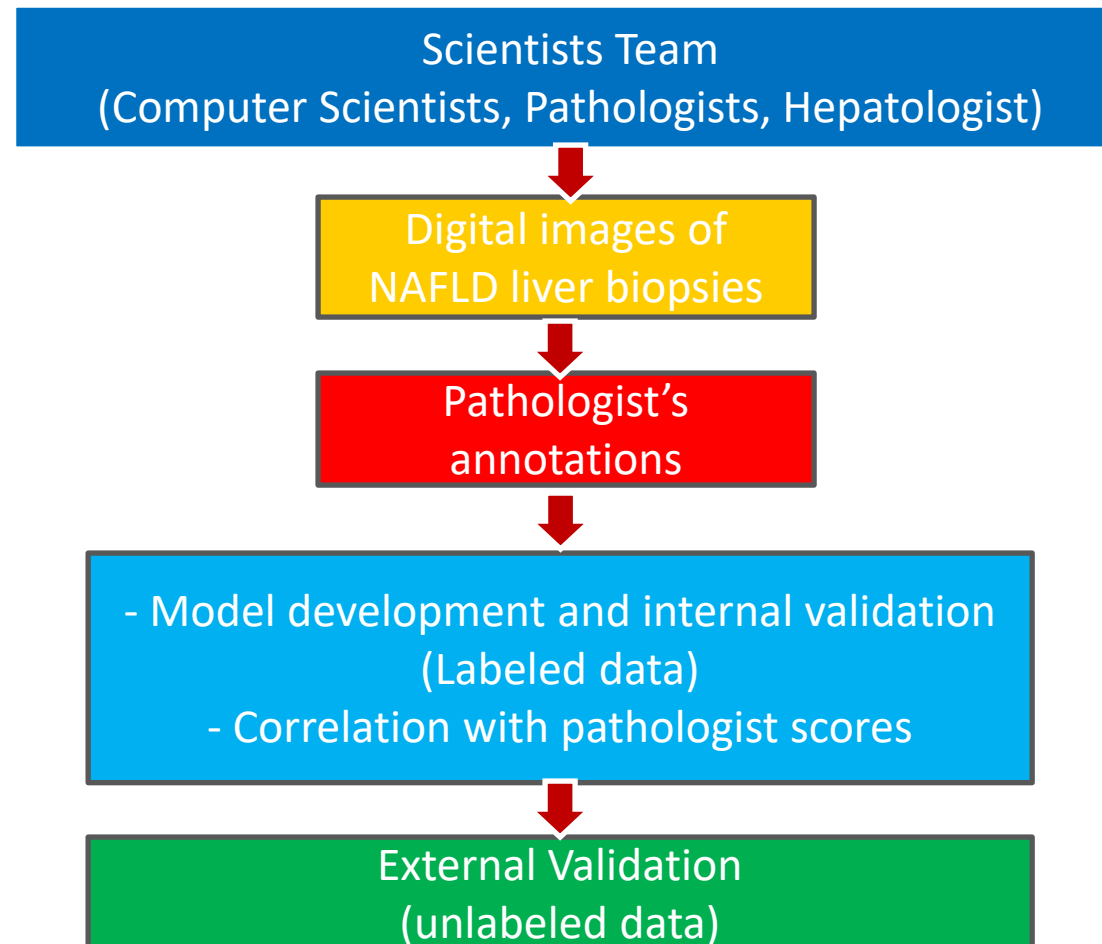
	Machine learning	Algorithm-based	Other (Adobe- or color-based tools)
What's quantified	NAFLD lesions	Correlates with NAFLD lesions	Usually fibrosis and/or steatosis assessment
Stains	H&E, Mason trichrome, Sirius red	Unstained/stained slides	H&E, Mason trichrome, Sirius red
Digital Slide Scanner	√	√	+/-
Equipment additional to software	None	SHG/TPEF Microscope*	None

\* SHG: Second harmonic generation microscopy, TPEF: Two-photon excitation fluorescence microscopy

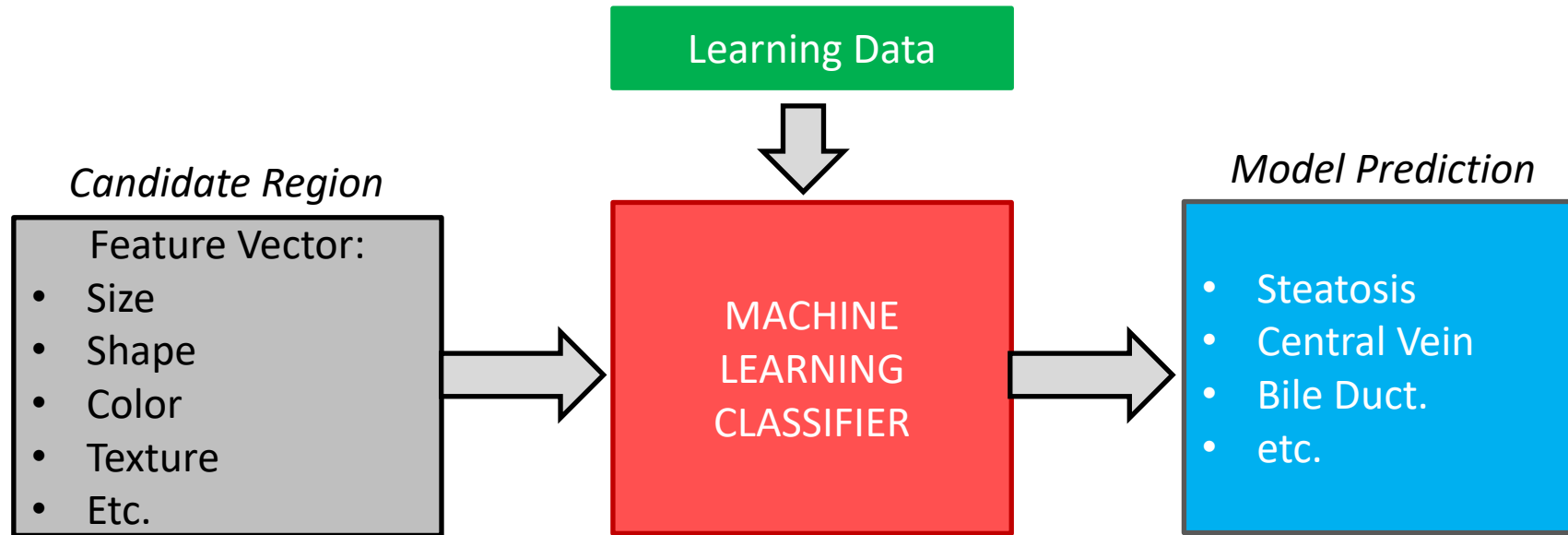
# Machine Learning (ML)/ Artificial Intelligence (AI)

- ML: algorithms and statistical models that learn from labelled training data, from which they are able to recognize and infer patterns
- General AI: ability of a machine to communicate, reason and operate independently in both familiar and novel scenarios in a similar manner to a human
- Commonly, ML interchangeable with AI

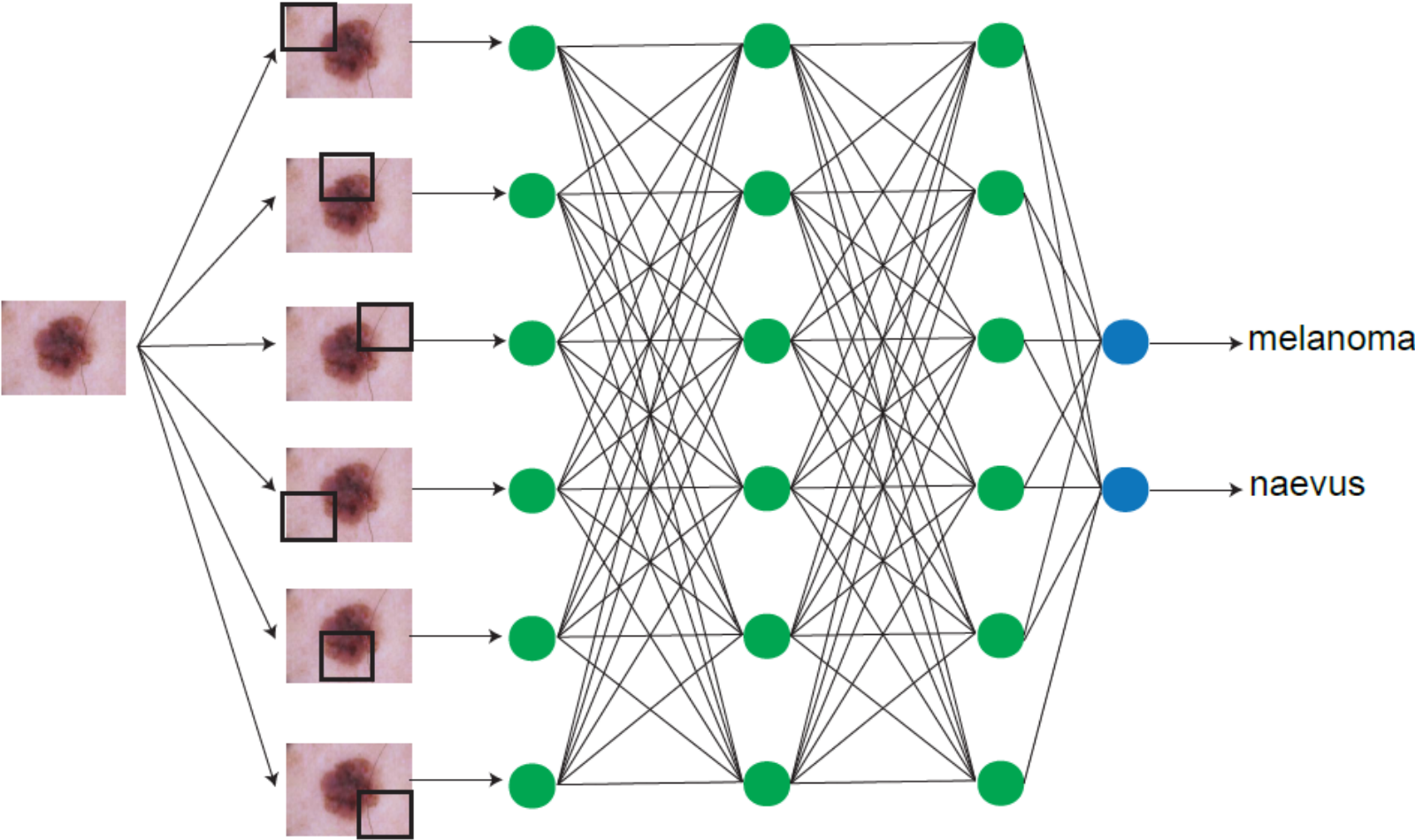
# General Approach to Developing ML Models for NAFLD Histology Analysis



# Supervised Machine Learning



# Deep Machine Learning and Neural Networks



Du-Harpur X et al. Br J Dermatol 2020



# Pathologist Annotations Software

Current Image: FLE157\_HE  
Liver Biopsy Labeling

Click to add labels

Index



(793,1384)



Select Current Type

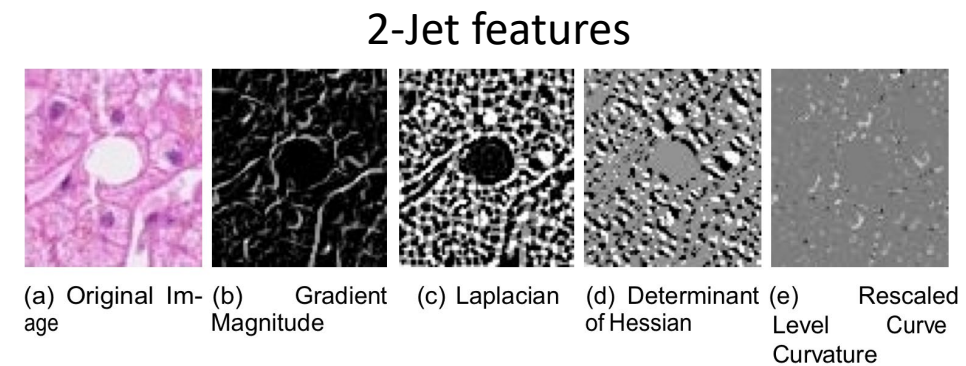
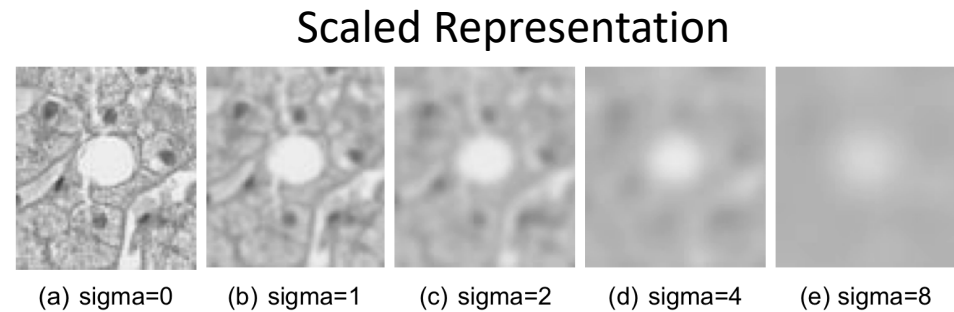
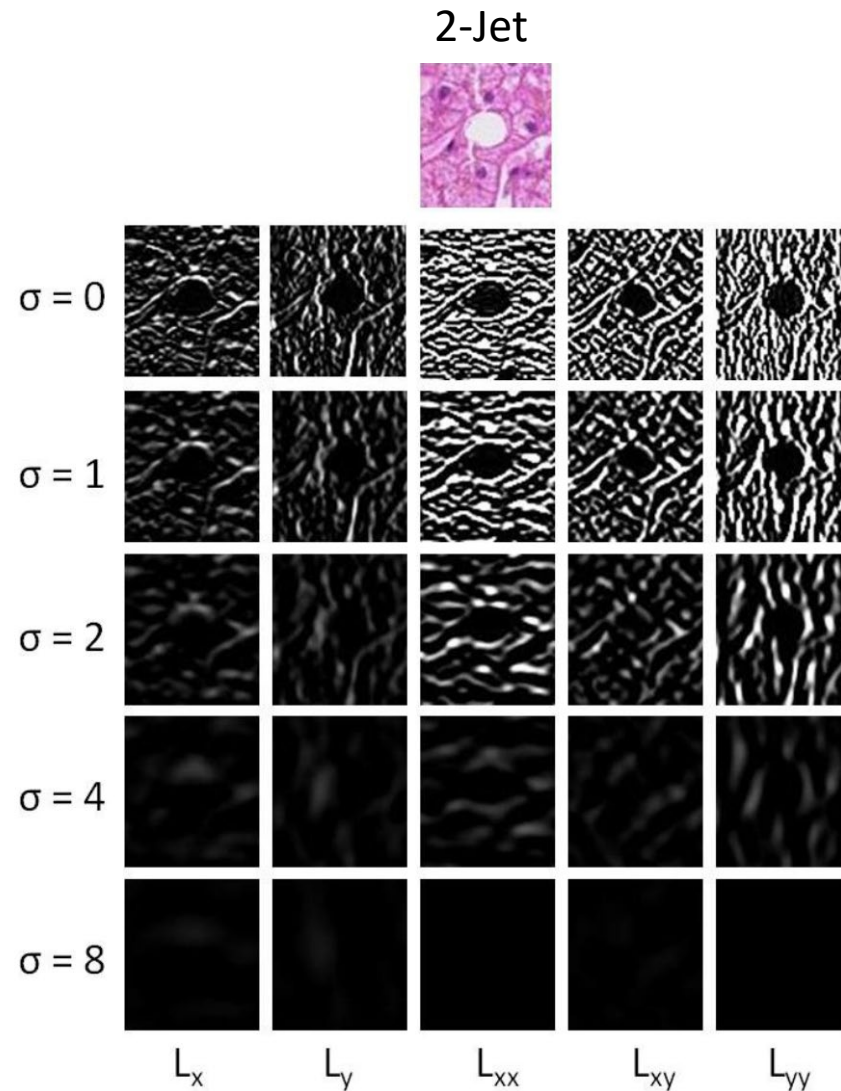
- Macro Fat (red)
- Micro Fat (orange)
- Portal Vein (blue)
- Central Vein (green)
- Bile Duct (magenta)
- Portal Artery (pink)
- Sinusoid (yellow)
- Other (black)
- Fibrosis (red\*\*\*)
- Lobular Inflammation (green\*\*\*)
- Portal Inflammation (yellow\*\*\*)
- Ballooning (magenta\*\*\*)

Undo Last Label

Zoom In Zoom Out

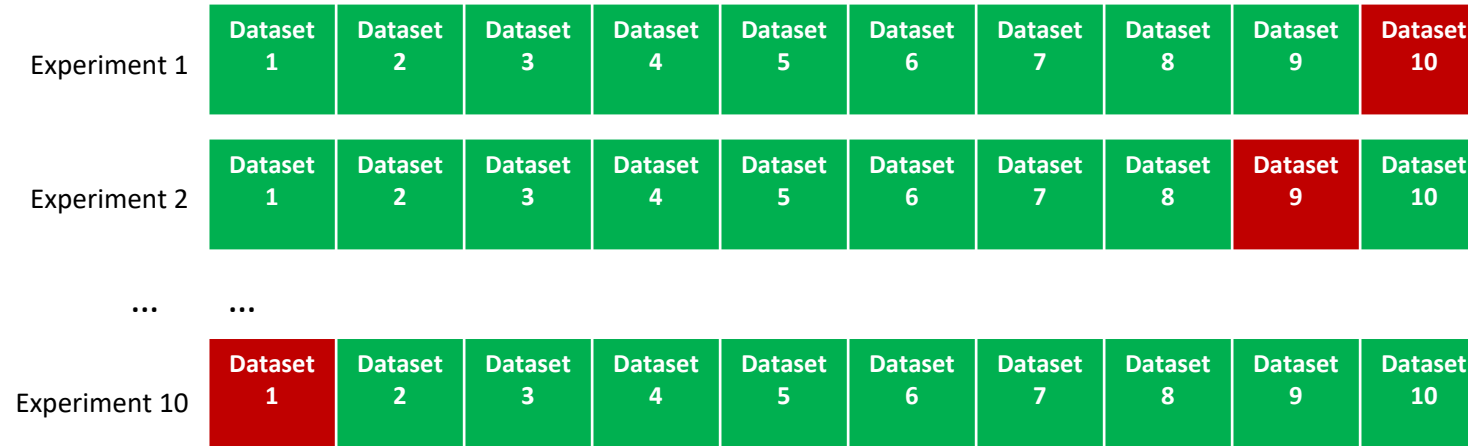


# Annotation Attributes (Feature Vectors) with Supervised ML

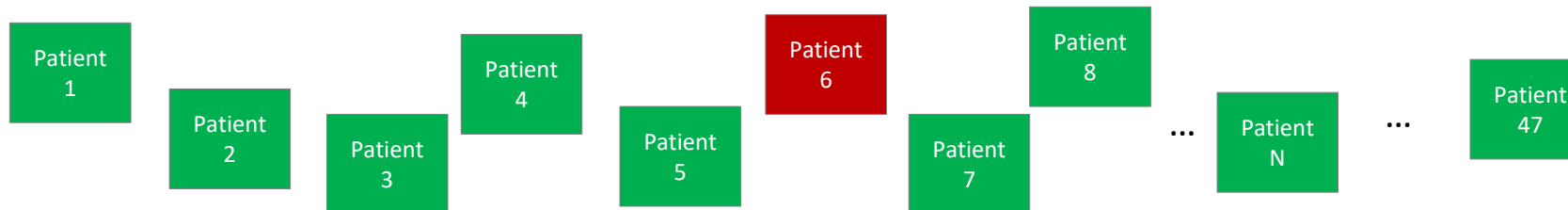




# Internal Testing/Validation

- 10 fold cross validation

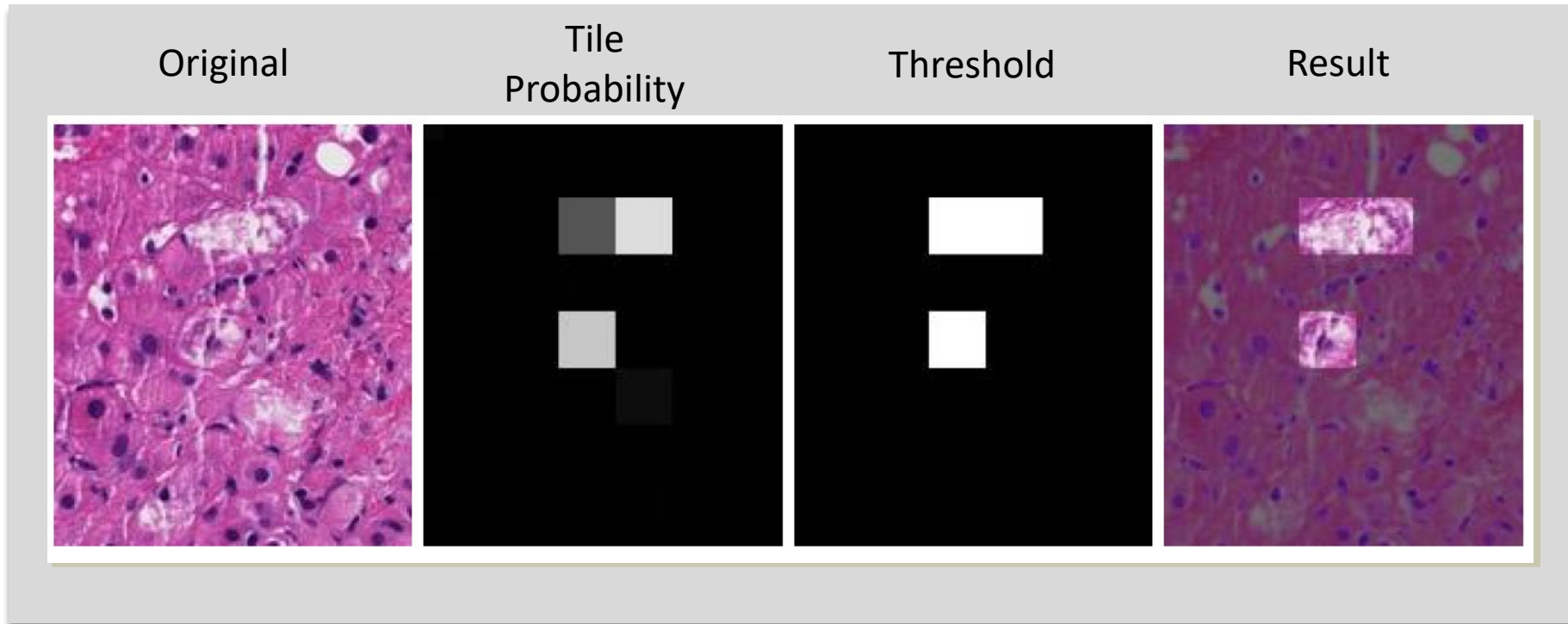


- Leave one out approach



 Data used to **LEARN**  
 Data used to **TEST**

# Hepatocyte Ballooning Example

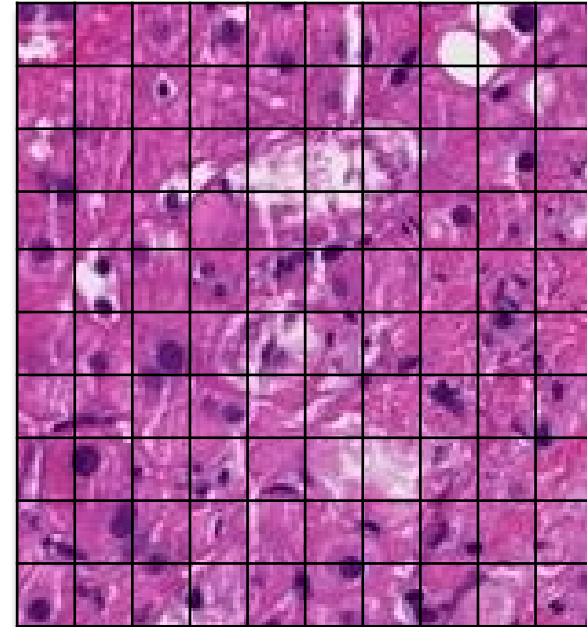


# Classification Approach

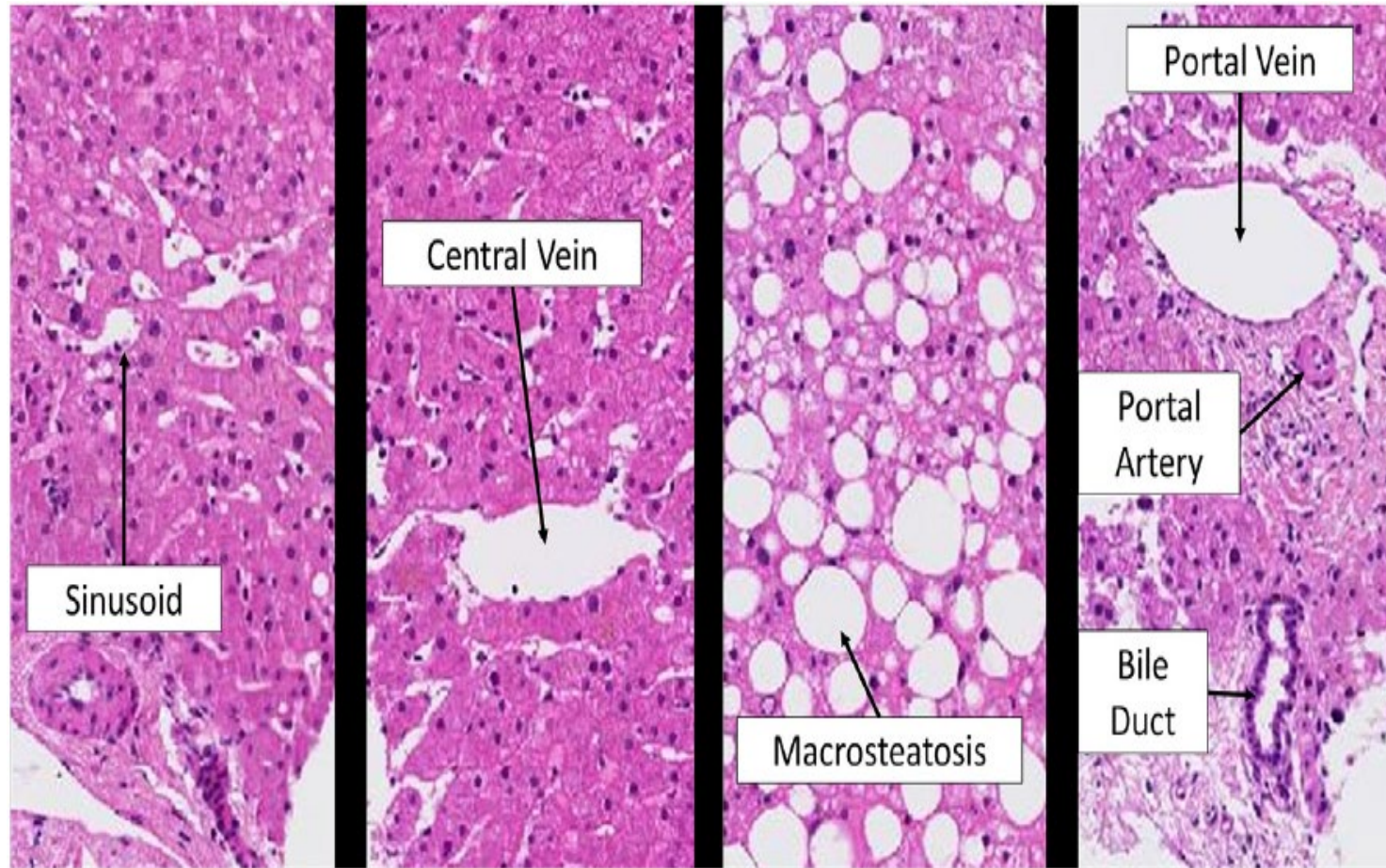
- Tile image into equal size pieces
- Classify each tile as either containing or NOT hepatocyte ballooning
- Calculate the total percent of tissue with hepatocyte ballooning

$$\% \text{ Ballooning} = \frac{\text{total area ballooning tiles}}{\text{total tissue area}}$$

- Similar procedure for other lesions



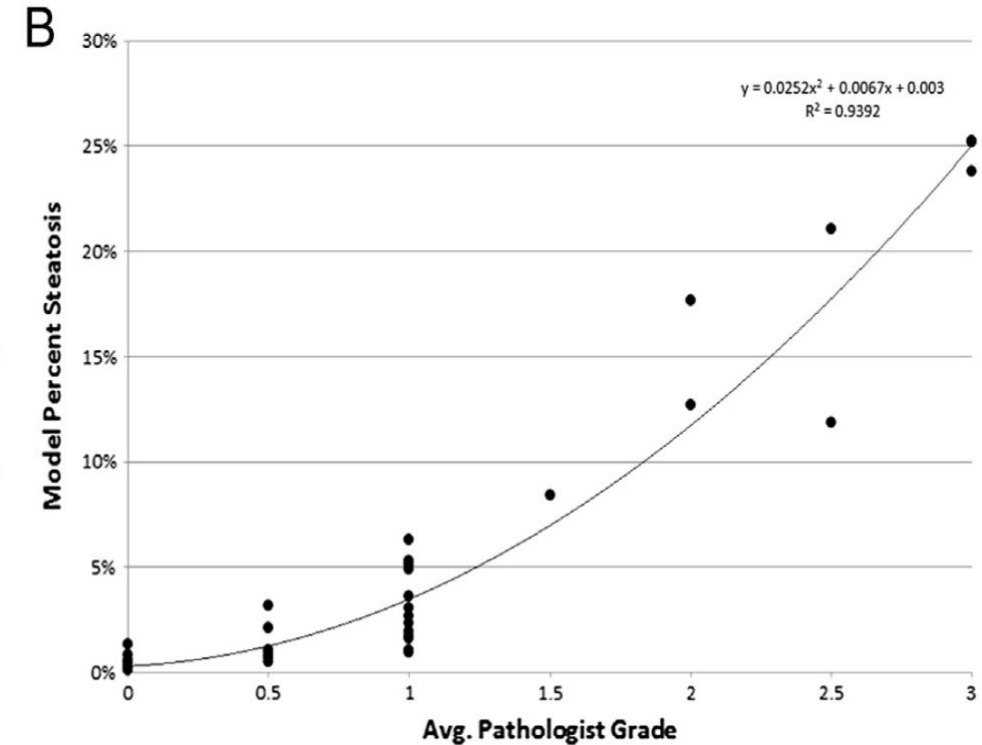
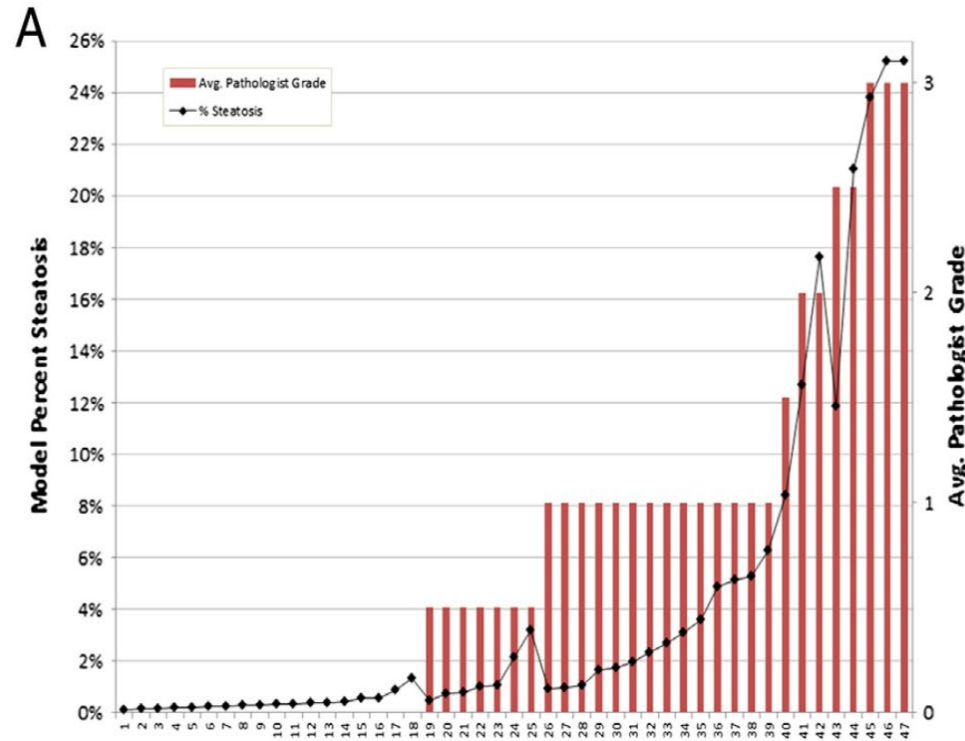
# White Regions on H&E Stained Liver Biopsy Images



# Automated Detection of Liver Microscopic Anatomic Landmarks

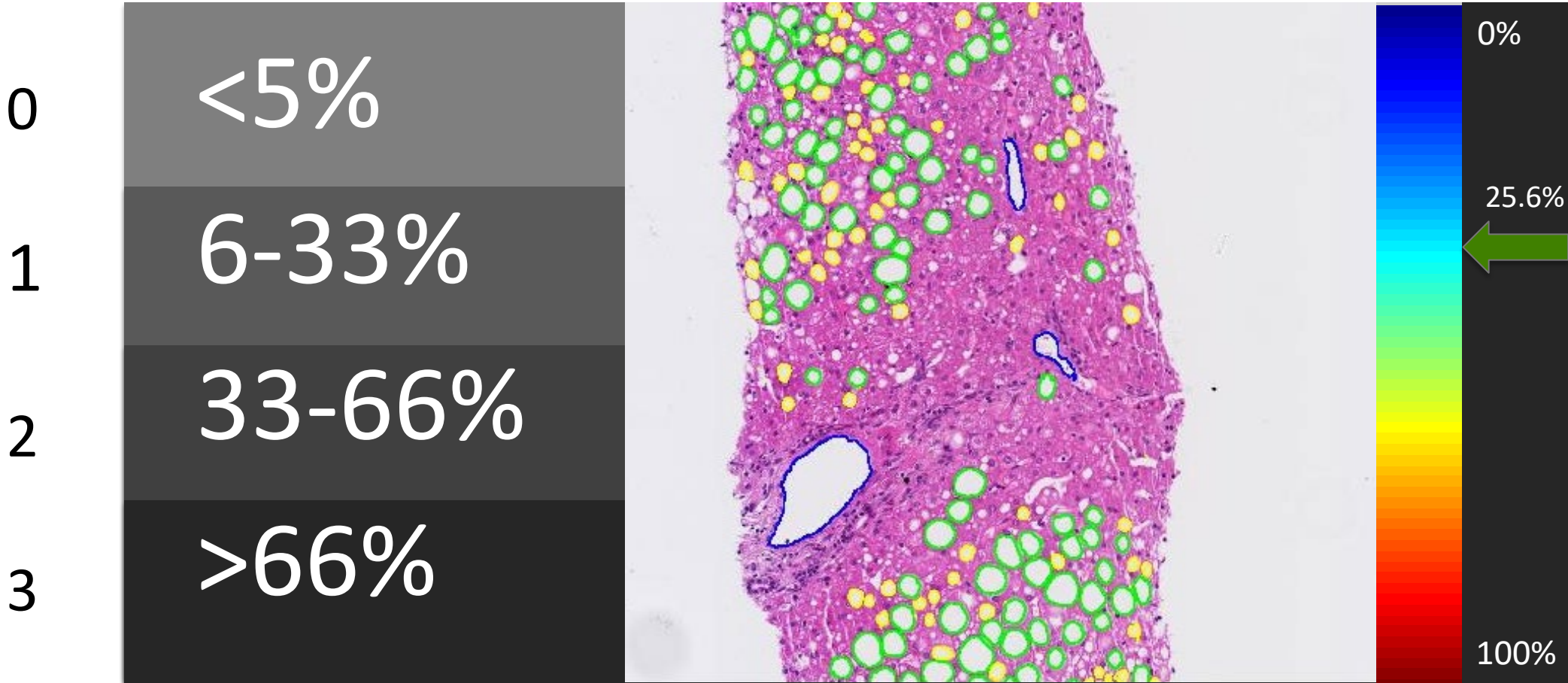
Feature	Pathologist (R.K.)			Pathologist (D.E.K.)			Combined model		
	Precision	Recall	ROC area	Precision	Recall	ROC area	Precision	Recall	ROC area
→ Bile duct	0.92	0.87	0.98	1.00	0.56	0.93	0.911	0.82	0.99
Central vein	0.64	0.79	0.92	0.67	0.64	0.83	0.615	0.63	0.83
Macrosteatosis	0.98	0.99	0.98	0.92	0.94	0.96	0.957	0.98	0.97
Other	1.00	0.86	0.99	0.76	0.69	0.90	0.860	0.63	0.94
Portal artery	0.85	0.77	0.97	0.67	0.18	0.94	0.667	0.59	0.96
→ Portal vein	0.91	0.77	0.97	0.81	0.88	0.97	0.825	0.84	0.97
Sinusoid	0.90	0.89	0.96	0.80	0.79	0.92	0.859	0.86	0.94

# Correlation and Relationship of Model Calculated Percent Steatosis with the Average of Pathologists Grade

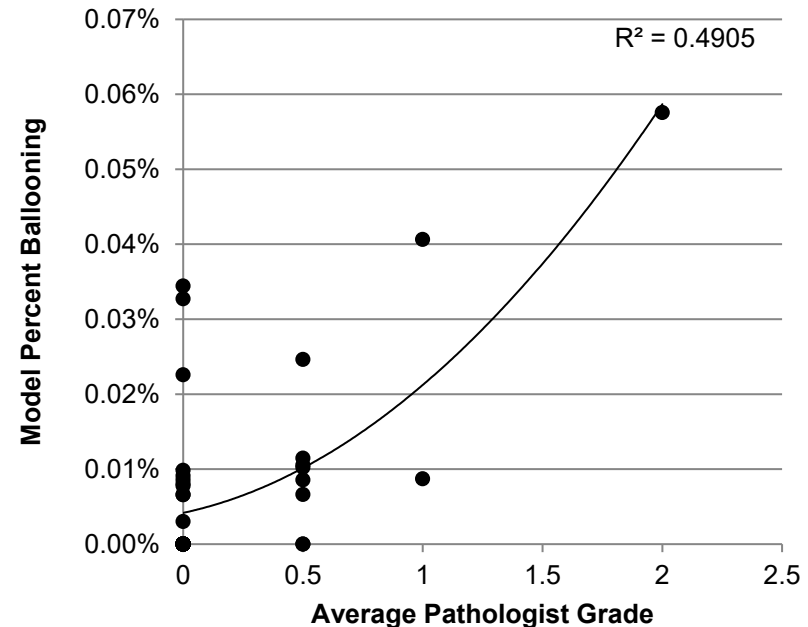
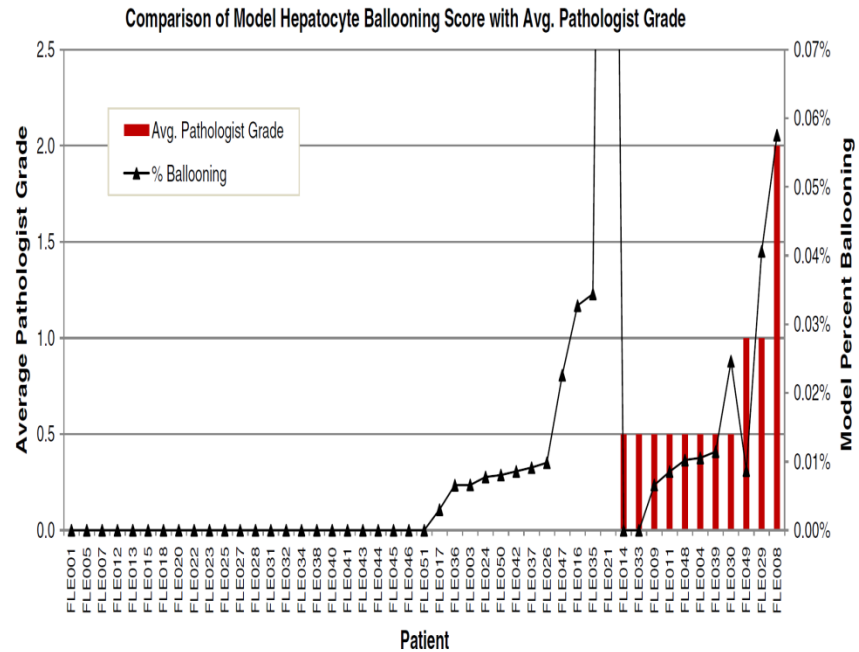




# Automated Continuous Quantification of Macrosteatosis

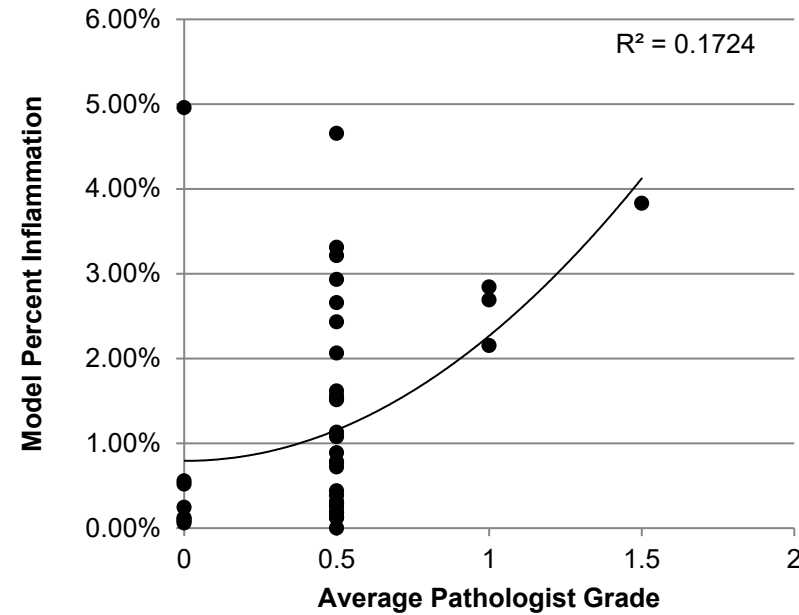
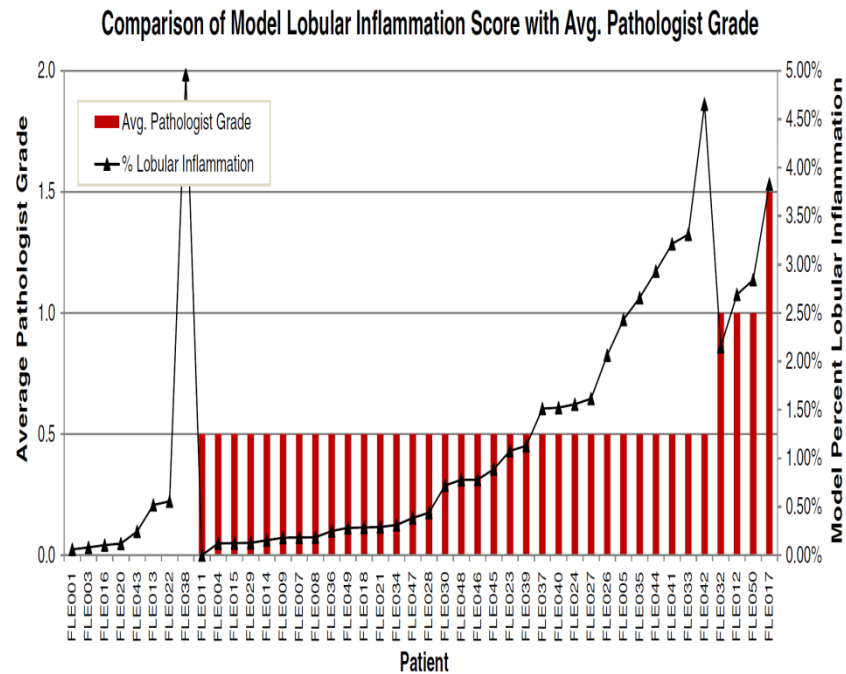


# Correlation of Model Calculated Percent *Hepatocyte Ballooning* with the Average of Pathologists Grade



Feature	Precision (PPV)	Recall (Sensitivity)	ROC Area
Hepatocyte Ballooning	<b>0.912</b>	0.542	0.983
Not-Hepatocyte Ballooning	0.990	0.999	0.983
<b>OVERALL</b>	<b>0.989</b>	<b>0.989</b>	<b>0.983</b>

# Correlation of Model Calculated Percent *Lobular Inflammation* with the Average of Pathologists Grade

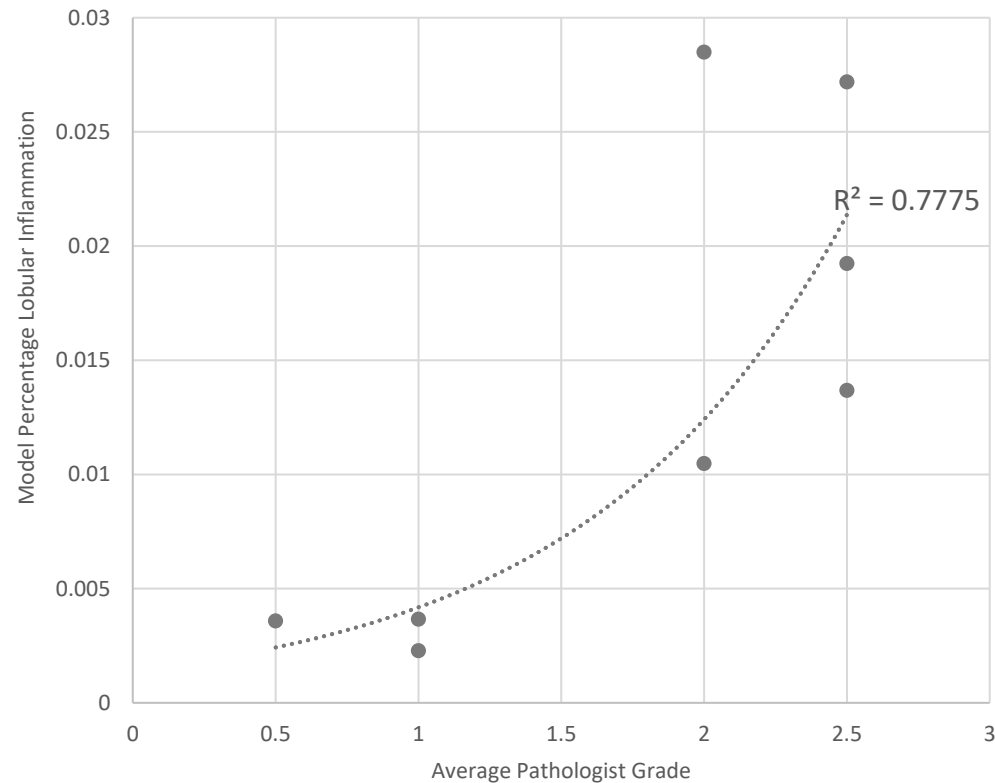


Feature	Precision (PPV)	Recall (Sensitivity)	ROC Area
Lobular Inflammation	0.696	0.489	0.946
Not-Lobular Inflammation	0.968	0.986	0.946
<b>OVERALL</b>	<b>0.952</b>	<b>0.956</b>	<b>0.946</b>

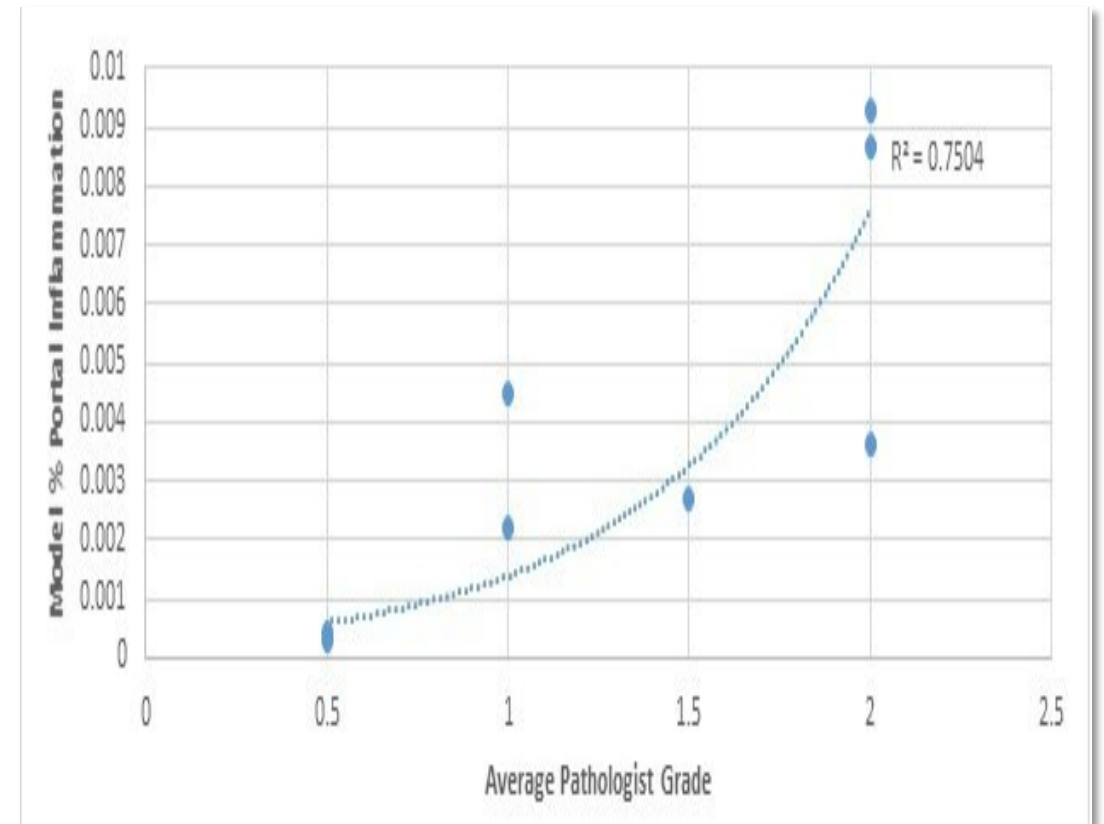
Vanderbeck S et al. Hum Pathol 2015

# Refined Model for **Lobular** Inflammation and New Model for **Portal** Inflammation in Human NAFLD

**AUROC – 97.4%, Precision – 79.3%,  
Sensitivity – 81.3%,**



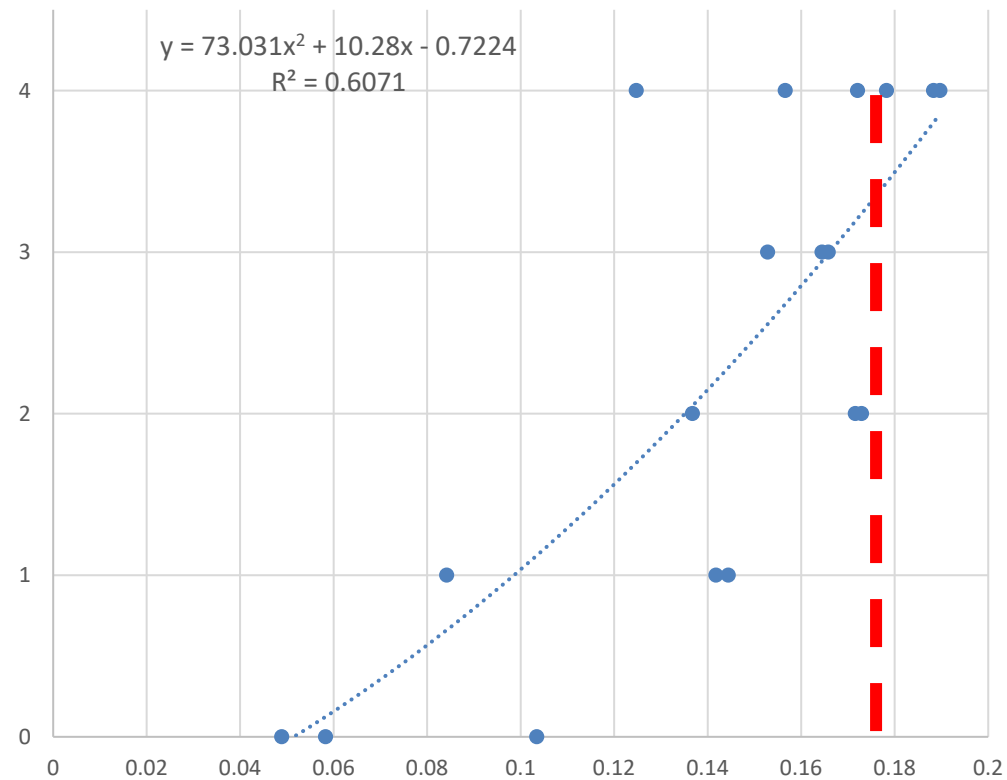
**AUROC – 97.9%, Precision – 82.1%,  
Sensitivity – 88.3%**



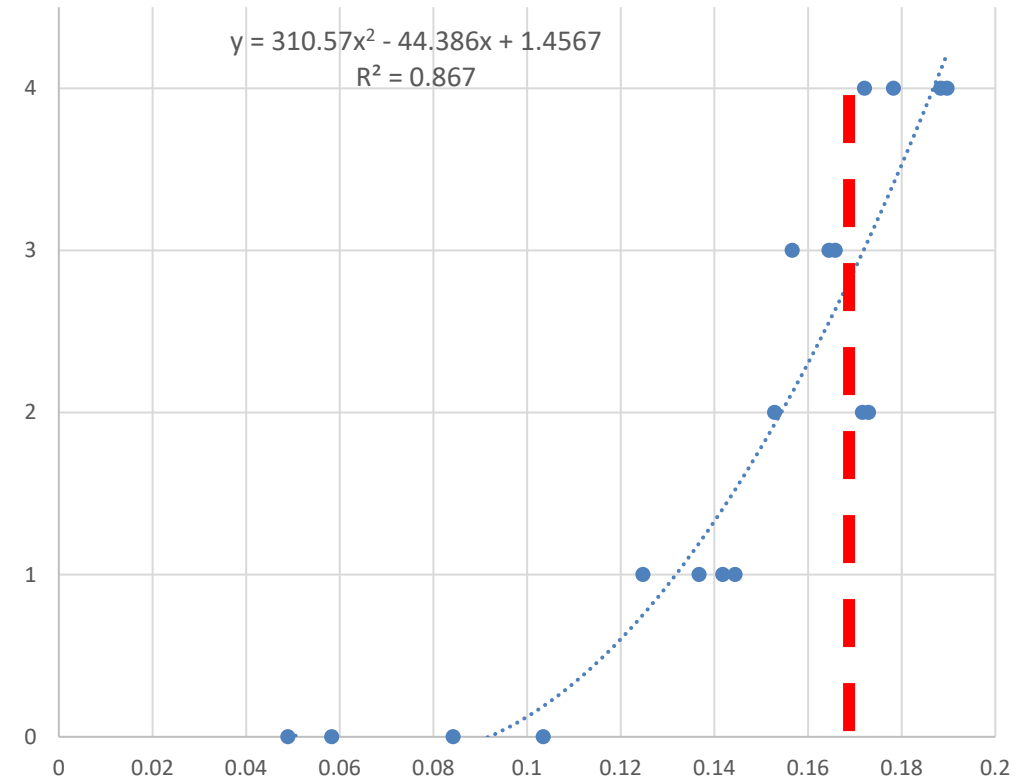
Unpublished data

# Automated Fibrosis Assessment in Human NAFLD

CPA Correlation with Pathologist DEK



CPA Correlation with Pathologist OWC

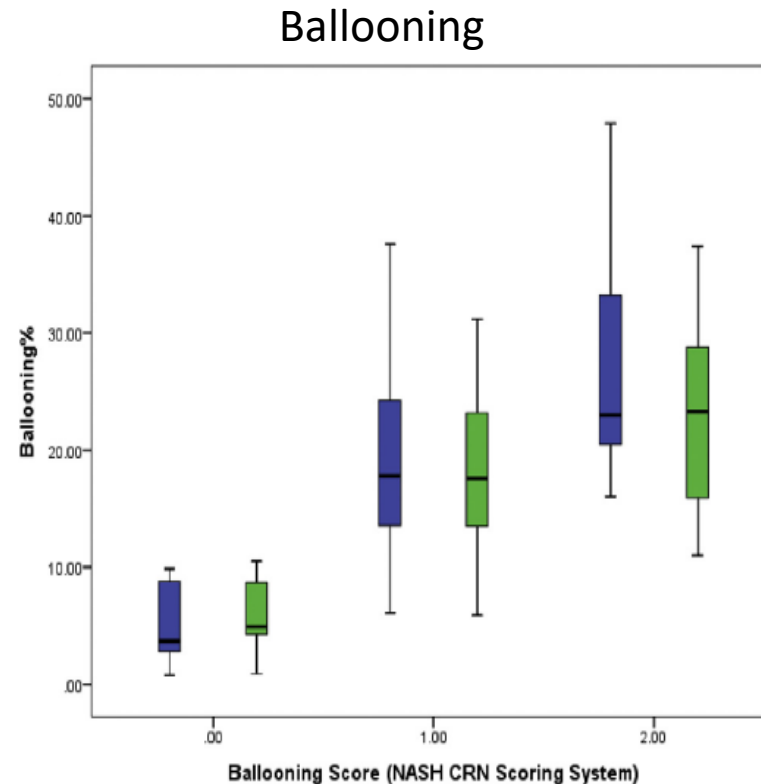
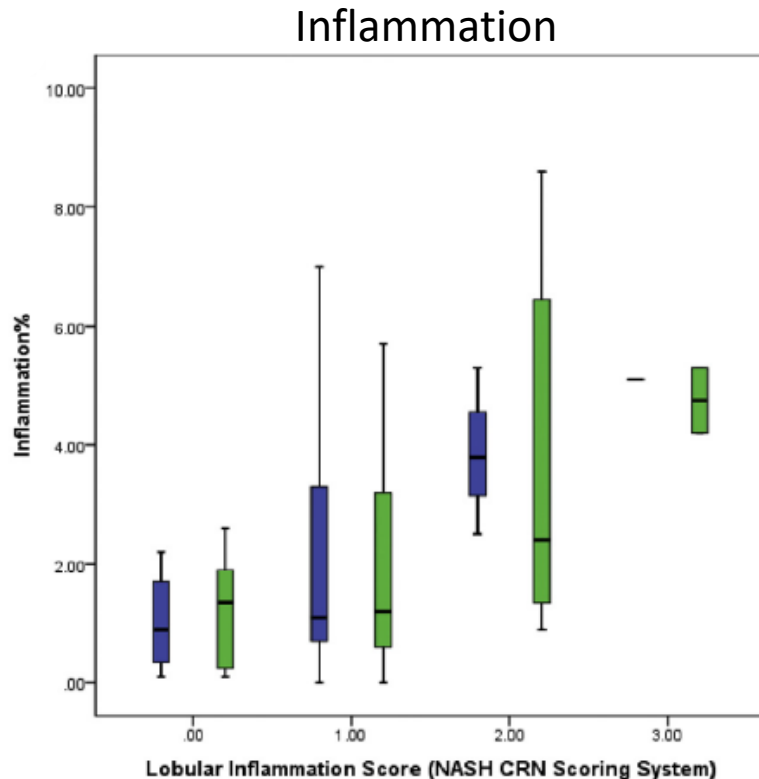


# Automated Identification of Architectural Type of Liver Fibrosis in Human NAFLD Liver Biopsies

<b>Fibrosis Type</b>	<b>Precision (%)</b>	<b>Recall (%)</b>	<b>AUROC (%)</b>
<b>Normal</b>	85.6	83.3	91.9
<b>Pericellular</b>	76.6	82.7	83.3
<b>Periportal</b>	72.1	76.9	78.6
<b>Portal</b>	77	84.4	86.4
<b>Bridging</b>	84.9	91.7	93
<b>Nodule</b>	89.8	91.6	95.4

# Machine Learning for Automated NAFLD Histology Assessment

246 patients with NAFLD (**190 with NASH**, 56 with simple steatosis)



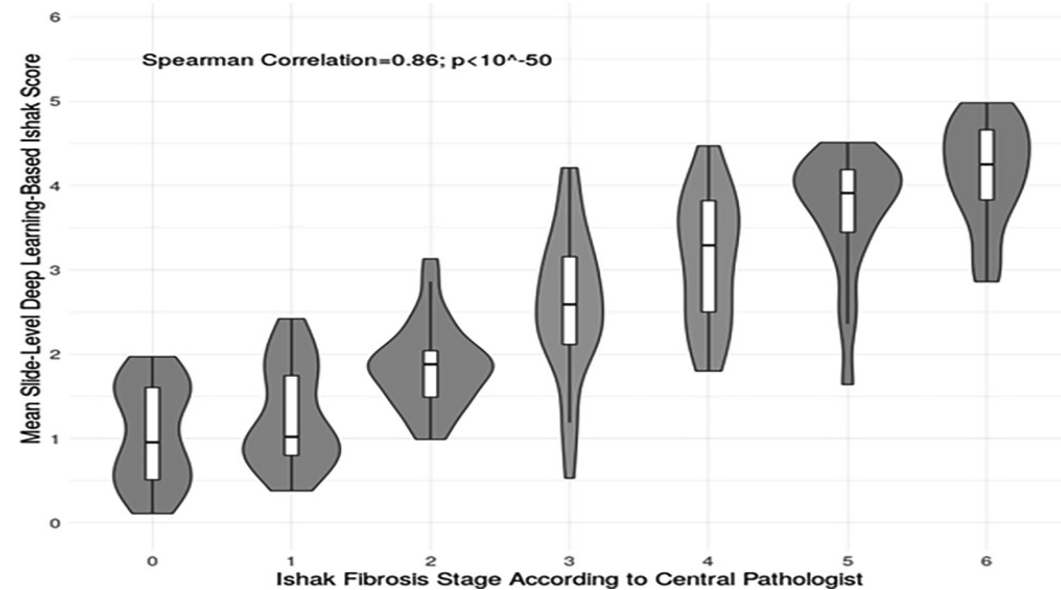
### Correlation (rho) with Pathologist's scores

- Steatosis: 0.66
- Inflammation: 0.36
- Ballooning: 0.52
- Fibrosis: 0.57

**AUROC for CPA for F $\geq$ 3: 0.82**

# Deep Machine Learning for Automated NASH Histology Assessment

- 834 liver biopsies from subjects screened for a phase 3 trial of selonsertib (STELLAR-4) (advanced cirrhotic NASH)
- CNN with over 20 layers and 8 million parameters using over 68,000 annotations collected from 75 board-certified pathologists



## Correlations ( $\rho$ ) with average of two pathologists scores

Steatosis: 0.86

Inflammation: 0.56

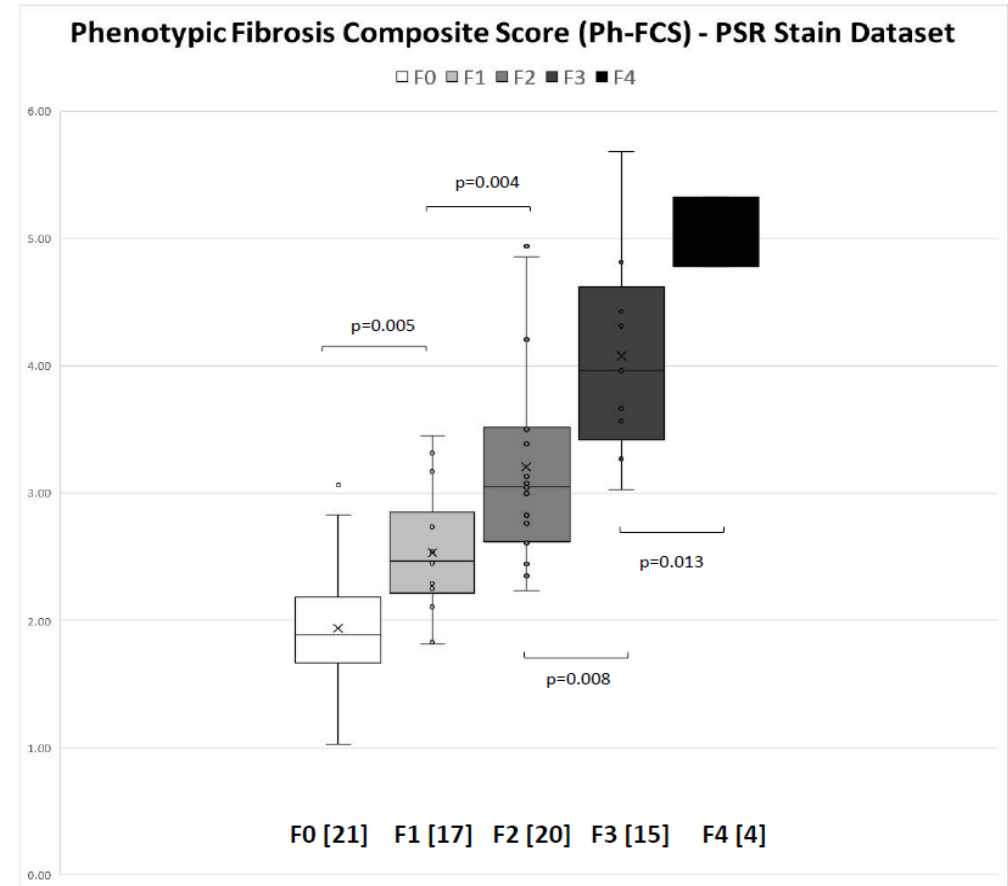
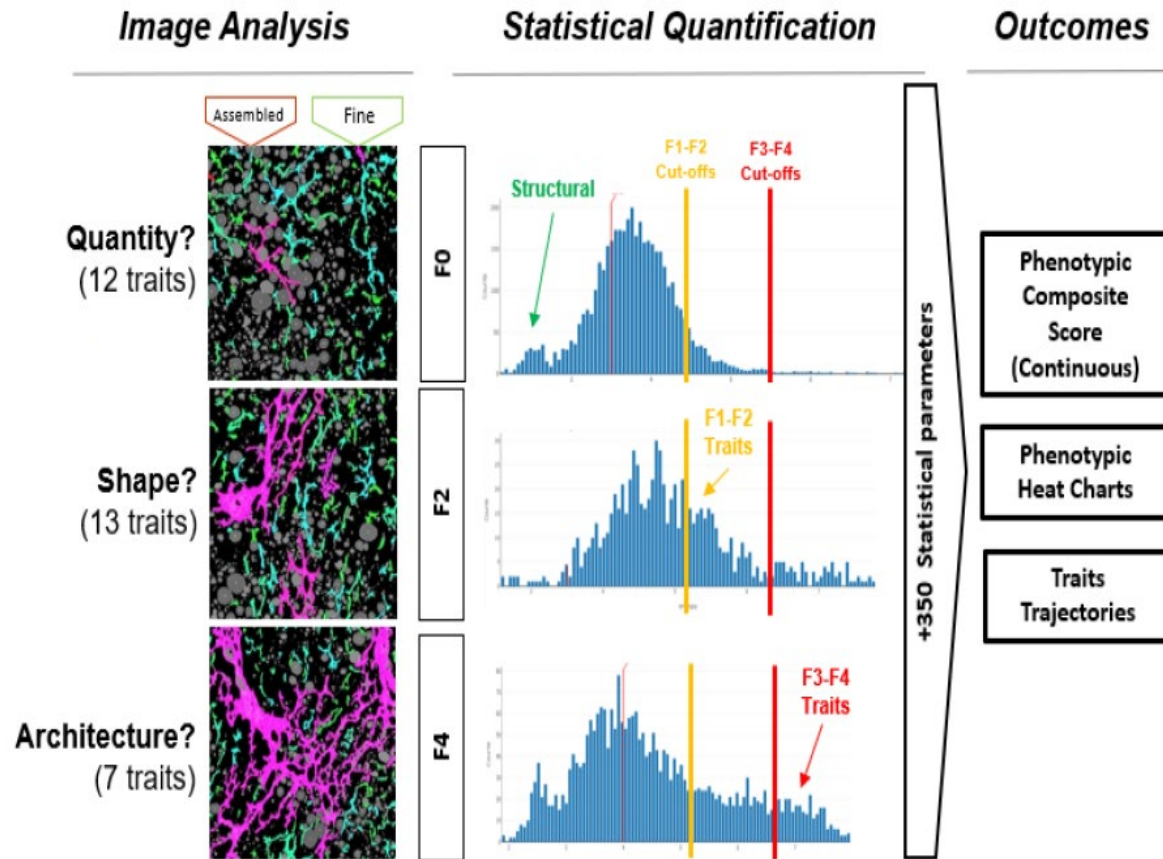
Ballooning 0.68

Fibrosis 0.83 (CRN), 0.86 (Ishak)



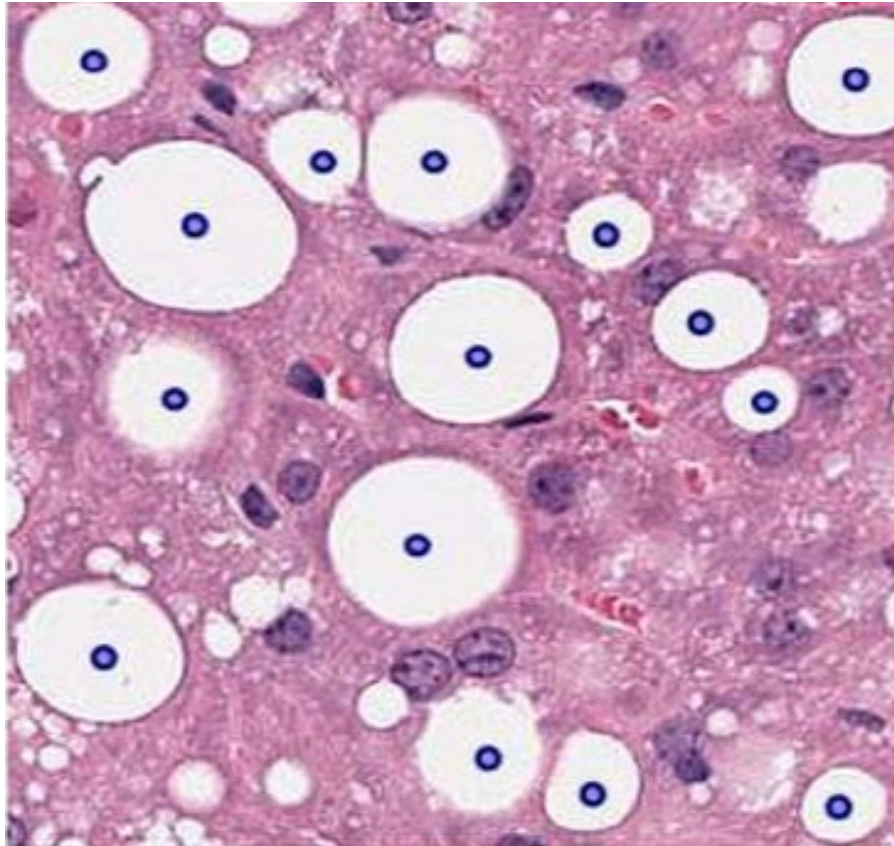
# Phenotypic Fibrosis Composite Score by FibroNest Image Analysis Platform

77 NASH Biopsies; unstained-SHG/TPEF or stained by Sirius red or trichrome-digitally scanned



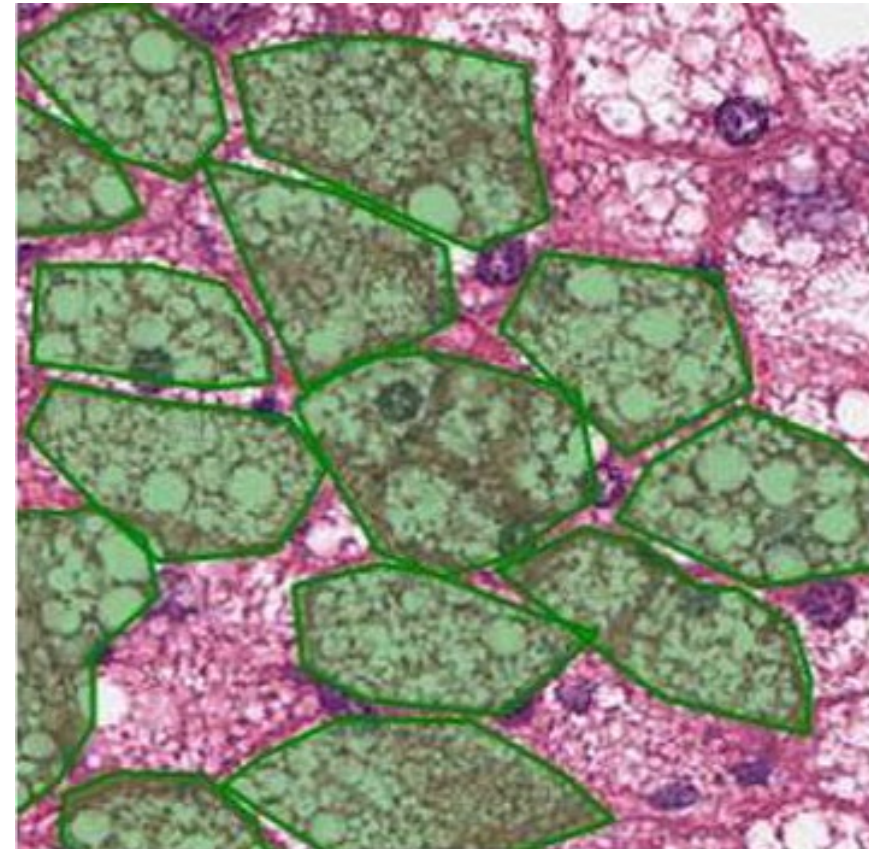
# External Validation of Steatosis Classifiers in Murine NAFLD

Macrosteatosis



Accuracy of labels: 100%

Microsteatosis



63%

Sethunath D et al. Plos1 2018

# Automated algorithm- and SHG microscopy-based assessment of NAFLD histology (qFIBS)

	AUROC	95% CI	P Value	Cutoff (Youden's Index)	Sensitivity	Specificity	PPV	NPV
<b>qFibrosis</b>								
F0 vs. F≥1	0.870	0.787-0.953	<0.001	0.761	94%	63%	84%	83%
F≤1 vs. F≥2	0.881	0.804-0.959	<0.001	0.882	97%	58%	65%	96%
F≤2 vs. F≥3	0.945	0.891-0.999	<0.001	1.491	96%	76%	66%	97%
F≤3 vs. F4	0.951	0.905-0.996	<0.001	2.395	87%	91%	72%	96%
<b>qInflammation</b>								
0 vs. ≥1	0.838	0.752-0.924	0.105	1.251	83%	100%	100%	14%
≤1 vs. ≥2	0.820	0.726-0.913	<0.001	1.357	93%	58%	58%	93%
≤2 vs. 3	0.831	0.729-0.933	0.112	1.503	100%	79%	12%	100%
<b>qBallooning</b>								
0 vs. ≥1	0.844	0.731-0.957	0.011	1.086	71%	100%	100%	20%
≤1 vs. 2	0.813	0.708-0.918	<0.001	1.266	60%	89%	67%	85%
<b>qSteatosis</b>								
0 vs. ≥1	0.986	0.959-1.000	<0.001	0.796	99%	100%	100%	50%
≤1 vs. ≥2	0.948	0.903-0.993	<0.001	1.572	91%	85%	83%	92%
≤2 vs. 3	0.939	0.867-1.000	<0.001	2.210	67%	98%	86%	95%

## Correlation (r) with Pathologist's scores

- qSteatosis: 0.802
- qInflammation: 0.557
- qBallooning: 0.533
- qFibrosis: 0.776

# Considerations for Development of Automated Methods for NAFLD Histology Analysis

- Minimum acceptable standards for liver biopsy sample
- Derivation biopsy cohort
  - Minimizing bias in biopsy selection: Representation of the entire histological spectrum of NAFLD
  - Large number of annotations by expert NAFLD pathologists to train the model (Different severity/typical/atypical variety of each lesion)
  - Number of expert pathologists involved? 1 vs more, observer agreement, other factors...
- Trade offs of how thresholds are set {High sensitivity/High specificity/Optimal(Youden)}



# Considerations for Development of Automated Methods for NAFLD Histology Analysis

- External validation
  - Verification by experts pathologists of accuracy of lesion identification on unseen biopsy images
  - Validation of performance in different cohorts
- ? Weight of strength of correlation with semi-quantitative assessments/scores
- Lack of explainability factor for deep learning networks and regulatory approvals:
  - Unknown what attributes of a lesion are used/contribute to decision making process in nodes/networks (Black box factor)
- These tools may be viewed as complimentary decision aids/guides, not replacements, to pathologists



# Current and Future States of NAFLD Histology Analysis

## Current

- Manual
- Semi-quantitative data
- Limited scale
- Intra- & inter-observer variability
- Limited pool of experienced NAFLD pathologists
- Limited access to experienced NAFLD pathologists

## Future: Optimize/Maximize Extracted Data

- Automated
- Continuous data
- Large scale
- Precise
- Reproducible
- Available
- Accessible

**Pathologists are key partners in leading us through this transformation of the field**



# Acknowledgements

- Pathologists
  - David Kleiner, MD, PhD (National Cancer Institute)
  - Richard Komorowski, MD (Medical College of Wisconsin)
  - Oscar W. Cummings, MD (Indiana University School of Medicine)
- Computer Engineers
  - Scott Vanderbeck, M.S. and Joe Bockhurst, PhD (University of Wisconsin – Milwaukee)
  - Mihran Tuceryan, PhD (Indiana University-Purdue University-Indianapolis)

