Complexity of ballooned hepatocyte feature recognition: Defining a training atlas for artificial intelligence-based imaging in NAFLD

Liver Forum, April 22, 2022

Elizabeth M Brunt, MD

Emeritus Professor, Pathology and Immunology
Washington University School of Medicine
St Louis, MO

Brunt et al, J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Research Article
NAFLD and Alcohol-Related Liver Diseases

JOURNAL
OF HEPATOLOGY

Complexity of ballooned hepatocyte feature recognition: Defining a training atlas for artificial intelligence-based imaging in NAFLD

Elizabeth M. Brunt[1,*,†], Andrew D. Clouston[2], Zachary Goodman[3], Cynthia Guy[4], David E. Kleiner[5], Carolin Lackner[6], Dina G. Tiniakos[7,8], Aileen Wee[9], Matthew Yeh[10], Wei Qiang Leow[11], Elaine Chng[12], Yayun Ren[12], George Goh Boon Bee[13], Elizabeth E. Powell[14,15], Mary Rinella[16], Arun J. Sanyal[17], Brent Neuschwander-Tetri[18], Zobair Younossi[19], Michael Charlton[20], Vlad Ratziu[21], Stephen A. Harrison[22,23], Dean Tai[11,*,†], Quentin M. Anstee[7,24,*,†]

[1]Department of Pathology and Immunology, Washington University School of Medicine, Saint Louis, Missouri, USA; [2]Molecular and Cellular Pathology, University of Queensland and Envoi Specialist Pathologists, Brisbane, Australia; [3]Pathology Department, and Center for Liver Diseases, Inova Fairfax Hospital, Falls Church, Virginia, USA; [4]Division of Pathology, Duke University Medical Center, Durham, NC, USA; [5]Laboratory of Pathology; Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA; [6]Institute of Pathology, Medical University of Graz, Graz, Austria; [7]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK; [8]Dept of Pathology, Aretaieion Hospital, National and Kapodistrian University of Athens, Greece; [9]Department of Pathology, Yong Loo Lin School of Medicine, National University of Singapore, National University Hospital, Singapore; [10]Department of Pathology, University of Washington, Seattle, Washington, USA; [11]Department of Anatomical Pathology, Singapore General Hospital, Singapore & Duke-NUS Medical School, Singapore; [12]HistoIndex Pte Ltd, Singapore; [13]Department of Gastroenterology and Hepatology, Singapore General Hospital, Singapore; [14]Centre for Liver Disease Research, Faculty of Medicine, University of Queensland, Translational Research Institute, Brisbane, Queensland, Australia; [15]Department of Gastroenterology and Hepatology, Princess Alexandra Hospital, Brisbane, Queensland, Australia; [16]Division of Gastroenterology and Hepatology, Feinberg School of Medicine, Northwestern University, Chicago, USA; [17]Department of Internal Medicine, School of Medicine, Virginia Commonwealth University, Richmond, Virginia, USA; [18]Division of Gastroenterology and Hepatology, Saint Louis University, Saint Louis, Missouri, USA; [19]Betty and Guy Beatty Center for Integrated Research, Inova Health System, Falls Church, Virginia, USA; [20]Center for Liver Diseases, and Transplantation Institute, University of Chicago, Chicago, Illinois, USA; [21]Department of Hepatology, Sorbonne University and Pitié-Salpêtrière Hospital, Paris, France; [22]Pinnacle Clinical Research, San Antonio, USA; [23]Hepatology, Radcliffe Department of Medicine, University of Oxford, Oxford, UK; [24]Newcastle NIHR Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

Brunt et al, J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

COI

Advisory Board: Alnylam/Regeneron, Pfizer.

Consultancy agreements:
Arrowhead, Cymabay, Histoindex, Intercept, Medpace, NGM, Perspectum Diagnostics.

Special thanks to

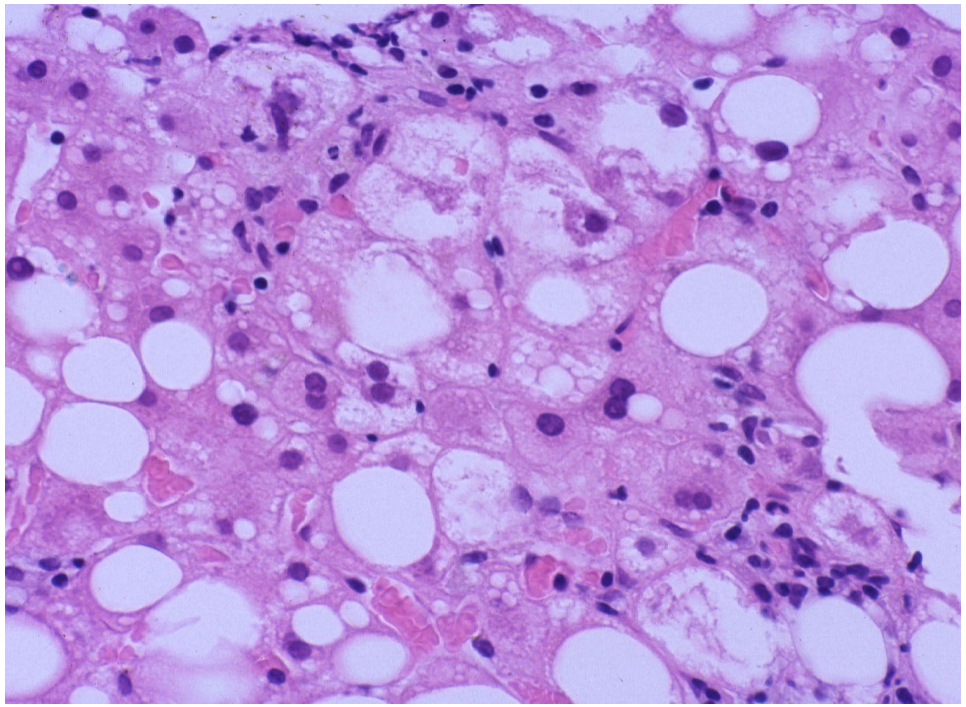
- The **amazing pathologists** who did this work
- **Quentin Anstee**: stats and intellectual (hepatology) input
- **Dean Tai**: stats and intellectual (engineer) input
  - Both for sharing some of these slides
- **Clinical hepatology colleagues**

Current Concept

Necessary for dx of ***NASH***

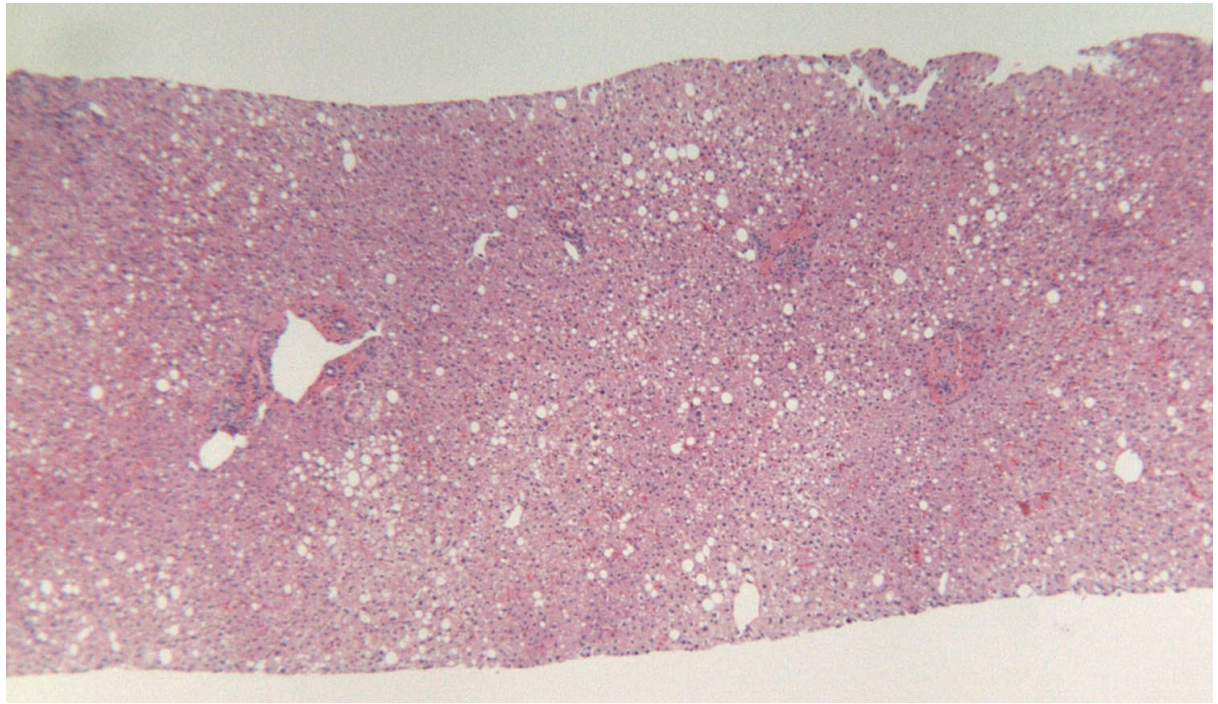
NOT NASH
NASH Resolution

Ballooning 1-2



Steatosis, Inflammation, +/- Fibrosis

Ballooning 0

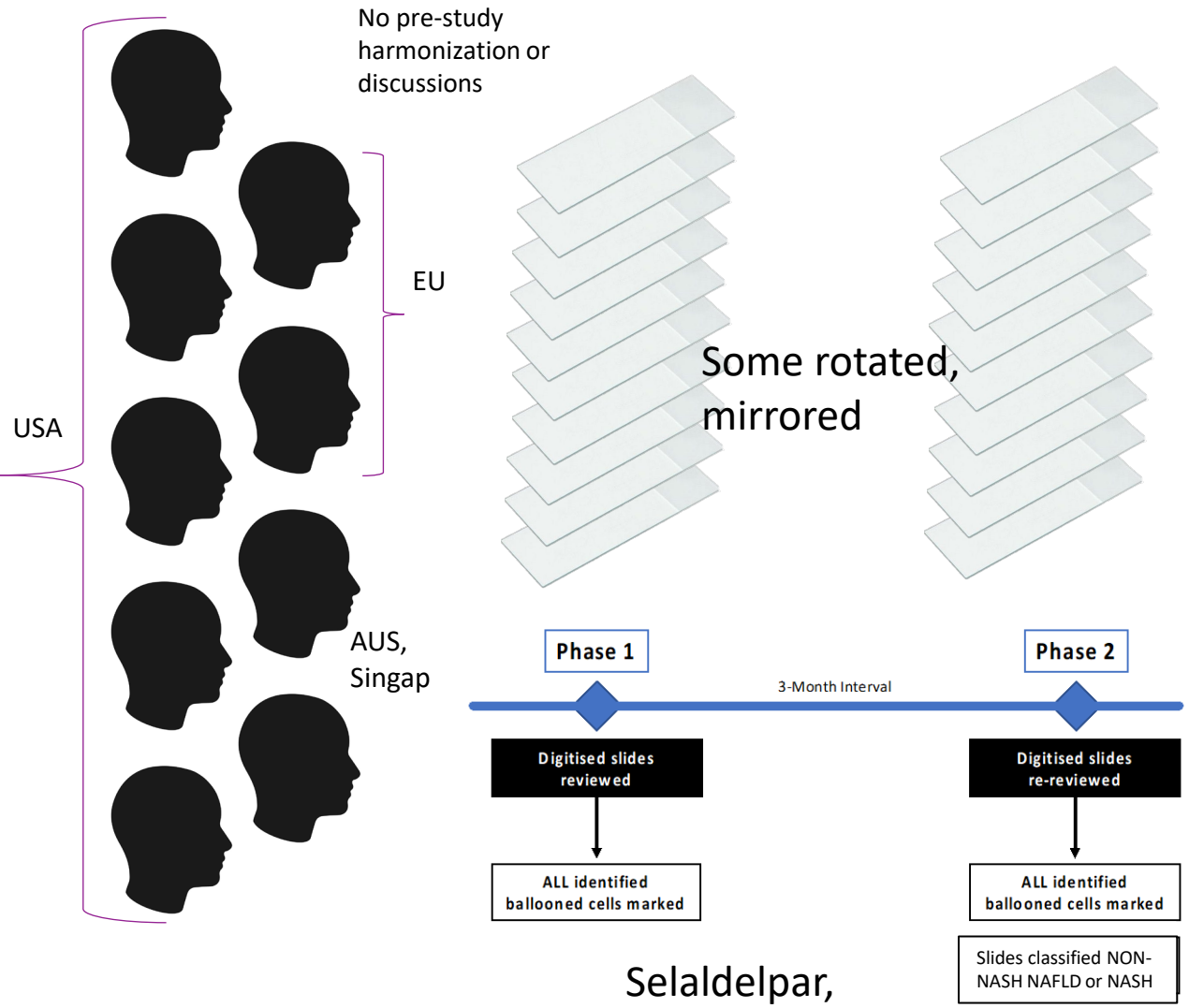


Steatosis, Inflammation, +/- Fibrosis

Background

- Increase in clinical trials in NASH → questions:
  - Reproducibility of pathologists' interpretations on "routine stains"
    - Ballooning
    - Diagnosis: NASH
    - Features for NASH Resolution
  - Role of Machine learning/AI for "consistency"
    - ? Replace pathologists
    - ? Guide pathologists

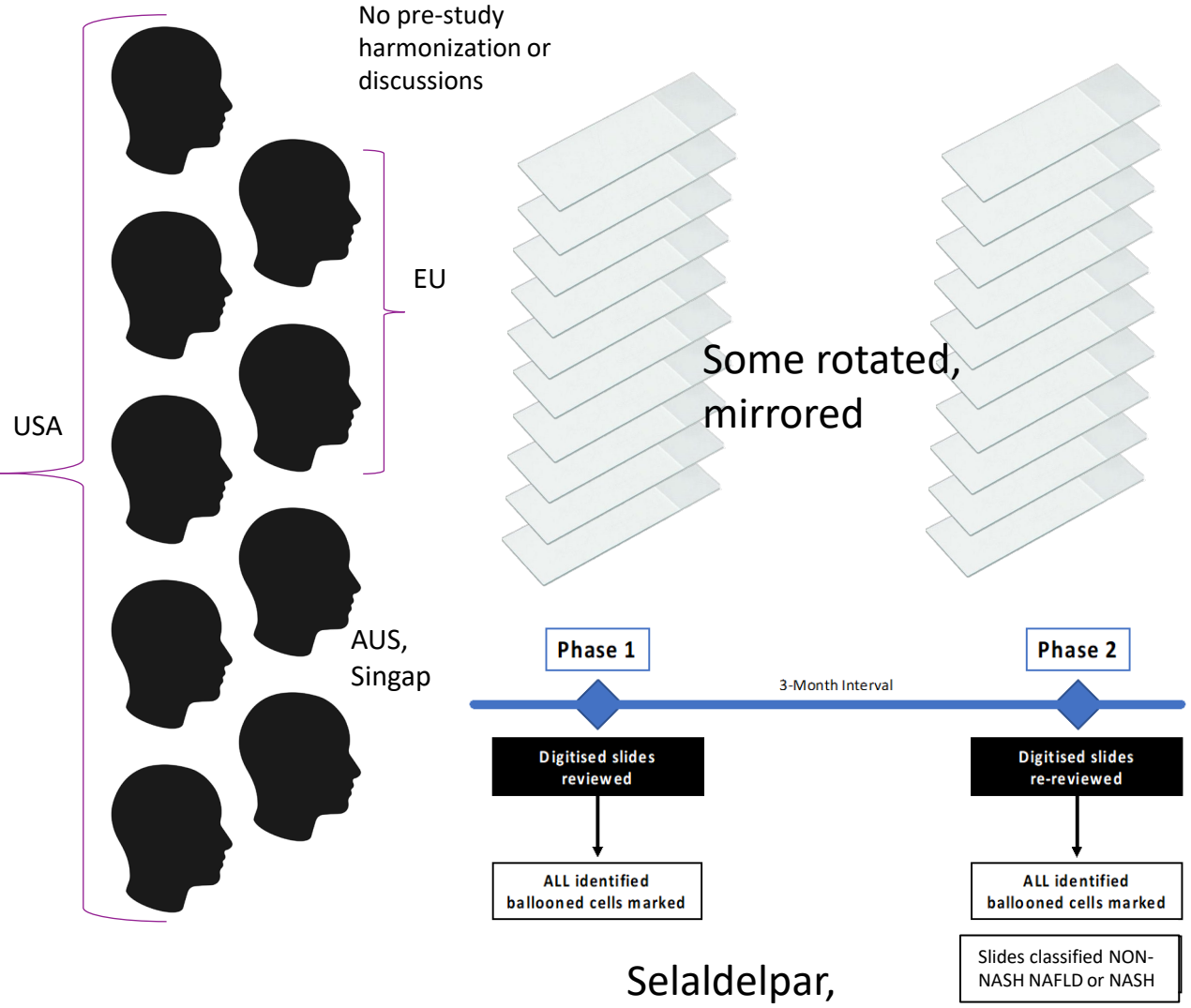
9 International Expert Hepatopathologists ❯ **Independent Annotation of All Ballooned Cells in 10 Digitised Liver Biopsy Images** ❯



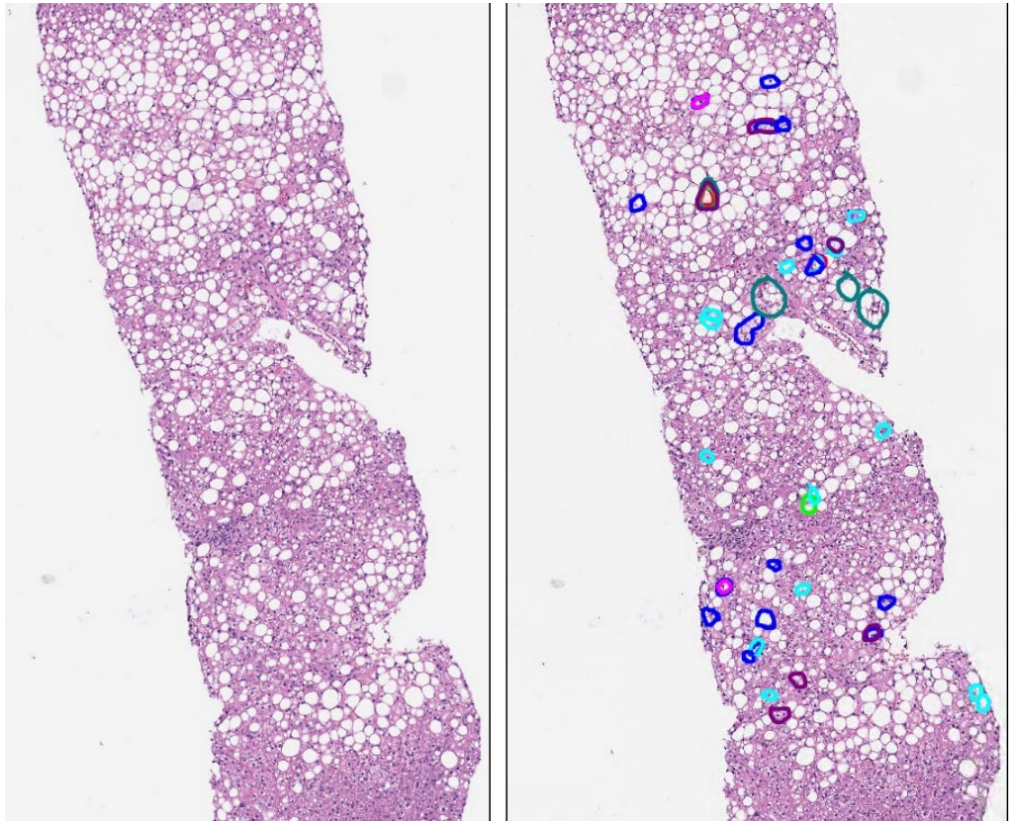
Selaldelpar, [NCT03551522]

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

9 International Expert Hepatopathologists ❯ **Independent Annotation of All Ballooned Cells in 10 Digitised Liver Biopsy Images** ❯ **Establish a 'Concordance Atlas' of Annotated Ballooned Cell Images** ❯ **Training of SHG/TPE AI Algorithm**



Selaldelpar, [NCT03551522]



J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

M/M

- ROI of digitized slides chosen (by EMB) to
  - Equalize amount of liver tissue/slide for each bx
  - Represent the spectrum of NAFLD; steatosis (**B0**) → active steatohepatitis (**B1-B2**)
  - Represent the spectrum of slide quality as seen in clinical trials

- Agreement statistics: 3 binary conditions
  - Presence of ANY ballooned hepatocytes
  - Presence of ≥ 5 ballooned hepatocytes
  - Non-NASH v NASH diagnosis

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Inter-Observer Concordance between Pathologists for Number of Ballooned Cells Identified.

B0 = no ballooning

Range of ballooned hepatocytes observed after rotation of slides by pathologists: **32%-91%**

"NO BALLOONING"

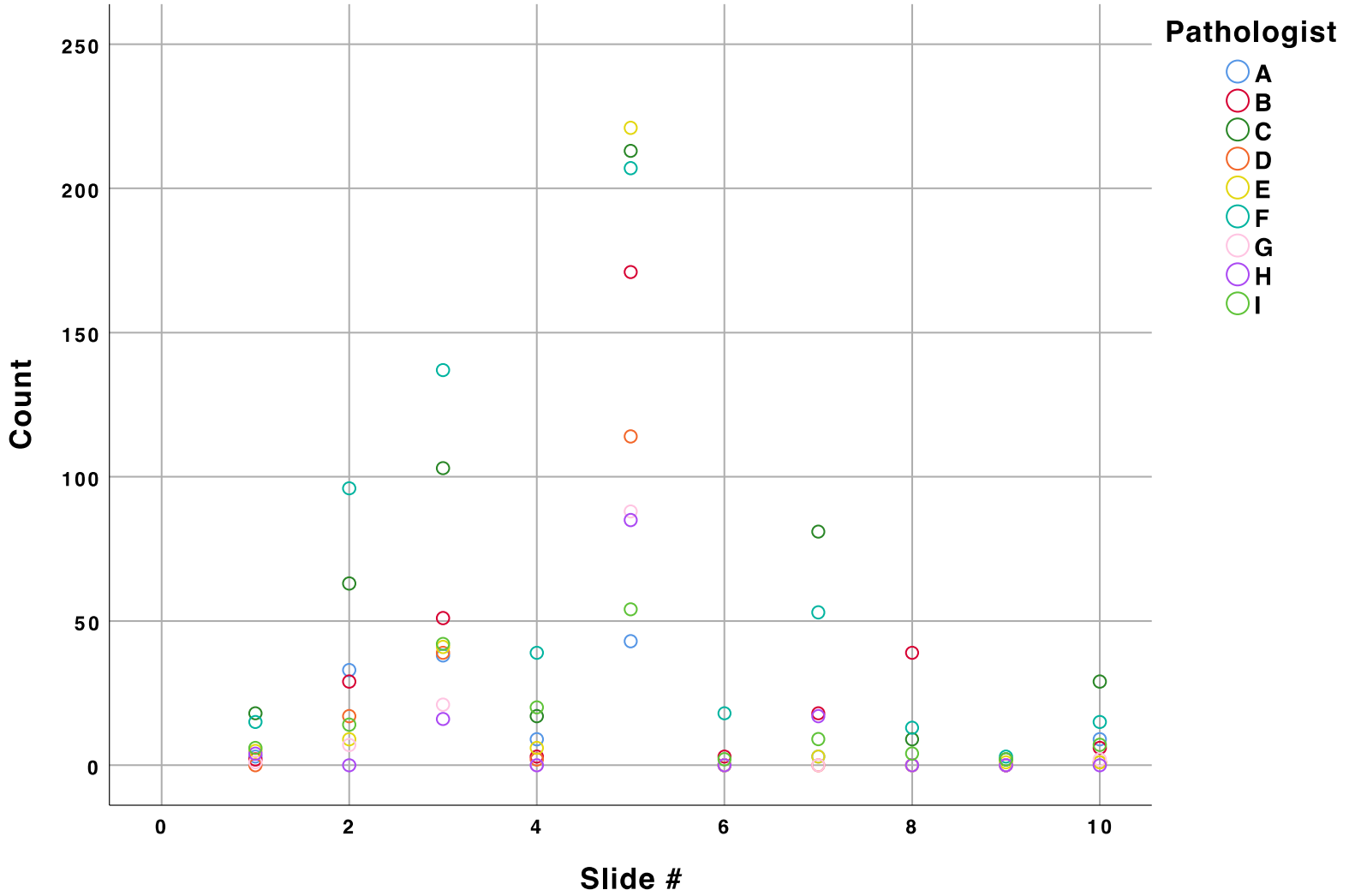
-**2** pathologists agreed **B0: #4, #7**

-**5** pathologists agreed **B0: #8, #9**

-**6** pathologists agreed **B0: #6**

-At least one pathologist recorded B0 for all slides except #3, #5

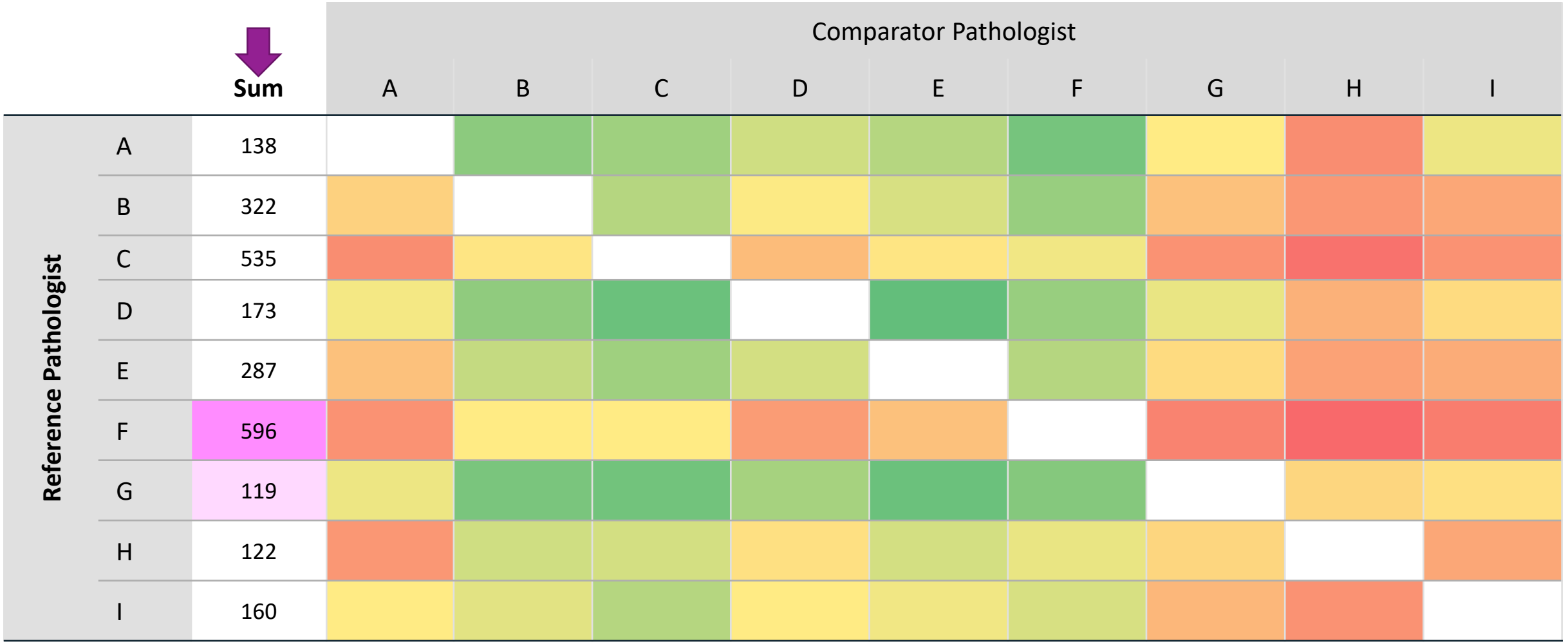
-All pathologists except two (H, I) recorded B0 at least once



Intraclass correlation coefficient for consistency: 0.718 (0.511-0.900): moderate for ballooning burden

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Inter-Observer Concordance between Pathologists for Ballooned Cells Identified.



Mean: 173

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Inter-Observer Concordance between Pathologists for Ballooned Cells Identified.

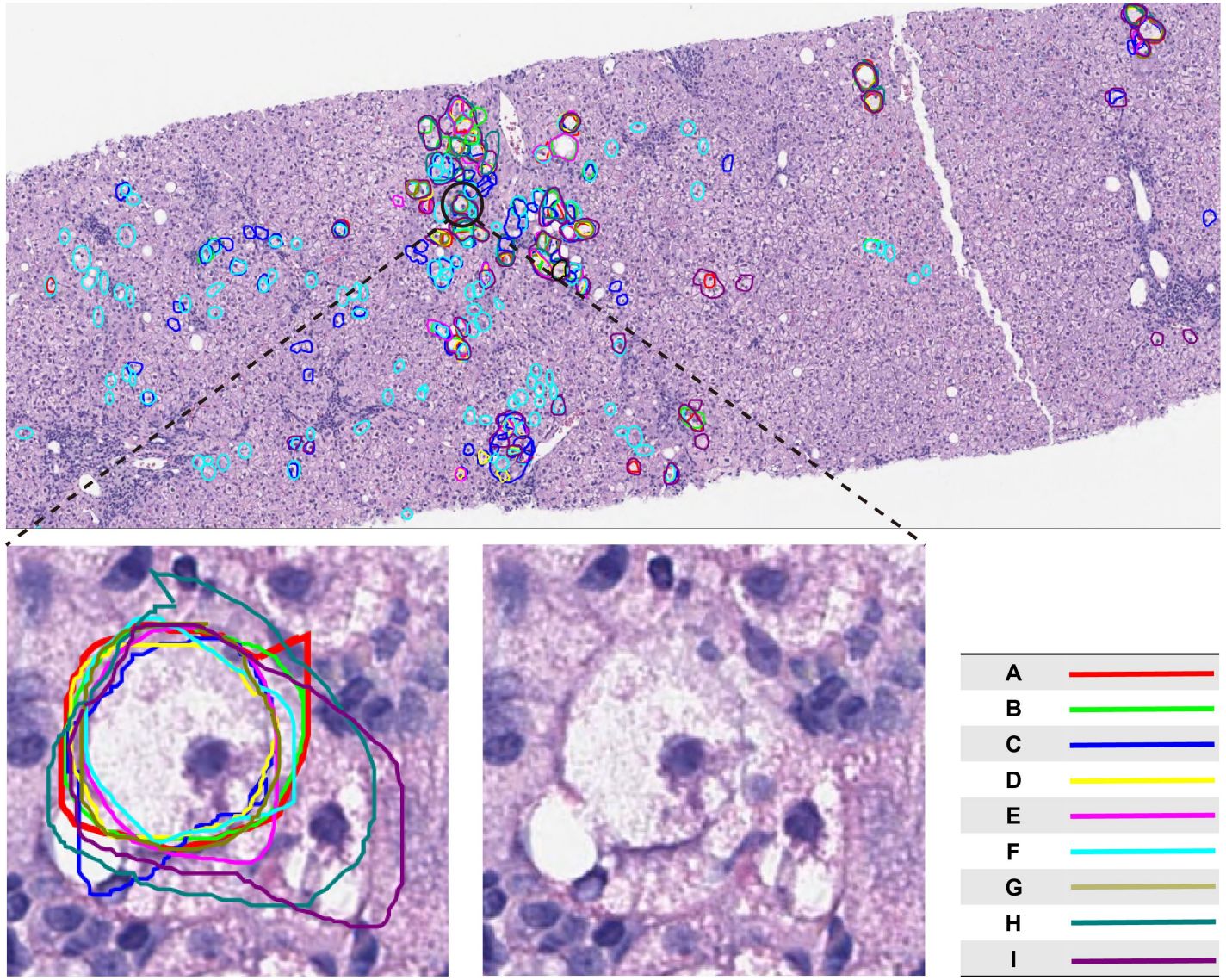
Reference Pathologist		Sum	Comparator Pathologist A	B	C	D	E	F	G	H	I
	A	138	-	89 (64%)	81 (59%)	63 (46%)	73 (53%)	97 (70%)	45 (33%)	21 (15%)	52 (38%)
	B	322	89 (28%)	-	171 (53%)	109 (34%)	142 (44%)	196 (61%)	82 (25%)	56 (17%)	65 (20%)
	C	535	81 (15%)	171 (32%)	-	126 (24%)	169 (32%)	197 (37%)	84 (16%)	55 (10%)	85 (16%)
	D	173	63 (36%)	109 (63%)	126 (73%)	-	130 (75%)	106 (61%)	68 (39%)	38 (22%)	52 (30%)
	E	287	73 (25%)	142 (49%)	169 (59%)	130 (45%)	-	151 (53%)	87 (30%)	55 (19%)	59 (21%)
	F	596	97 (16%)	196 (33%)	197 (33%)	106 (18%)	151 (25%)	-	79 (13%)	47 (08%)	71 (12%)
	G	119	45 (38%)	82 (69%)	84 (71%)	68 (57%)	87 (73%)	79 (66%)	-	35 (29%)	37 (31%)
	H	122	21 (17%)	56 (46%)	55 (45%)	38 (31%)	55 (45%)	47 (39%)	35 (29%)	-	25 (20%)
	I	160	52 (33%)	65 (41%)	85 (53%)	52 (33%)	59 (37%)	71 (44%)	37 (23%)	25 (16%)	-

Heatmap: green is high agreement of comparator to reference pathologist for # of ballooned cells; red is low

Variation: 8%-75%; kappa: 0.197; 0.395 for >5 cells

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

> 8,000 cells evaluated (by counting nuclei); 1188 ballooned cells circled.



The **ONLY** ballooned hepatocyte upon which all 9 agreed

Inter-Observer Concordance between Pathologists for Number of Ballooned Cells Identified: Trends Identified

Item Statistics

	Mean	Std. Deviation	N
A	13.80	17.184	10
B	32.20	51.873	10
C	53.50	66.185	10
D	17.30	36.237	10
E	28.70	68.668	10
F	59.60	67.160	10
G	11.90	27.534	10
H	12.20	26.452	10
I	16.00	17.969	10

Propensity to “see” ballooning:

More	**Fewer**
F > C	G, H

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Inter-Observer Concordance between Pathologists for Number of Ballooned Cells Identified: Trends Identified

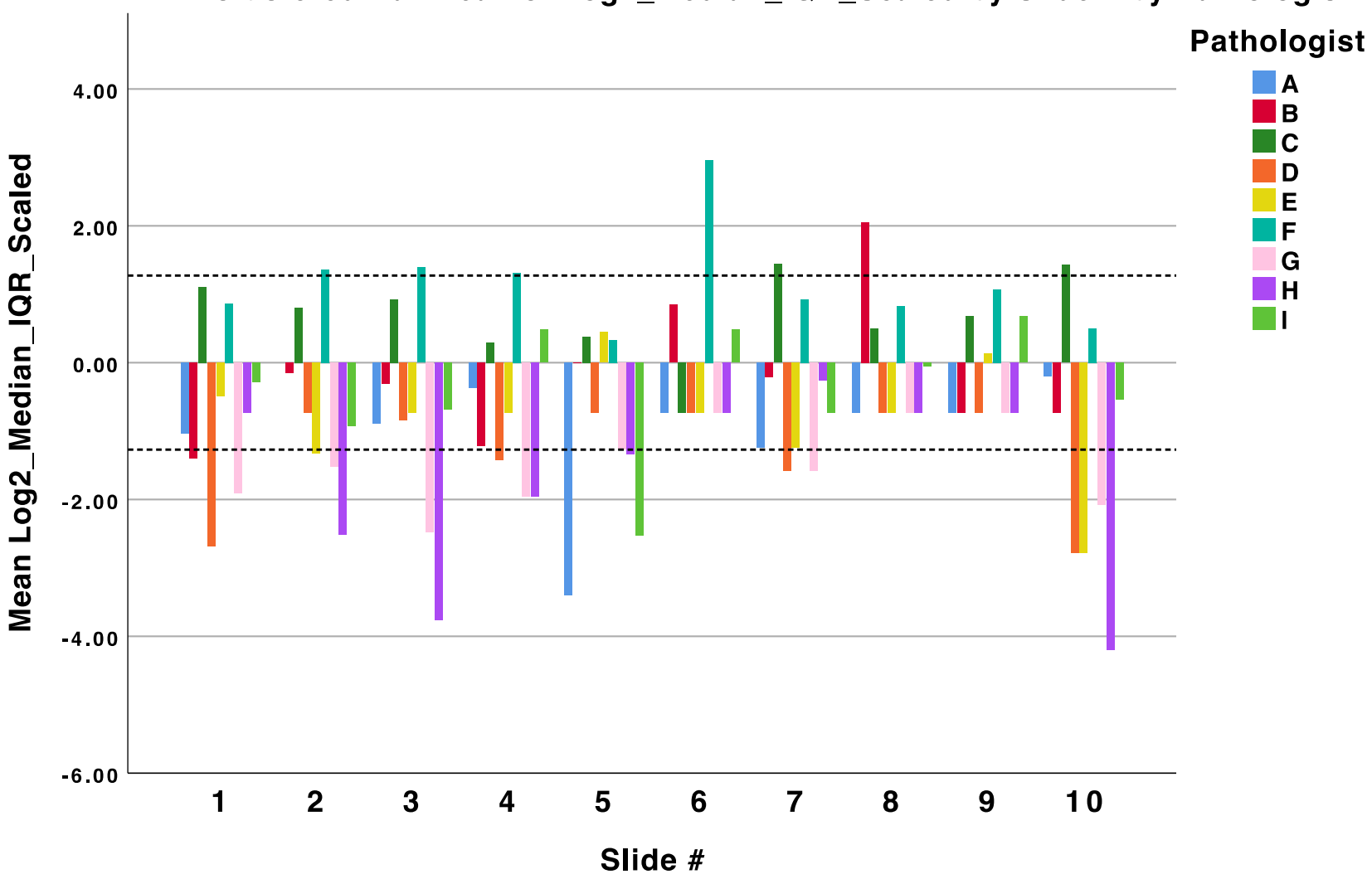
Not related to country or prior work together

Item Statistics

	Mean	Std. Deviation	N
A	13.80	17.184	10
B	32.20	51.873	10
C	53.50	66.185	10
D	17.30	36.237	10
E	28.70	68.668	10
F	59.60	67.160	10
G	11.90	27.534	10
H	12.20	26.452	10
I	16.00	17.969	10

Propensity to “see” ballooning:

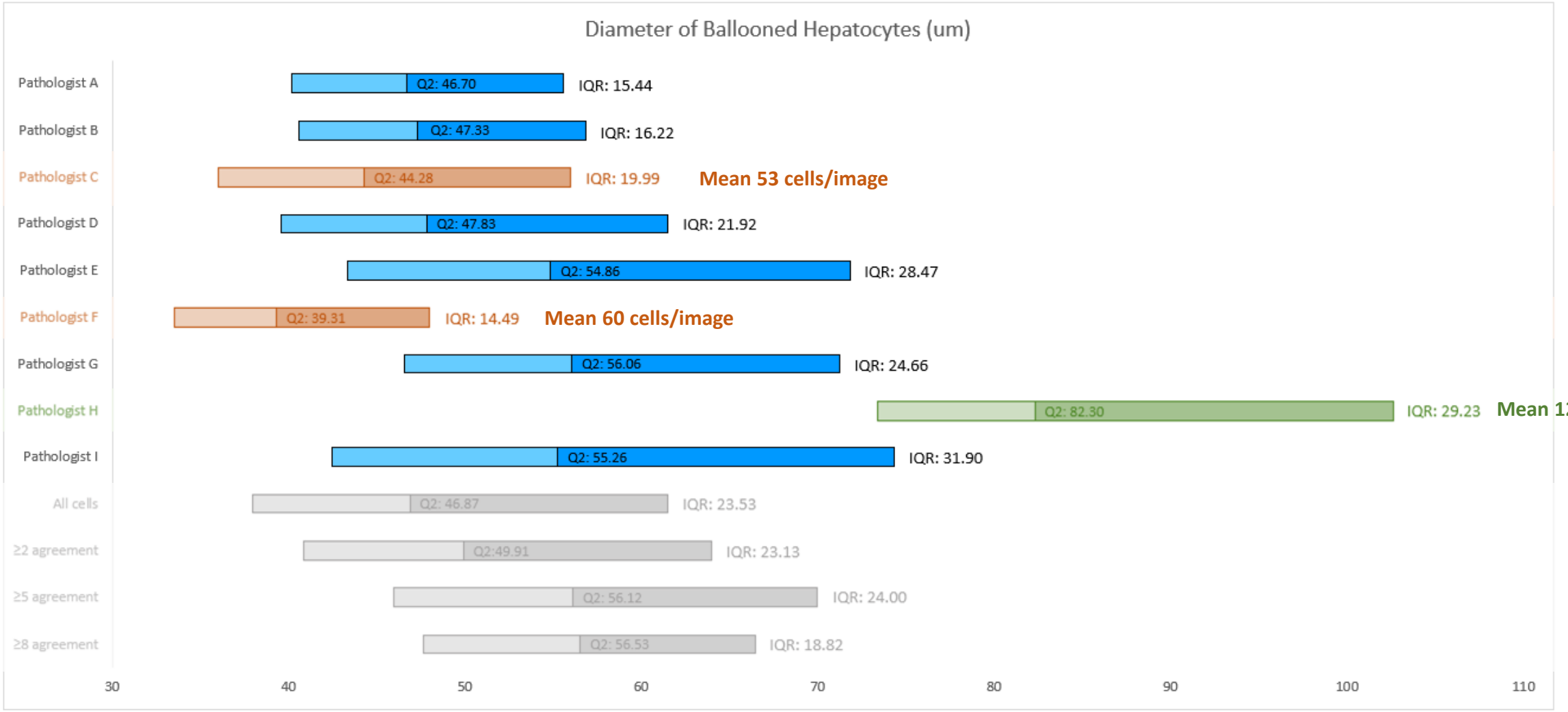
More	**Fewer**
F > C	G, H



Log2 transformed graph showing Median-IQR Robust Scaled Cell Count. Longer lines imply greater deviation by given pathologist from group median for each slide. Dotted lines 1 SD.

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Ballooned hepatocyte diameter by Pathologist: Is this the key to the differences?

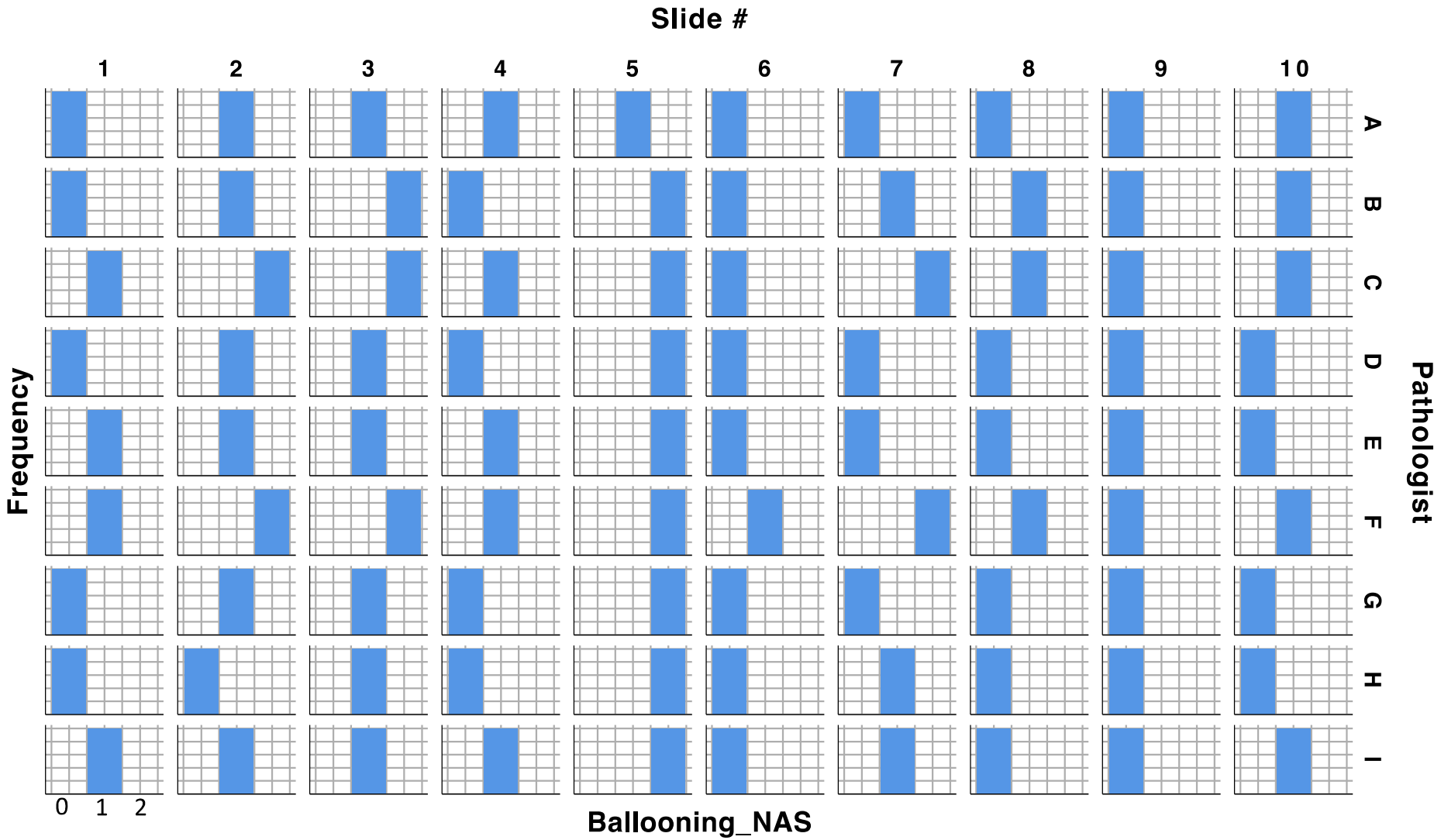


SQBS: semi-quantitative balloon score created

- 0-2 to align with NAS and SAF scores
- Transformed ballooned hepatocyte count/image/pathologist
  - 0 < 5
  - 1 = 5-75
  - 2 >75
- Cut-off b/w 1-2 derived from overall mean +1SD of ballooned cells reported per slide

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Trend of Semi-Quantitative Ballooning Score (0-2) by Slide and Pathologist

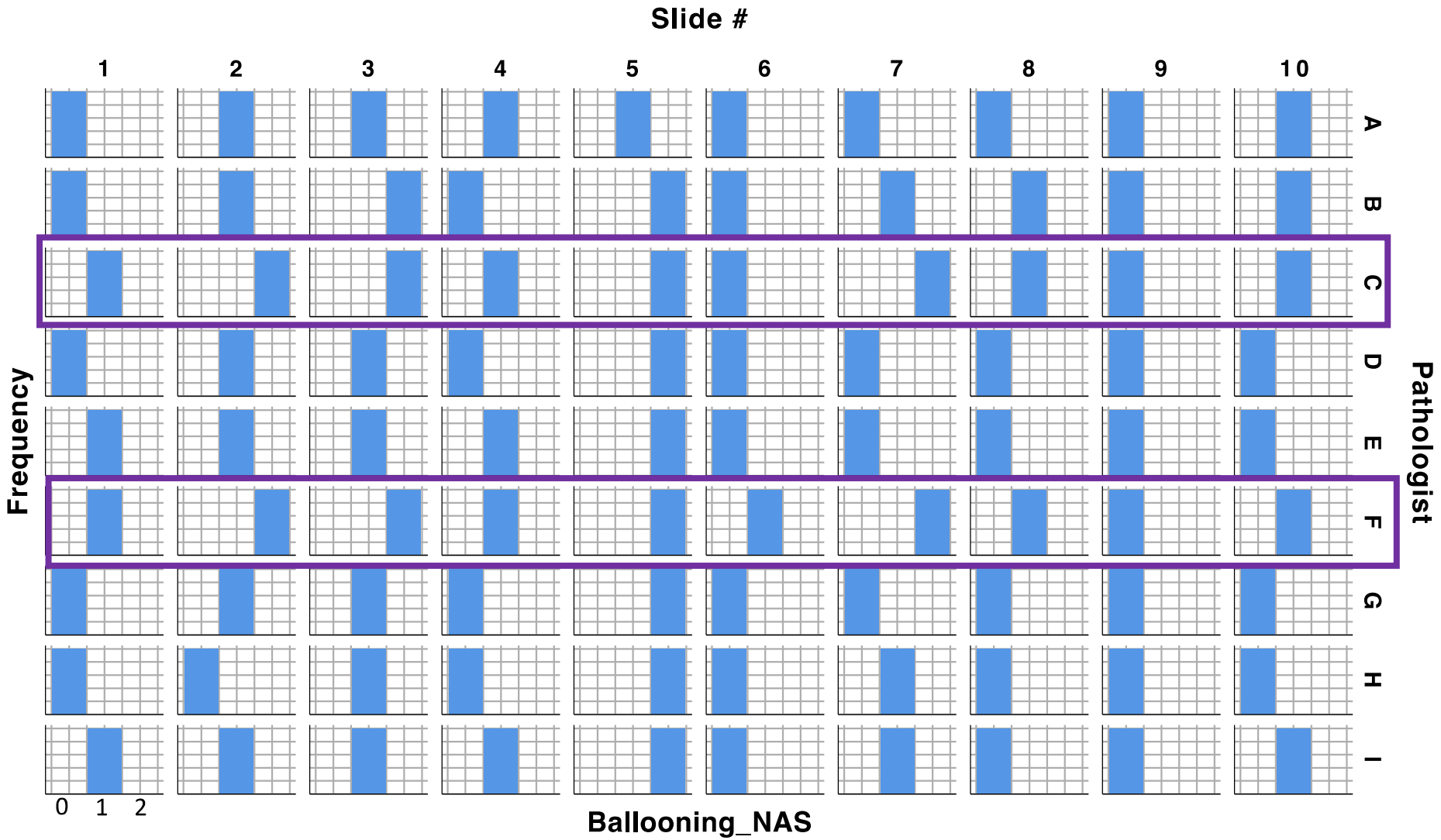


Modal Score:	0	1	1	1	2	0	0	0	0	1

Overall “fair” agreement **kappa 0.29** (95% CI 0.210-0.371); **pairwise observer kappa 0.231-1.000**.

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Trend of Semi-Quantitative Ballooning Score (0-2) by Slide and Pathologist

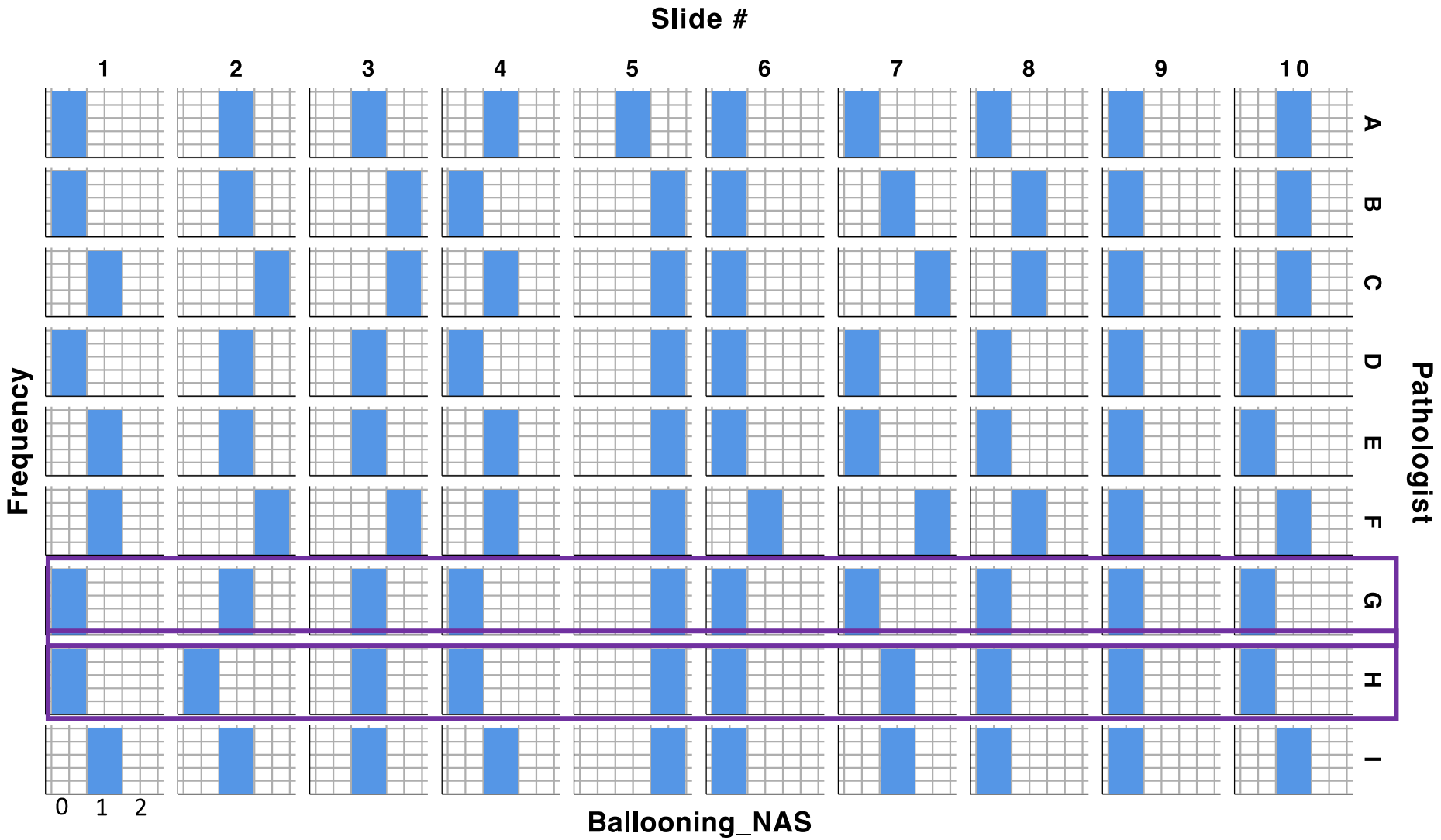


Modal Score:	0	1	1	1	2	0	0	0	0	1

Overall “fair” agreement **kappa 0.29** (95% CI 0.210-0.371); **pairwise observer kappa 0.231-1.000**.

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Trend of Semi-Quantitative Ballooning Score (0-2) by Slide and Pathologist

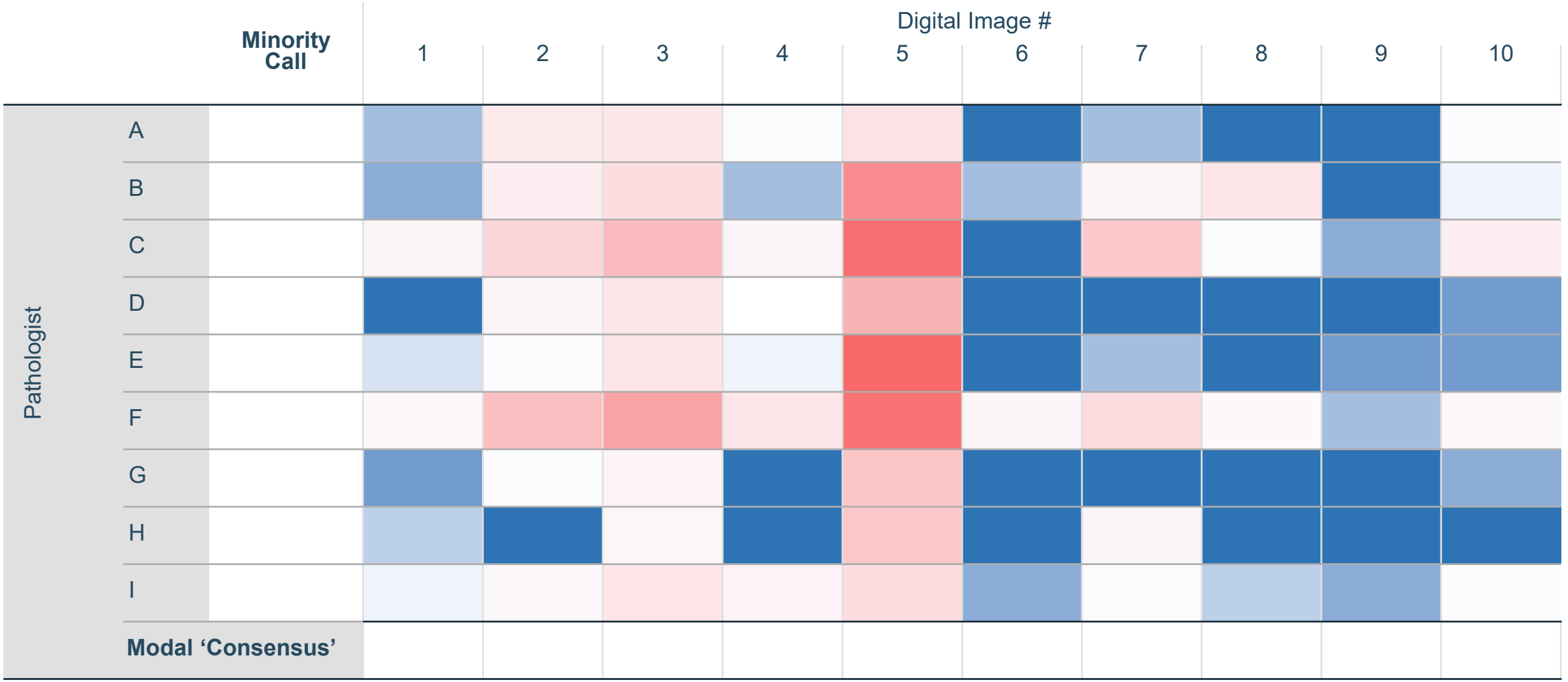


Modal Score: **0** **1** **1** **1** **2** **0** **0** **0** **0** **1**

Overall “fair” agreement **kappa 0.29** (95% CI 0.210-0.371); **pairwise observer kappa 0.231-1.000**.

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Comparison of ‘non-NASH NAFL’ vs. ‘NASH’ diagnostic call by Pathologist and Image: Correlation with ballooning.



Blue = no ballooned hepatocytes
Red = many ballooned hepatocytes

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

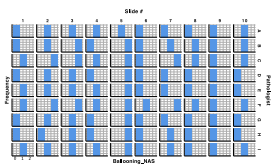
Comparison of ‘non-NASH NAFL’ vs. ‘NASH’ diagnostic call by Pathologist and Image: Correlation with ballooning.

Pathologist	Minority Call	Digital Image # 1	2	3	4	5	6	7	8	9	10
A	1/10	NASH	NASH	NASH	NASH	NASH	NASH	NASH	Not NASH	NASH	NASH
B	3/10	Not NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH
C	2/10	NASH	NASH	NASH	NASH	NASH	Not NASH	NASH	Not NASH	Not NASH	Not NASH
D	2/10	Not NASH	NASH	NASH	NASH	NASH	Not NASH	NASH	Not NASH	Not NASH	NASH
E	1/10	NASH	Not NASH	NASH	NASH	NASH	NASH	NASH	Not NASH	Not NASH	NASH
F	1/10	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	Not NASH	NASH
G	2/10	NASH	NASH	NASH	NASH	NASH	Not NASH	Not NASH	Not NASH	Not NASH	NASH
H	7/10	Not NASH	Not NASH	Not NASH	Not NASH	NASH	Not NASH	Not NASH	Not NASH	Not NASH	Not NASH
I	2/10	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH
Modal ‘Consensus’		6/9 NASH	7/9 NASH	8/9 NASH	7/9 NASH	9/9 NASH	5/9 NASH	7/9 NASH	6/9 Not NASH	6/9 Not NASH	7/9 NASH

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

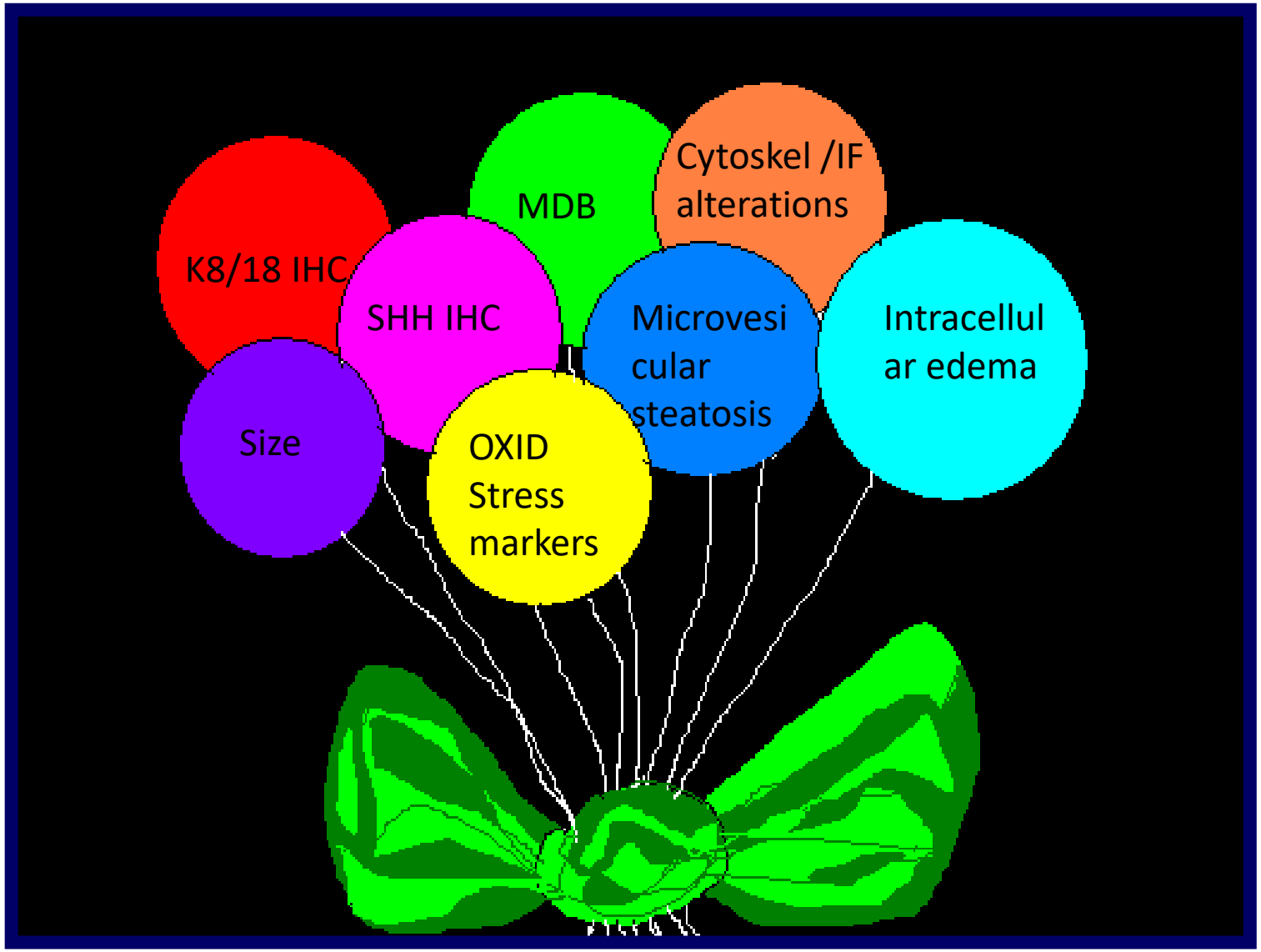
Is there a “strong” correlation of presence of ballooning with diagnosis of NASH by pathologists?

Pathologist	Minority Call	Digital Image # 1	2	3	4	5	6	7	8	9	10
A	1/10	NASH	NASH	NASH	NASH	NASH	NASH	NASH	Not NASH	NASH	NASH
B	3/10	Not NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH
C	2/10	NASH	NASH	NASH	NASH	NASH	Not NASH	NASH	Not NASH	Not NASH	Not NASH
D	2/10	Not NASH	NASH	NASH	NASH	NASH	Not NASH	NASH	Not NASH	Not NASH	NASH
E	1/10	NASH	Not NASH	NASH	NASH	NASH	NASH	NASH	Not NASH	Not NASH	NASH
F	1/10	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	Not NASH	NASH
G	2/10	NASH	NASH	NASH	NASH	NASH	Not NASH	Not NASH	Not NASH	Not NASH	NASH
H	7/10	Not NASH	Not NASH	Not NASH	Not NASH	NASH	Not NASH	Not NASH	Not NASH	Not NASH	Not NASH
I	2/10	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH	NASH
Modal ‘Consensus’		6/9 NASH	7/9 NASH	8/9 NASH	7/9 NASH	9/9 NASH	5/9 NASH	7/9 NASH	6/9 Not NASH	6/9 Not NASH	7/9 NASH
Modal SQBS Score:		**0**	**1**	**1**	**1**	**2**	**0**	**0**	**0**	**0**	**1**



J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

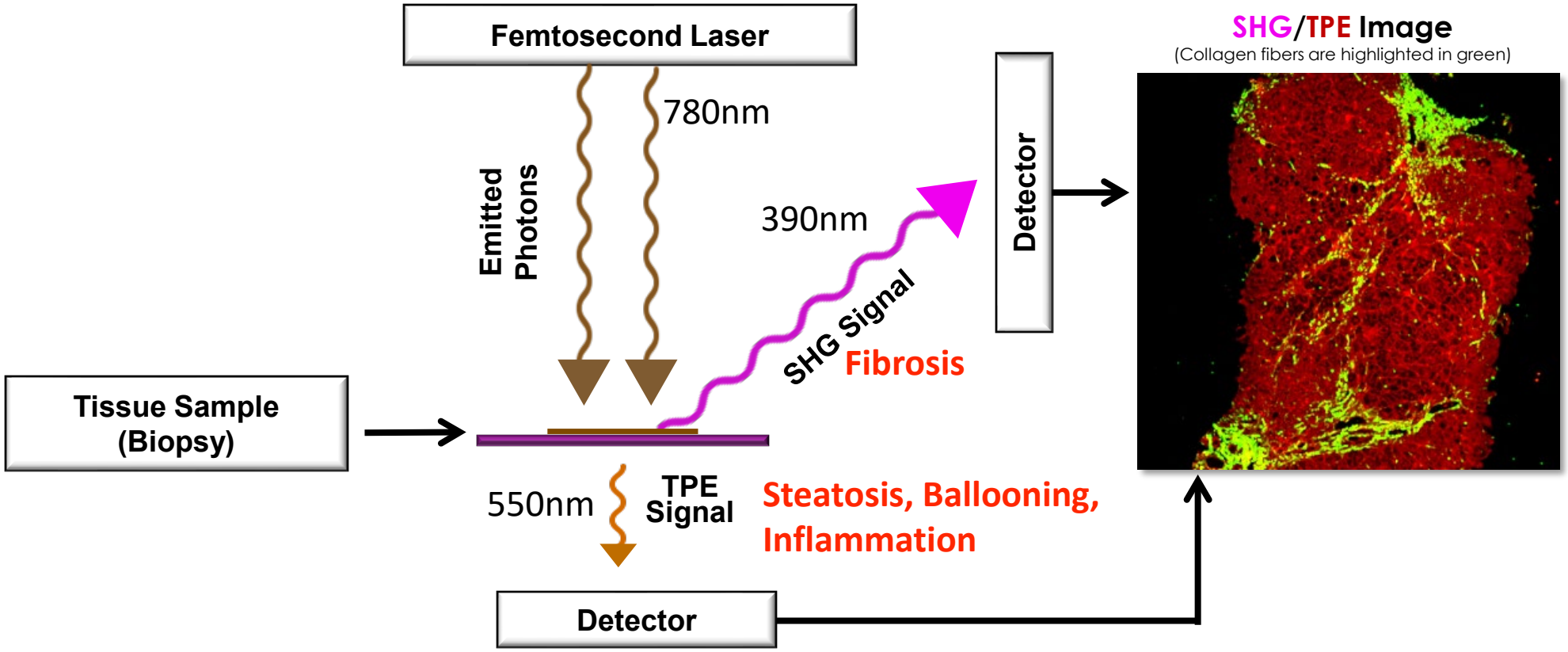
Is there a way to improve/ standardize interpretation of ballooning in NASH?



Courtesy V Desmet

SHG/TPEF: on Unstained Slides: Machine Learning…AI

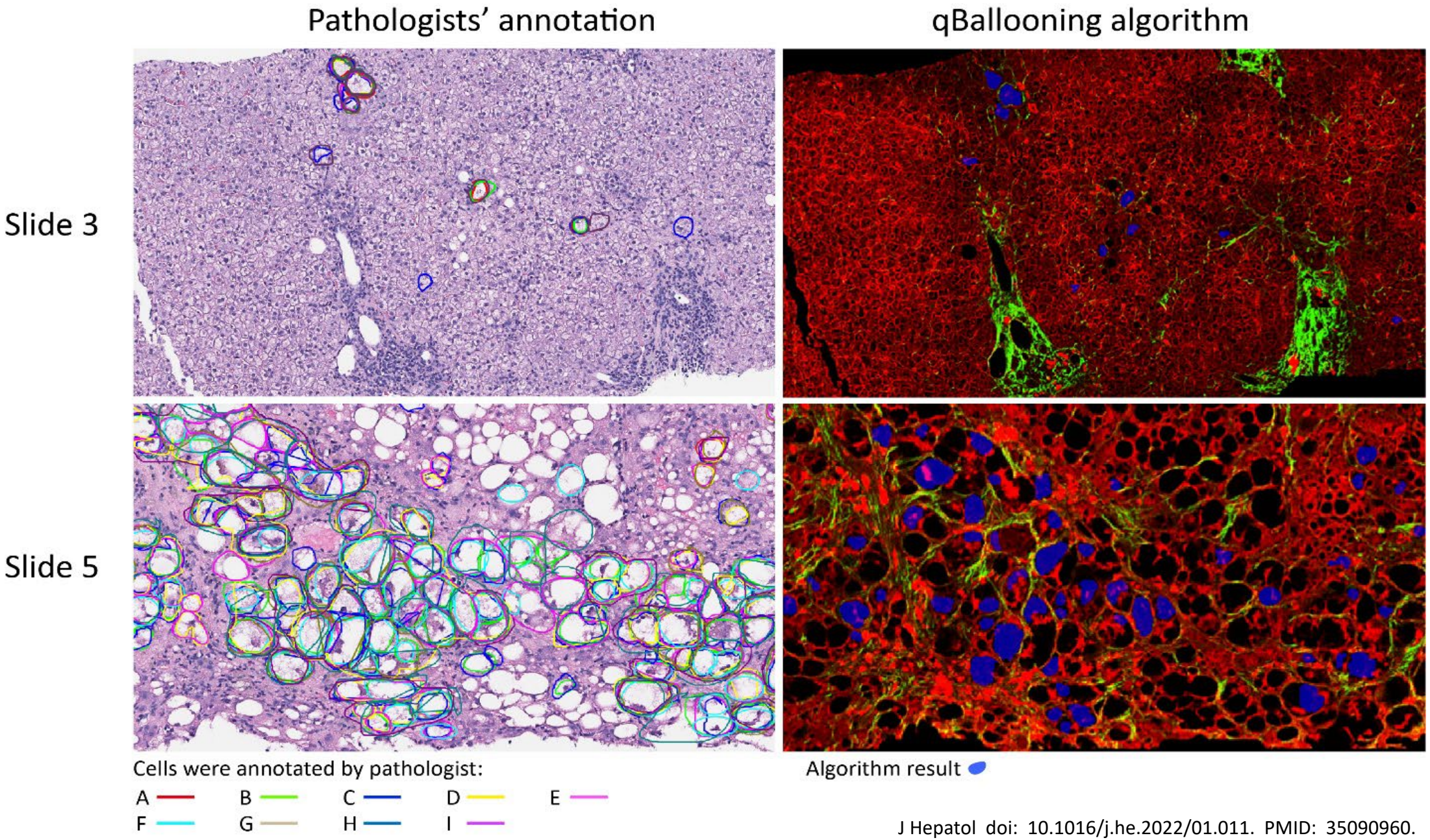
- **Reading of tissues' endogenous biomarkers and unique signature without staining**
- **Combination of both morphology and biomarkers information**



Quantitative, objective and repeatable data with reliable continuous readouts

Courtesy Dean Tai, CSO Histoindex

Algorithm: Out of 45 parameters, 7 were chosen: 1 is fibrosis related; 6 are ballooning related



J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

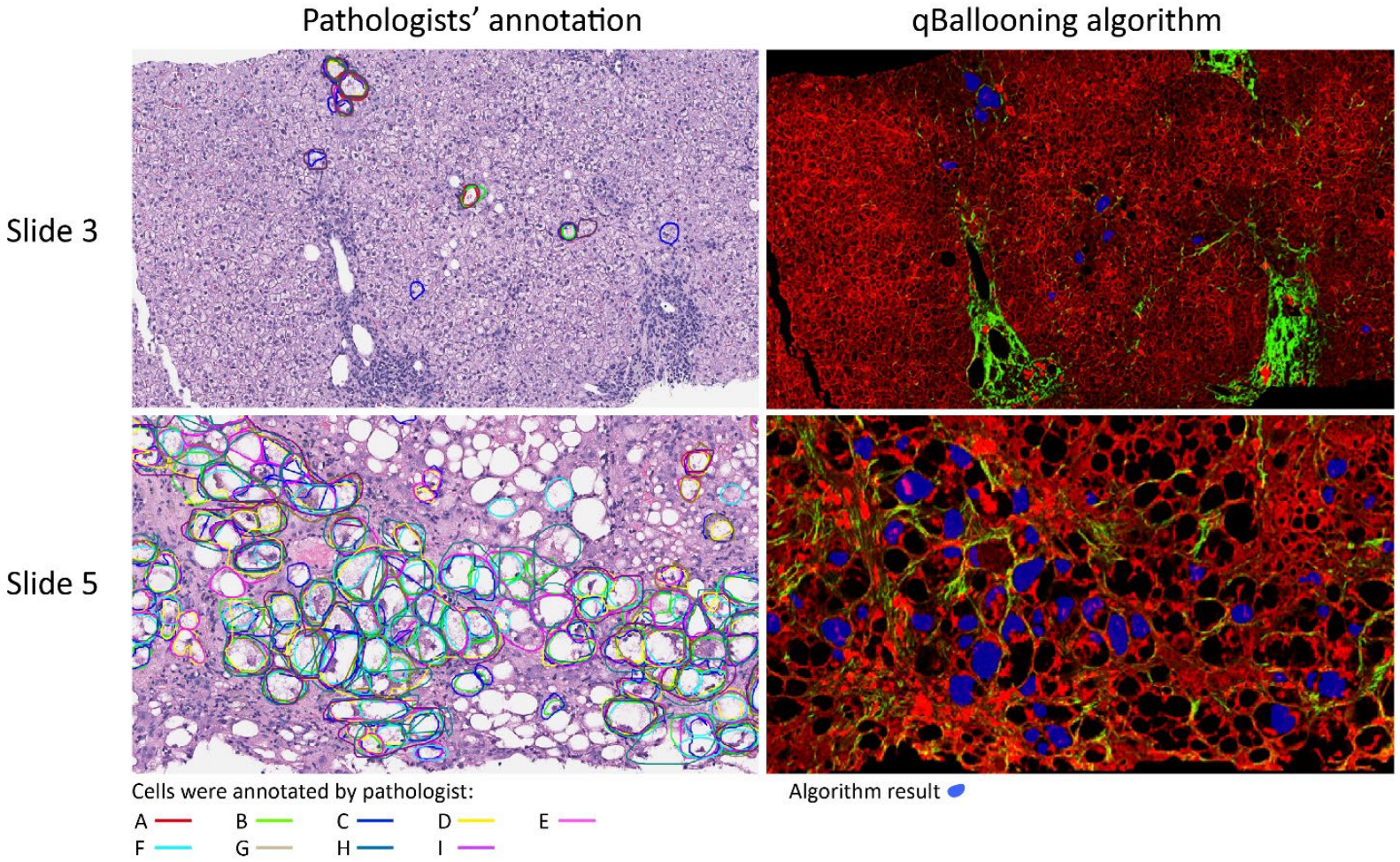
qBallooning2 training-set cell-selection criteria	Number of ballooned cells identified by Pathologists	Number of ballooned cells identified by qBallooning2	Overlap between qBallooning2 and majority concordance of ≥5-Pathologists	*Positive Predictive Value* Proportion of ballooned cells called by qBallooning2 are 'True Positive' *	*False Discovery Rate* Proportion of ballooned cells called by qBallooning2 are 'False Positive' *	*True Positive Rate (Sensitivity)* Proportion of ballooned cells identified by qBallooning2 *	*False Negative Rate* Proportion of ballooned cells missed by qBallooning2 *
Agreement of any 1 pathologist	1188	346	54	54/346 (16%)	292/346 (84%)	54/133 (41%)	79/133 (59%)
Agreement of at least 2 pathologists	481	250	51	51/250 (20%)	199/250 (79.6%)	51/133 (38%)	82/133 (62%)
Agreement of at least 3 pathologists	284	170	37	37/170 (22%)	133/170 (78.2%)	37/133 (28%)	96/133 (72%)
Agreement of at least 4 pathologists	188	114	25	25/114 (22%)	89/114 (78%)	25/133 (19%)	108/133 (81%)
Agreement of at least 5 pathologists	133	88	22	22/88 (25%)	66/88 (75%)	22/133 (17%)	111/133 (83%)
Agreement of at least 6 pathologists	86	59	16	16/59 (27%)	43/59 (73%)	16/133 (12%)	117/133 (88%)
Agreement of at least 7 pathologists	59	40	15	15/40 (38%)	25/40 (62.5%)	15/133 (11%)	118/133 (89%)
Agreement of at least 8 pathologists	26	24	5	5/24 (21%)	19/24 (79%)	5/133 (4%)	128/133 (96%)

Table comparing the performance of qBallooning2 in the development dataset. The algorithm was optimized to detect ballooned cells using data derived from each level of interobserver concordance and shows how the level of interobserver concordance stipulated affects the performance of the algorithm. * Relative to majority concordance of ≥5-pathologists.

- qballooning2 was compared with **every pathologist** to fine tune the algorithm
- had pairwise overlap with individual pathologists ranging from 19% (with Pathologist F) to 42% (with Pathologist G), which was comparable to the level of inter-observer variation between pathologists of **8-75%**
- For studies, a simple majority (>5) was chosen

‘Ground Truth’ Atlas

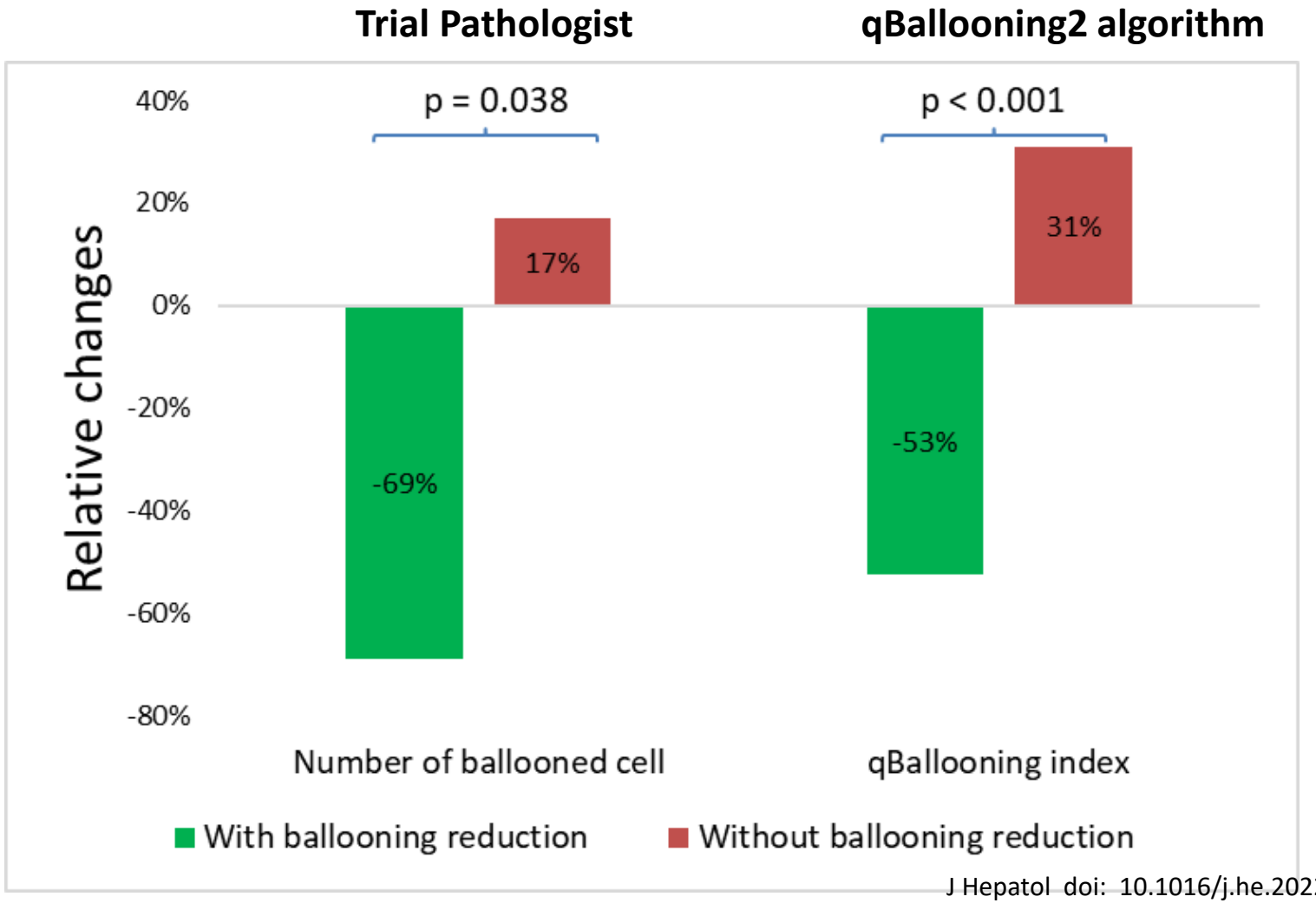
qBallooning2 training-set cell-selection criteria	Number of ballooned cells Identified by Pathologists
Agreement of any 1 pathologist	1188
Agreement of any 2 pathologists	481
Agreement of any 3 pathologists	284
Agreement of any 4 pathologists	188
Agreement of any 5 pathologists	133
Agreement of any 6 pathologists	86
Agreement of any 7 pathologists	59
Agreement of any 8+ pathologists	26



* Relative to majority consensus of ≥5-pathologists. § Based on an estimated me

Proof of Principle: Can the algorithm/machine detect change?

Quantification data showing the relative change in the number of ballooned hepatocytes and the qBallooning index for patients with and without ballooning reduction.



22 cases, paired biopsies
Resmetirom, Phase 2
[NCT02912260]

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Conclusions -1-

- Hepatocyte ballooning is recognized as THE key to distinguishing NASH from other forms of NAFLD within the spectrum

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Conclusions -1-

- Hepatocyte ballooning is recognized as THE key to distinguishing NASH from other forms of NAFLD within the spectrum

- Loss of ballooning (B0) has been identified as necessary, along with decrease in inflammation and no worsening of fibrosis, for “NASH resolution” in clinical trial assessment (FDA, EMA)

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Conclusions -1-

- Hepatocyte ballooning is recognized as THE key to distinguishing NASH from other forms of NAFLD within the spectrum

- Loss of ballooning (B0) has been identified as necessary, along with decrease in inflammation and no worsening of fibrosis, for “NASH resolution” in clinical trial assessment (FDA, EMA)

- Pathologists define ballooning similarly…but do we see it similarly in slides?

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Conclusions -2-

- Broad divergence in assessment of hepatocyte ballooning amongst expert pathologists
  - Not based on level of training or geographic location
  - Suggests there are nuances in ballooned hepatocytes differently appreciated by pathologists

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Conclusions -2-

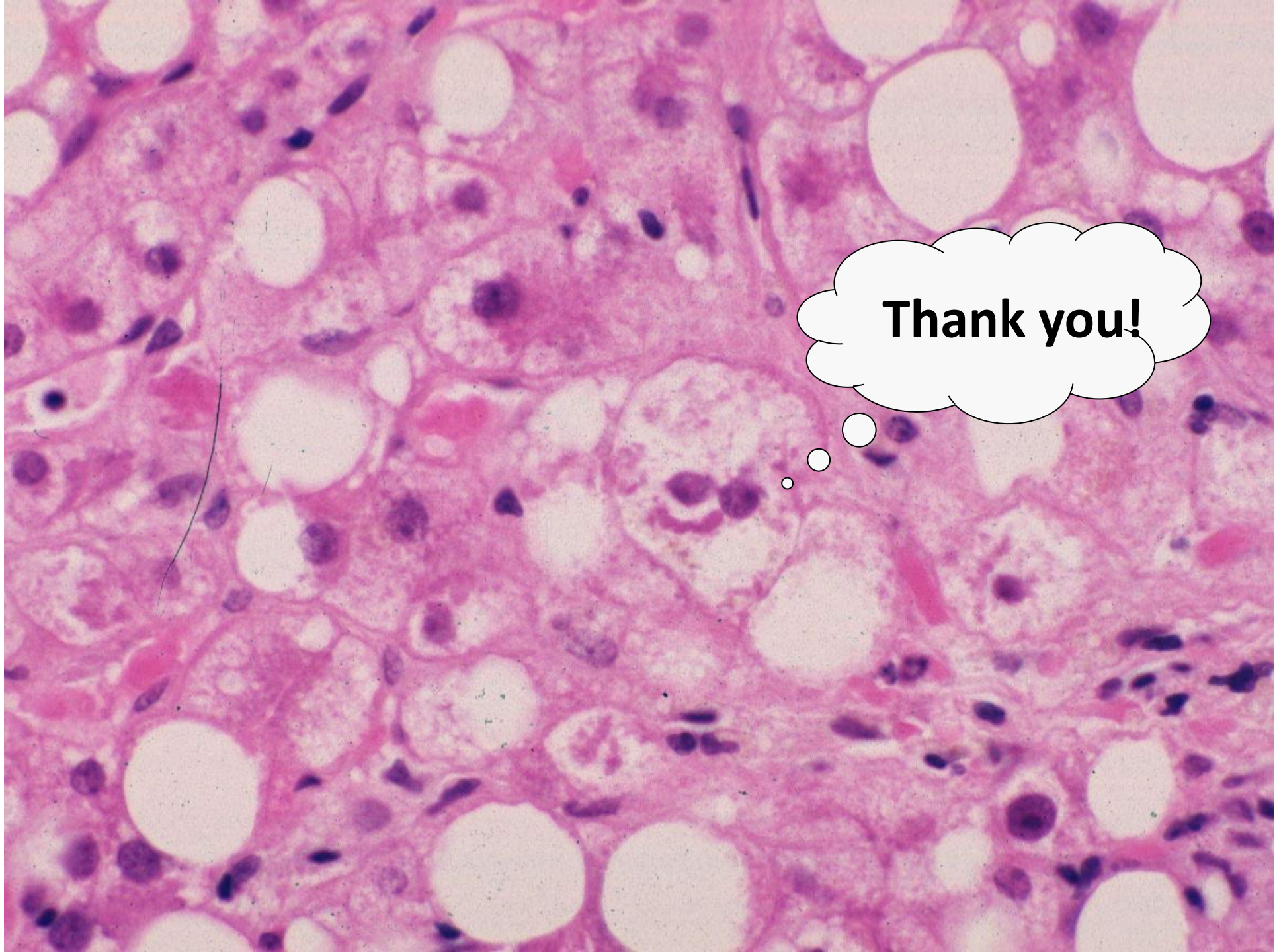
- Broad divergence in assessment of hepatocyte ballooning amongst expert pathologists
  - Not based on level of training or geographic location
  - Suggests there are nuances in ballooned hepatocytes differently appreciated by pathologists
- Implies that requiring ballooning score 0 for endpoint efficacy in trials is quite possibly unrealistic
  - It may be more realistic to look for trends in ballooning burden between pre and post intervention biopsies

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.

Conclusions -2-

- Broad divergence in assessment of hepatocyte ballooning amongst expert pathologists
  - Not based on level of training or geographic location
  - Suggests there are nuances in ballooned hepatocytes differently appreciated by pathologists
- Implies that requiring ballooning score 0 for endpoint efficacy in trials is quite possibly unrealistic
  - It may be more realistic to look for trends in ballooning burden between pre and post intervention biopsies
- Training of AI/machine learning by concordance atlas is doable and can give appropriate PPV; use of such may bring a more standardized means of assessing efficacy in clinical trials

J Hepatol doi: 10.1016/j.he.2022/01.011. PMID: 35090960.


Thank you!