ORIGINAL RESEARCH—CLINICAL

Assessment of Esophageal and Gastric Varices in Patients With Cirrhosis for Clinical Trials: A Centralized Blinded Evaluation System

Naim Alkhouri,[1,2] Raj Vuppalanchi,[3] Mazen Noureddin,[1,4,5,6] Eric Lawitz,[7] Edward Mena,[8] Nadege T. Gunn,[9] Khurram Jamil,[10] Seth Zuckerman,[10] Stephen A. Harrison,[1,†] and Naga Chalasani[11]

[1]Summit Clinical Research, San Antonio, Texas; [2]Clinical Research Institute of Ohio, Westlake, Ohio; [3]Division of Gastroenterology and Hepatology, Department of Medicine, Indiana University, Indianapolis, Indiana; [4]Department of Medicine, Houston Methodist Hospital, Houston, Texas; [5]Sherrie and Alan Conover Center for Liver Disease and Transplantation, JC Walter Jr Transplant Center, Houston Methodist Hospital, Houston, Texas; [6]Houston Research Institute, Houston, Texas; [7]Texas Liver Institute, San Antonio, Texas; [8]California Liver Research Institute, Pasadena, California; [9]Velocity Clinical Research, Waco, Texas; [10]Galectin Therapeutics, Norwood, Georgia; and [11]Division of Gastroenterology and Hepatology, Department of Medicine, Indiana University School of Medicine, Indianapolis, Indiana

BACKGROUND AND AIMS: Development of varices is an important milestone in the natural history of patients with cirrhosis, and yet the data are sparse in terms of how best to assess for gastroesophageal varices as a clinical trial outcome in multicenter studies. Here we describe a centralized upper endoscopy (esophagogastroduodenoscopy [EGD]) reading process for assessing esophageal varices (EVs) and gastric varices (GVs) and to investigate inter-reader agreement between experienced endoscopists on the presence/size of varices. **METHODS:** Patients with compensated metabolic dysfunction–associated steatohepatitis cirrhosis evaluated for inclusion in the NAVIGATE phase 2b/3 trial (NCT04365868) underwent EGD by local endoscopists, video recordings of which were centrally read by a pool of 6 qualified, trained reviewers. Two initial reviewers determined the presence/absence and size of varices, and in cases of disagreement, a third adjudicating reviewer assisted with the final determination. Agreement between the reviewers was analyzed using Cohen's kappa. **RESULTS:** Structured central blinded adjudication of varices was achieved at the participating centers across the globe. Each assigned reviewer completed their review within 24 hours of assignment. Of the 1006 EGDs reviewed, 216 (21.5%) had confirmed EVs, including 115 (53.2%) small, 71 (32.9%) medium, and 30 (13.9%) large varices. GVs were identified among 20 (2.0%) EGDs. Adjudication was required in 399 (39.7%) cases, with the third reviewer confirming varices in 216 (54.1%) cases. Percent agreement between reader pairs for EVs ranged from 40.0% to 100% (kappa 0.118–1.000), and inter-reader agreement for GVs varied between 81.8% and 100% (kappa 0.000–1.000). **CONCLUSION:** Centralized review of EGD video recordings, coupled with a structured adjudication process, can be implemented in large multicenter trials to provide reliable varices assessment in metabolic dysfunction–associated steatohepatitis clinical trials.

Keywords: Esophageal Varices; Gastric Varices; Endoscopy; Gastrointestinal; Liver Cirrhosis

Introduction

Metabolic dysfunction–associated steatotic liver disease (MASLD) is an increasingly important public health concern. MASLD currently affects 38% of adults worldwide, a number that is expected to increase to 55% by 2040.[1,2] If left unaddressed, MASLD can progress to metabolic dysfunction–associated steatohepatitis (MASH), where inflammation and hepatocyte injury can cause progressive liver fibrosis, which may eventually develop into cirrhosis.[3] In patients with cirrhosis, fibrotic tissue restricts the flow of blood through the liver, resulting in portal hypertension and increased blood flow through collateral vessels. The elevated pressure in the gastrointestinal vessels can lead to the development of varices, most prominently in the esophagus (esophageal varices [EVs]), stomach (gastric varices [GVs]), and duodenum.

Interventions are available to treat varices, including medical management through medications or procedures such as balloon tamponade or banding.[4] However, receiving these interventions depends on having an accurate diagnosis of the varices, which can be challenging and is often subjective, relying on esophagogastroduodenoscopy (EGD) for screening. Additionally, there is no standardized scoring system for varices in the context of MASH clinical trials, unlike other outcomes such as liver histology, further

[†]Deceased.

Abbreviations used in this paper: **EGD, esophagogastroduodenoscopy; EVs, esophageal varices; GEVs, gastroesophageal varices; GVs, gastric varices; IQR, interquartile range; MASH, metabolic dysfunction–associated steatohepatitis; MASLD, metabolic dysfunction–associated steatotic liver disease.**

Most current article

2772-5723
https://doi.org/10.1016/j.gastha.2026.101004

complicating systematic scoring of varices. Variability in endoscopic interpretation—arising from differences in technique, experience, and subjective assessment—can introduce misclassification that may impact endpoint ascertainment and treatment effect estimates.

This subjectivity is particularly consequential in the clinical trial setting, where variceal size classification directly determines eligibility and endpoint adjudication. Despite the drawbacks of EGD-based variceal scoring, there is a great need for a standardized approach to variceal scoring, both in the clinical and research setting. Clinical trials in patients with MASH cirrhosis are ongoing (eg, NCT04365868, NCT06528314, NCT06632457, NCT06419374), and progression to large varices is now an accepted endpoint by the US Food and Drug Administration, making the accurate diagnosis of varices and their size of utmost clinical importance. As development of pharmacotherapies for MASH cirrhosis progresses, there is a real need to systematically measure the presence and size of varices.

The purpose of the study is to describe the standardized blinded central EGD assessment and adjudication process implemented in NAVIGATE trial. We also aimed to explore the rates of inter-reader variability in terms of determining the presence and size of varices among experienced endoscopists. Subjects were evaluated for enrollment in the phase 2b/3 NAVIGATE trial examining the effects of belapectin on the prevention of EVs in patients with MASH cirrhosis and evidence of portal hypertension (NCT04365868). Belapectin is a galectin-3 inhibitor[5,6] and has been shown to reduce the hepatic venous pressure gradient and the development of varices in patients without EVs at baseline.[7] This phase 2b/3 trial utilized a centralized approach for quantifying the presence and size of varices, the effectiveness of which has not been previously studied.

Methods

Study records of all patients screened for inclusion in the NAVIGATE phase 2b/3 trial were reviewed. The target population for the study was patients with MASH-related cirrhosis with portal hypertension but without known esophageal and gastric varices. Evidence of portal hypertension was defined with either one of the following: (1) platelet count $<150{,}000/mm^3$ or (2) documented hepatic vein pressure gradient measurement >6 mmHg or (3) at least 2 of the following: spleen size ≥ 14 cm (documented by ultrasound, magnetic resonance imaging, or computed tomography scan), abdominal collateral circulation (documented by ultrasound, magnetic resonance imaging, or computed tomography scan or physical examination, ie, caput medusae), documented liver transient elastography ≥ 20 kPa, or aspartate aminotransferase/alanine transaminase ratio >1. Patients who failed to randomize into the NAVIGATE study because they had varices at baseline EGD (screen failures) were also included in the current analysis. The analysis for this manuscript was performed on deidentified data; therefore, ethics board approval was not required for the study.

As part of the enrollment process, during the screening period, an EGD was performed without therapeutic intent in patients considered eligible for the NAVIGATE study. All patients underwent elective EGD by an experienced local endoscopist. Per the schedule of assessments, randomized participants would also have an EGD at completion (or early termination) of stage 1 of the study (phase 2B week 78). If patients continued into stage 2 of the study, an additional EGD was conducted at completion (or early termination) of the study. Each local endoscopist continuously recorded the EGD on video using GI Hawkeye recording software (v1.2.1, GI Reviewers, Brookline, Massachusetts). This software records the color video feed produced by the endoscope in real time and saves the finished video file locally (USB flash drive). Users can easily black out areas of the video that display identifiable information with the software.

The esophagus was visualized by recording the insertion of the endoscope through the esophagus. Visualization of the stomach was made with forward and retroflexed views of the fundus. When possible, the gastroesophageal junction was visualized during the endoscopy of the stomach; otherwise, it was visualized as the endoscope was removed. Video of the distal esophagus was also recorded during removal of the endoscope.

Videos were saved and submitted to the central processor on a secure web-based system for central review within 24 hours of the procedure. The submitted footage underwent a quality control check prior to being anonymized via the submission system. Poor-quality and missing data were identified and addressed by the central processor to ensure the technical adequacy of the submitted EGD footage. Identifying information was also removed at this stage, prior to sharing the videos with reviewers. A key aspect of central blinded assessment was to adjudicate on size of varices if present. This was predefined criteria based on the EGD protocol as follows:

- Large >5 mm in diameter, occupying more than 1/3 of esophageal lumen
- Medium >5 mm in diameter, occupying less than 1/3 of esophageal lumen
- Small <5 mm in diameter, minimally elevated above esophageal mucosa.

Central Endoscopy Evaluation Process

A team of 6 qualified gastroenterologists participated in the central EGD review process. Qualification included at least 15 years of practice and experience in performing EGDs. All were board-certified gastroenterologists in the United States. Principal investigators of the trial were not eligible to be EGD reviewers. The reviewers underwent a training process to familiarize themselves with the centralized review process and study-specific read rules.

Reviewers were blinded to patient information, drug treatment, EGD sequence, time point designation, and local read results. Each video was reviewed by 2 initial reviewers in a parallel fashion (Figure 1). If the 2 initial reviewers were in agreement about the presence/absence and size (small, medium, or large) of varices, a final diagnosis was made. If there was disagreement between the initial reviewers, a third adjudicating reviewer was utilized to establish the final diagnosis by agreeing with one of the initial reviewers' diagnoses. If a patient underwent more than one EGD during the course of the study, the same initial reviewers were utilized for subsequent reads. All 6 reviewers participated as initial reviewers for study cases; 3 of the reviewers also served as adjudicating reviewers.

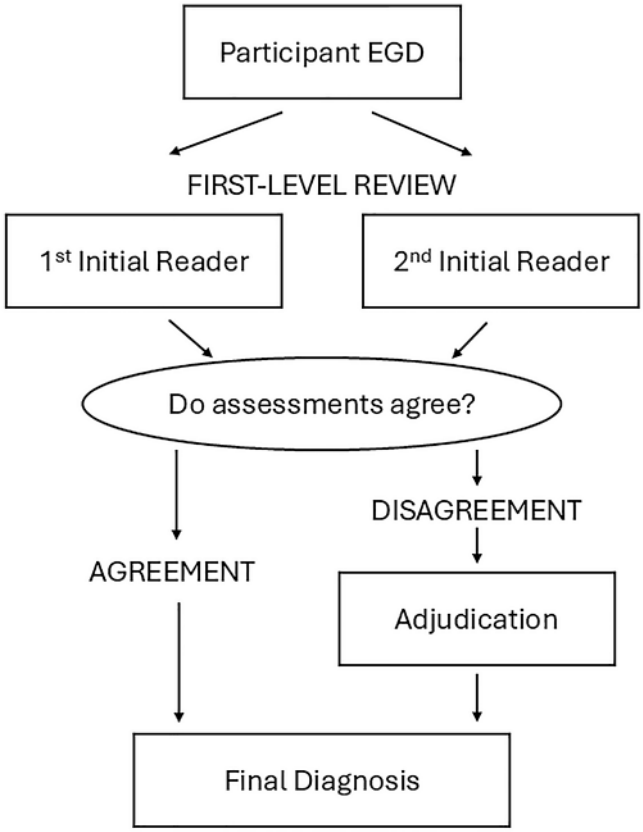

Figure 1. Visualization of esophagogastroduodenoscopy review and adjudication process.

Statistical Analysis

Inter-reader variability in ratings was measured using Cohen's kappa statistic (κ), which quantifies the level of agreement between 2 reviewers beyond that expected by chance.[8] Lower κ values indicate less agreement, while higher values indicate stronger agreement. Kappa values were interpreted using conventional categories: ≤ 0 as poor, 0.01–0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement. Kappa values, standard errors, and 95% confidence intervals were calculated using the AGREE option in PROC FREQ in SAS. The agreement expected by chance was calculated by multiplying the probability of each score between readers, then summing the products across categories. Categorical data are presented as n (%), and continuous data are presented as median, interquartile range [IQR]. SAS v8.5 (SAS Institute, Cary, NC, USA) was used for all statistical analyses.

Results

Patient Population

A total of 1006 EGDs were included in the analysis, which included screening and end-of-trial readings. All EGD video reviews occurred within 24 hours of the reviewer receiving the data. Demographics and clinical characteristics of 621 unique study participants are described in Table 1. A majority of patients were female (62.0%), white (90.1%), and non-Hispanic or Latino (64.3%). Baseline liver function tests are also reported in Table 1. The median Model for End-stage Liver Disease score at baseline was 7 (IQR, 6–8). A comparison of individuals with and without EVs is shown in Table 1.

Prevalence of Varices

Overall, EVs were noted in 216 (21.5%) cases from across time points and phase 2b and phase 3 studies, 115 (53.2%) of which were graded as small, 71 (32.9%) as medium, and 30 (13.9%) as large (Table 2). GVs were identified in 20 (2.0%) cases, and red wales were identified in 20 (2.0%) cases. Rates of varices identification when the reviewer served as an initial reader are presented in Supplementary Table 2. During the initial review, both reviewers agreed on the presence and size of varices in 216 number of EGDs. The remaining 399 (39.7%) cases required adjudication by the third reviewer, which confirmed the presence of gastroesophageal varices (GEVs) in 216 cases (54.6% of adjudicated cases; Figure 2). The proportion of initial GEV diagnoses confirmed by adjudication varied between 14.5% for reviewer 2 to 72.3% for reviewer 4 (median, 62.5%; IQR, 53.1%–63.3%). EVs were confirmed in a median of 63.9% (IQR, 58.1%–70.6%) of adjudicated cases, with confirmation rates ranging from 47.9% for reviewer 2 to 73.0% for reviewer 3. GVs were confirmed in a median of 83.3% (IQR, 54.2%–100.0%) of cases, with reviewer confirmation rates between 16.7% (reviewer 6) and 100% (reviewers 1, 3, and 6).

In patients for whom phase 2b week 78 data were available (n = 286), 44 (15.4%) patients had EVs. Of those cases, 32 (72.7%) were small, 11 (25.0%) were medium, and 1 (2.3%) was large. In the phase 3 study, 8 (13.8%) patients with available week 78 data (n = 58) developed EVs (small EV: 3 and medium EV: 5).

Inter-Reader Variability

When assessing EVs, the overall agreement between the 2 initial readers was 76.0%, (>60% expected by chance) with an unweighted kappa of 0.401 (95% CI, 0.339–0.463) (Table 3). Comparing large and medium varices to small EVs, the overall agreement rate between the initial readers was 65.3%, with a kappa of 0.299 (95% CI, 0.172–0.426). For large EVs alone, the kappa was 0.247 (95% CI, 0.082–0.411), with an agreement rate of 81.9%.

Percent agreement between reader pairs ranged from 40.0% to 100%, with values of kappa ranging from 0.118 to 1.000 (Supplementary Figure 1). For GVs, the overall agreement between readers was 92.1%, slightly higher than the 88.2% expected by chance (Table 3). Inter-reader agreement varied between 81.8% and 100%, with kappa values between 0.000 and 1.000 (Supplementary Figure 2).

Reviewer 2 had the largest number of cases requiring adjudication and the largest number of cases that ultimately were revised after adjudication (Figure 2). This reviewer

Table 1. Demographics and Clinical Characteristics of Patients Undergoing Endoscopy at the Time of Screening, Stratified by Whether Esophageal Varices Were Detected

	All subjects (N = 621)	EV detected (N = 157)	No varices (N = 164)	*P* value
Age (y), median (IQR)	61 (55–66)	63 (57–66)	61 (55–66)	.09
Sex, n (%)				.53
Female	385 (62.0%)	94 (59.9%)	291 (62.7%)	
Male	236 (38.0%)	63 (40.1%)	173 (37.3%)	
Race, n (%)				.10
American Indian or Alaska Native	27 (4.3%)	4 (2.5%)	23 (5.0%)	
Asian	13 (2.1%)	3 (1.9%)	10 (2.2%)	
Black	8 (1.3%)	0	8 (1.7%)	
Other	9 (1.4%)	0	9 (1.9%)	
White	564 (90.8%)	150 (95.5%)	414 (89.2%)	
Ethnicity, n (%)				**.04**
Hispanic or Latino	220 (35.4%)	63 (40.1%)	157 (33.8%)	
Not Hispanic or Latino	399 (64.3%)	92 (58.6%)	307 (66.2%)	
Not reported	1 (0.2%)	1 (0.6%)	0	
Unknown	1 (0.2%)	1 (0.6%)	0	
ALT (U/L), median (IQR)	33 (24–49)	34 (26–46)	33 (23–50)	.085
AST (U/L), median (IQR)	37 (27–52)	40 (31–51)	36.5 (26–52)	.936
Bilirubin (mg/dL), median (IQR)	0.6 (0.5–0.9)	0.7 (0.5–1.0)	0.6 (0.5–0.9)	.049
HbA1c (%), median (IQR)	6.2 (5.4–7.2)	6.3 (5.6–7.2)	6.0 (5.3–6.8)	**.087**
Platelets (platelets/μL), median (IQR)	127 (101–154)	111 (87–141)	132 (105–163)	**<.001**
Creatinine (μmol/L), median (IQR)	67 (58–79)	67 (56–79)	67 (59–79)	.25
Sodium (μmol/L), median (IQR)	140 (139–142)	140 (139–142)	140 (139–142)	.198
MELD score, median (IQR)	7 (6–8)	7 (7–9)	7 (6–8)	.087

Bold values denote statistical significance ($P < .05$).
ALT, alanine aminotransferase; AST, aspartate aminotransferase; HbA1c, hemoglobin A1C; MELD, Model for End-stage Liver Disease.

was more likely to identify EVs and GVs and more likely to record varices as larger than the adjudicated size. When considering EV cases in which reviewer 2 was not an initial reviewer, the kappa was 0.448 (95% CI, 0.385–0.510).

Discussion

This study describes the design and operational execution of a structured centralized endoscopic review process for the assessment of GEVs in the NAVIGATE trial—a large, multicenter, multinational clinical trial of patients with MASH cirrhosis. To our knowledge, this represents the first implementation and formal characterization of a centralized, blinded EGD adjudication system for variceal assessment in a MASH cirrhosis clinical trial, establishing a methodological framework for future cirrhosis clinical trials.

The centralized EGD review process was modeled on adjudication approaches previously validated for liver biopsy assessment in MASH trials, where discordant interpretations between primary reviewers are resolved by a third adjudicating pathologist to determine a final outcome. The present study extends this principle to endoscopic outcomes—which have not traditionally been incorporated into MASH trial endpoints despite their established clinical significance. By applying a structured, blinded, multi-reader adjudication framework to EGD-based variceal classification, this trial demonstrates that endoscopic outcomes can be operationalized in an efficient, rigorous, and reproducible manner even in a very large multinational clinical trial.

The adjudication process proved operationally efficient throughout the trial. Cases where primary readers reached discordant interpretations were systematically escalated to adjudication, and in each instance, the process produced a final classification based on majority. Notably, the adjudication framework successfully identified and corrected a pattern of systematic directional bias from a single reviewer—specifically, a tendency toward higher rates of variceal detection and assignment of larger variceal size—demonstrating that the process functioned precisely as intended: to safeguard endpoint integrity against outlier interpretations, whether arising from random variability or systematic divergence. The overall unweighted kappa for EV detection was 0.401 (95% CI, 0.339–0.463), which by conventional interpretation represents fair agreement—a finding that carries important implications for the design and analysis of MASH cirrhosis clinical trials that use variceal progression as an endpoint. Separately, while agreement for GV was higher than EV, because kappa is sensitive to disease prevalence, the relatively low prevalence of GV should be considered when interpreting pairwise kappa values for gastric variceal assessment.

From a clinical trial design perspective, the findings presented here carry meaningful implications for the evolving

Table 2. Variceal Detection Rates by Centralized Reviewers of Esophagogastroduodenoscopy Results

Reviewer ID	No. of assessments	Presence of esophageal varices (n, %)[a]	Presence of esophageal varices (accepted) (n, %)[b]	EV size: small (n, %)[b]	EV size: small (accepted) (n, %)[b]	EV size: medium (n, %)[b]	EV size: medium (accepted) (n, %)[b]	EV size: large (n, %)[b]	EV size: large (accepted) (n, %)[b]	Presence of red wales (n, %)[a]	Presence of red Wales (accepted) (n, %)[b]	Presence of gastric varices (n, %)[a]	Presence of gastric varices (accepted) (n, %)[b]
1	419	63 (15.0%)	26 (41.3%)	39 (61.9%)	17 (27.0%)	24 (38.1%)	9 (14.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	6 (1.4%)	3 (0.7%)
2	329	207 (62.9%)	41 (19.8%)	96 (46.4%)	13 (6.3%)	69 (33.3%)	12 (5.8%)	42 (20.3%)	16 (7.7%)	20 (6.1%)	9 (2.7%)	31 (9.4%)	5 (1.5%)
3	234	57 (24.4%)	34 (59.6%)	35 (61.4%)	20 (35.1%)	16 (28.1%)	9 (15.8%)	6 (10.5%)	5 (8.8%)	10 (4.3%)	5 (2.1%)	6 (2.6%)	3 (1.3%)
4	348	92 (26.4%)	46 (50.0%)	60 (65.2%)	28 (30.4%)	28 (30.4%)	16 (17.4%)	4 (4.3%)	2 (2.2%)	3 (0.9%)	2 (0.6%)	8 (2.3%)	4 (1.1%)
5	346	68 (19.7%)	34 (50.0%)	37 (54.4%)	16 (23.5%)	27 (39.7%)	16 (23.5%)	4 (5.9%)	2 (2.9%)	2 (0.6%)	1 (0.3%)	3 (0.9%)	2 (0.6%)
6	336	64 (19.0%)	35 (54.7%)	40 (62.5%)	21 (32.8%)	14 (21.9%)	9 (14.1%)	10 (15.6%)	5 (7.8%)	3 (0.9%)	3 (0.9%)	9 (2.7%)	3 (0.9%)
Total	1006		216 (21.5%)		115 (53.2%)		71 (32.9%)		30 (13.9%)		–		20 (2.0%)

[a]Percentages are based on each reviewer's total number of assessments (see "no. of assessments").
[b]Percentages are based on each reviewer's total number of detected varices.

landscape of MASH cirrhosis trials. Progression to large varices represents a clinically significant outcome—analogous to the development of ascites or hepatic encephalopathy—given its strong association with variceal hemorrhage and mortality. Despite this clinical relevance, GEV presence and size have not historically been incorporated as endpoints in MASH cirrhosis clinical trials. The NAVIGATE trial's centralized EGD review process offers proof-of-concept for the feasibility and reproducibility of GEV classification as a trial endpoint, supporting its inclusion in future randomized studies of portal hypertension in MASH cirrhosis. This work underscores the value of investing in robust central adjudication infrastructure for endoscopic outcomes and positions centralized EGD review as a scalable, clinically meaningful component of next-generation MASH trial design.

Several limitations of the centralized EGD review process in the NAVIGATE trial merit acknowledgment. Participating endoscopists were selected based on clinical expertise alone, without formal assessment of baseline inter-reader agreement or structured calibration exercises—practices that have shown promise in MASH liver biopsy trials and should be incorporated in future implementations. The adjudication framework relied on unanimity between 2 primary reviewers or majority decision when escalated for adjudication, which, while operationally practical, differs from the simultaneous consensus-based model employed by the nonalcoholic steatohepatitis (NASH) Clinical Research Network—a more thorough approach, though resource-intensive and potentially subject to reviewer dominance during consensus meeting. Finally, no prospective quality control mechanism was in place to detect systematic reader deviation in real time or periodically; future trials should incorporate periodic inter-reader performance monitoring to identify and address such divergence before it accumulates across the entirety of the trial.

This experience generates important operational insights for refining centralized EGD review in subsequent trials. Rather than rotating readers through primary and adjudicating roles, future implementations may benefit from preassigning dedicated "primary" and "adjudicating" reviewers for each case to preserve the integrity of the independent review hierarchy. Similar to liver biopsy, emerging artificial intelligence tools, particularly deep learning-based image recognition models, offer a promising avenue to supplement the centralized EGD review process in future trials. Additionally, prospective calibration exercises, pre-specified agreement thresholds, and periodic performance monitoring could be incorporated to minimize systematic divergence before it accumulates across a large case volume. These refinements would further strengthen the reliability and efficiency of the centralized review infrastructure.

Conclusion

Centralized EGD review with structured adjudication is an operationally feasible approach to variceal classification in large, multinational MASH cirrhosis trials. Large MASH outcome trials have not consistently incorporated the

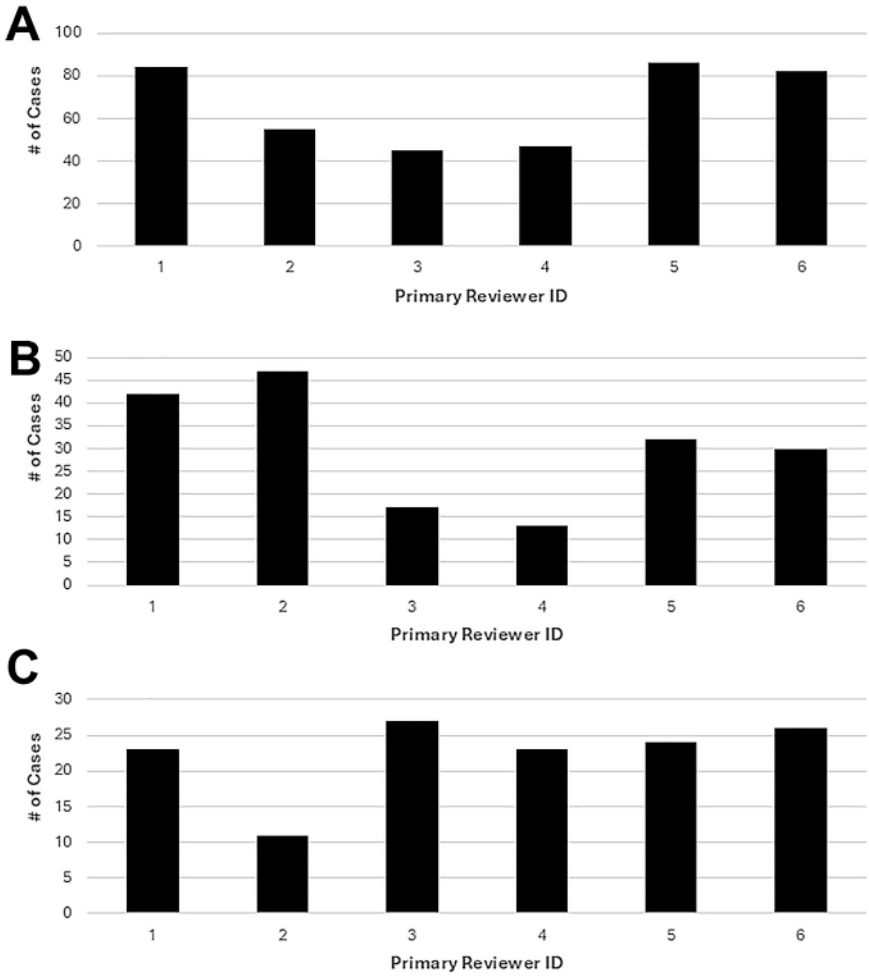

Figure 2. Number of adjudications per reviewer and outcomes of the adjudication process. (A) The number of adjudications required for each initial reviewer for all endoscopies. (B) The number of initial adjudications that were refuted by the adjudicating reviewer, grouped by initial reviewer. (C) The number of cases with esophageal varices confirmed by the adjudicating reviewer, grouped by initial reviewer.

development/progression to large EV as a hard clinical endpoint, representing a meaningful missed opportunity. If histologically confirmed cirrhosis is accepted as a hard outcome in F2/F3 populations, the development of large EV should be considered an equally definitive endpoint in cirrhosis trials, given its direct association with variceal

Table 3. Pairwise Inter-Reviewer Agreement for the Presence of Varices

Type of varices	Reviewer pair	Endoscopies assessed (N)[a]	% agreement	% expected by chance	Simple kappa	Standard error	Kappa 95% CI
Esophageal	Overall	1006	765 (76.0%)	604 (60.0%)	0.401	0.032	0.339–0.463
	R1 and R4	143	117 (81.8%)	98 (68.7%)	0.418	0.091	0.240–0.596
	R1 and R5	134	116 (86.6%)	97 (72.5%)	0.511	0.100	0.315–0.707
	R1 and R6	137	120 (87.6%)	94 (68.7%)	0.604	0.086	0.435–0.773
	R2 and R4	101	59 (58.4%)	51 (50.0%)	0.168	0.098	−0.024 to 0.359
	R2 and R5	115	63 (54.8%)	56 (48.9%)	0.115	0.077	−0.036 to 0.267
	R2 and R6	108	60 (55.6%)	54 (49.6%)	0.118	0.086	−0.050 to 0.286
Gastric	Overall	1006	963 (95.7%)	945 (93.9%)	0.295	0.078	0.143–0.447
	R1 and R4	143	142 (99.3%)	142 (99.3%)	0.000	0.000	0.000–0.000
	R1 and R5	134	133 (99.3%)	133 (99.3%)	0.000	0.000	0.000–0.000
	R1 and R6	137	130 (94.9%)	128 (93.6%)	0.199	0.187	−0.167 to 0.565
	R2 and R4	101	90 (89.1%)	85 (84.5%)	0.299	0.154	−0.003 to 0.601
	R2 and R5	115	103 (89.6%)	103 (89.9%)	−0.030	0.018	−0.065 to 0.005
	R2 and R6	108	101 (93.5%)	99 (92.0%)	0.192	0.187	−0.175 to 0.559

[a]Limited to reviewer pairs coreviewing 100 cases or more.

hemorrhage and mortality. The NAVIGATE trial addresses this gap by demonstrating that centralized adjudication of GEVs is not only feasible and operationally efficient but also yields consistent and clinically meaningful results—establishing a replicable framework for incorporating endoscopic endpoints into future large-scale MASH cirrhosis trials. These findings support the inclusion of EGD-based variceal endpoints in future clinical trials and underscore the importance of systematic adjudication workflows to ensure consistent, reliable endpoint assessment in this setting.

Supplementary Materials

Material associated with this article can be found, in the online version, at https://doi:10.1016/j.gastha.2026.101004.

References

1. Younossi ZM, Kalligeros M, Henry L. Epidemiology of metabolic dysfunction-associated steatotic liver disease. Clin Mol Hepatol 2025;31(Suppl):S32–S50.
2. Harrison SA, Gawrieh S, Roberts K, et al. Prospective evaluation of the prevalence of non-alcoholic fatty liver disease and steatohepatitis in a large middle-aged US cohort. J Hepatol 2021;75(2):284–291.
3. Diehl AM, Day C. Cause, pathogenesis, and treatment of nonalcoholic steatohepatitis. N Engl J Med 2017; 377(21):2063–2072.
4. Kaplan DE, Ripoll C, Thiele M, et al. AASLD Practice Guidance on risk stratification and management of portal hypertension and varices in cirrhosis. Hepatology 2024;79(5):1180–1211.
5. Al Attar A, Antaramian A, Noureddin M. Review of galectin-3 inhibitors in the treatment of nonalcoholic steatohepatitis. Expert Rev Clin Pharmacol 2021; 14(4):457–464.
6. Harrison SA, Marri SR, Chalasani N, et al. Randomised clinical study: GR-MD-02, a galectin-3 inhibitor, vs. placebo in patients having non-alcoholic steatohepatitis with advanced fibrosis. Aliment Pharmacol Ther 2016; 44(11–12):1183–1198.
7. Chalasani N, Abdelmalek MF, Garcia-Tsao G, et al. Effects of belapectin, an inhibitor of galectin-3, in patients with nonalcoholic steatohepatitis with cirrhosis and portal hypertension. Gastroenterology 2020; 158(5):1334–1345.e5.
8. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20(1):37–46.

Received April 20, 2026. Accepted May 5, 2026.

Correspondence:
Address correspondence to: Naga Chalasani, MD, Indiana University School of Medicine, 702 Rotary Circle, Suite 225, Indianapolis, Indiana 46202. e-mail: nchalasa@iu.edu.

Acknowledgments:
Assistance with the study: Gražina Binkauskiene, MS (Summit Clinical Research) assisted with the statistical analysis for this manuscript and Elizabeth W. Brombosz, PhD (Summit Clinical Research) assisted with drafting and preparing the manuscript. Both were funded by Galectin Therapeutics.

Authors' Contributions:
Naim Alkhouri and Khurram Jamil: Conception and design of the work. Naim Alkhouri, Stephen A. Harrison, and Khurram Jamil: Acquisition and analysis of the data. Naim Alkhouri: Drafting the manuscript. Naim Alkhouri, Raj Vuppalanchi, Mazen Noureddin, Eric Lawitz, Edward Mena, Nadege T. Gunn, Khurram Jamil, Seth Zuckerman, and Naga Chalasani: Critical review of the manuscript. Naim Alkhouri, Raj Vuppalanchi, Mazen Noureddin, Eric Lawitz, Edward Mena, Nadege T. Gunn, Khurram Jamil, Seth Zuckerman, and Naga Chalasani: Final approval of manuscript.

A preliminary version of this analysis was presented as a poster at the 2024 AASLD The Liver Meeting in San Diego, CA.

Conflicts of Interest:
These authors disclose the following: Naim Alkhouri received research support from 89bio, Akero Therapeutics, Arbutus Biopharma, AstraZeneca, BioAge, Boehringer Ingelheim, Bristol Myers Squibb, Corcept Therapeutics, Galectin Therapeutics, Genentech, Gilead Sciences, Hepagene Therapeutics, Intercept Pharmaceuticals, Inventiva Pharma, Ionis Pharmaceuticals, Ipsen, Lilly, Madrigal Pharmaceuticals, Merck, NGM Biopharmaceuticals, NorthSea Therapeutics, Novo Nordisk, Perspectum, Pfizer, PharmaIN, Poxel, Regeneron, Viking Therapeutics, and Zydus; consulting fees from 89bio, AbbVie, Akero, Boehringer Ingelheim, Echosens, Fibronostics, Gilead Sciences, HistoIndex, Intercept Pharmaceuticals, Ipsen, LiverRight, Madrigal Pharmaceuticals, NorthSea Therapeutics, Novo Nordisk, Perspectum, Pfizer, Regeneron, and Sonic Incytes; and speaking fees from AbbVie, AstraZeneca, Echosens, Gilead Sciences, Ipsen, Madrigal Pharmaceuticals, Novo Nordisk, and Perspectum. Raj Vuppalanchi reports research grants to his institution from Eli Lilly, GlaxoSmithKline (GSK), Takeda, and Zydus Therapeutics; prior institutional research support from Galectin Therapeutics, Kowa, and Gilead, which are no longer active; reports participation on advisory boards, adjudication committees or data safety monitoring boards for Fortrea, Medpace, Thermo Fisher, Madrigal Pharmaceuticals, Regeneron, 89Bio, GSK, Agios, Takeda, and Cogent Biosciences, with payments made to him; and prior advisory roles with Intercept Pharmaceuticals and Akero Therapeutics, which are no longer active; and reports stock ownership in Viking Therapeutics. Mazen Noureddin received research support from Allergan, Altimmune, Akero, BI, BMS, Boston Pharma, Conatus, Corcept, Gilead, Galectin, Genfit, GSK, Kowa, Enanta, Madrigal, Lilly, Merck, Novartis, Novo Nordisk, Rivus, Shire, Takeda, Terns, Viking, and Zydus; consulting fees from Akero, Altimmune, Alligos, AstraZeneca, BI, Boston Pharma, Curve Bioscience, Cytodyn, GSK, Histoindex, Kryia, Lilly, Madrigal, Merck, Novo Nordisk, OPKO, Rivus, Sagimet, Terns, and Takeda; speaking fees from Madrigal and Novo Nordisk; has leadership roles at the journal *Clinical Gastroenterology and Hepatology* (associate editor) and the American Association for the Study of Liver Diseases (chair of the metabolic dysfunction–associated steatotic liver disease special interest group); and stock in Rivus Pharma, OPKO, Kryia, and Akero. Eric Lawitz received research support from 89Bio, Akero, Alnylam, AstraZeneca, Bioscience, Boehringer Ingelheim, Boston, Bristol Myers Squibb, Cour, Eli Lilly, Enanta, Enyo, Exalenz, Galectin, Galmed, Gasherbrum, Genfit, Gilead, GlaxoSmithKline, Hanmi, Hightide, Intercept, Inventiva, Ipsen, Madrigal, Merck, Neurobo, NGM, Novartis, Novo Nordisk, Regeneron, Rivus, Sagimet, Takeda, Terns, Viking, and Zydus; speaking fees from Gilead, Intercept, and Novo Nordisk; consulting fees from 89Bio, AstraZeneca, Boehringer Ingelheim, Corcept, Eli Lilly, Inventiva, Merck, Novo Nordisk, Organovo, Regeneron, and Sagimet. Edward Mena received speaking fees from Gilead, Novo Nordisk, AbbVie, Eli Lilly, and Madrigal Pharmaceuticals. Khurram Jamil is employed by Galectin Therapeutics and owns stock options in Galectin Therapeutics. Seth Zuckerman is employed by Galectin Therapeutics. Naga Chalasani received research funding from Madrigal; consulting fees from Madrigal, Zydus, GSK, Chugai, Biomea Fusion, Insitro, Altimmune, Janssen, and Ipsen; and ownership in Heligenics, a drug discovery start-up company and Avant Santé, a contract research organization. The remaining authors disclose no conflicts.

Funding:
This study was funded by Galectin Therapeutics.

Ethical Statement:
The analysis for this manuscript was performed on deidentified data; therefore, ethics board approval was not required for the study.

Data Transparency Statement:
Data may be available from the authors on reasonable request pending approval from Galectin Therapeutics. Participant-level data containing identifiable data will not be shared as it is covered by privacy laws.

Reporting Guidelines:
This manuscript follows the Uniform Requirements for Manuscripts Submitted to Biomedical Journals set forth by the ICMJE.
This study does not report the direct results of a clinical trial and therefore the CONSORT checklist is not applicable.