coreNASH: Multi-stakeholder Consensus on Core Outcomes for Decision Making About Nonalcoholic Steatohepatitis Treatment

Elizabeth Clearfield ,[1] Veronica Miller,[2] Joseph Nadglowski,[3] Katherine Barradas ,[2] Jennifer Al Naber ,[1] Arun J. Sanyal,[4] Brent A. Neuschwander-Tetri ,[5] and Donna A. Messner[1], on behalf of the coreNASH Panel

The increasing prevalence and burden of nonalcoholic steatohepatitis (NASH) has spurred the development of new treatments and a need to consider outcomes used for NASH treatment decision making. Development of a NASH core outcome set (COS) can help prioritize outcomes of highest importance by incorporating the perspectives from a variety of decision makers. coreNASH was an initiative to develop a COS for NASH using a modified Delphi consensus process with a multi-stakeholder voting panel. A candidate outcome list was created based on a literature review and key informant interviews. The candidate outcome list was then condensed and prioritized through three rounds of online voting and through discussion at an in-person meeting. Outcomes were retained or eliminated based on predetermined consensus criteria, which included special weighting of patients' opinions in the first two voting rounds. The coreNASH Delphi panel included 53 participants (7 patients, 10 clinicians and researchers, 7 health technology assessors, 22 industry representatives, 2 regulators, and 5 payers) who considered outcomes for two NASH-related COS: one for NASH without cirrhosis (F2-F3) and one for NASH with cirrhosis (F4). The initial candidate outcome list for both disease stages included 86 outcomes. The panel agreed on including two core outcomes for NASH without cirrhosis and nine core outcomes for NASH with cirrhosis in the COS. *Conclusion:* A consensus-based COS has been developed that can be used across the life cycle of NASH treatments. Outcomes included can contribute to decision making for regulatory, market access, and on-market decision making. Including the coreNASH COS in clinical development programs will facilitate improved comparisons and help decision makers assess the value of new products. (*Hepatology Communications* 2021;5:774-785).

Nonalcoholic steatohepatitis (NASH) is a progressive type of nonalcoholic fatty liver disease (NAFLD) characterized by inflammation and liver injury associated with the accumulation of fat in the liver. In the United States, the estimated population of patients with NASH in 2015 was 16.52 million; by 2030, the prevalence of people with NASH is expected to increase to about 27 million.[1] The rise in the prevalence of NASH is correlated with the obesity epidemic,[2] and NASH is

Abbreviations: COMET, Core Outcomes Measures in Effectiveness Trials; COS, core outcome set; HTA, health technology assessment; MELD, Model for End-Stage Liver Disease; NCT, national clinical trial; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis.

Received September 29, 2020; accepted December 26, 2020.

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep4.1678/suppinfo.

Supported by unrestricted participation fees from Allergan, Astra-Zeneca, Boehringer Ingelheim, GENFIT, Intercept Pharmaceuticals, NGM Biopharmaceuticals, Novartis, Novo Nordisk Inc., NuSirt Biopharma, Pfizer, and Sanofi (to the Green Park Collaborative).

View this article online at wileyonlinelibrary.com.

DOI 10.1002/hep4.1678

expected to soon become the leading indication for liver transplant.[3] Complications of NASH include chronic inflammation and fibrosis that can progress to cirrhosis, liver failure, cancer, and death.[4,5]

There are currently no approved pharmacologic therapies for NASH; the increasing burden of NASH and corresponding lack of effective therapies has spurred the development of many new treatments.[6,7] New drug candidates in the pipeline aim to block or reverse the metabolic perturbations, cell injury, and inflammatory and fibrogenic pathways that are active during disease progression.[8-14] With a new generation of possible breakthrough therapies for NASH in development and the very large population of patients who may be eligible for treatment, significant attention is being directed to regulatory oversight of these products. Case definitions,[15] baseline parameters,[16] and endpoints[17] for clinical trials have been recommended and concentrate on the regulatory perspective. Draft guidance has been issued by regulatory agencies providing direction on outcomes of interest for regulatory decisions.[18-20] Phase 3 trials are ongoing or providing readout. Because of the potential for new treatments to affect disease management, we are at a key juncture to document opinions about outcome priorities from a diverse group of stakeholders who soon may use or make decisions about using these new treatments. It will be important to understand how these treatments will be evaluated by payers, health technology assessment (HTA) organizations, and other "postregulatory" decision makers.

The objective of the coreNASH initiative was to develop a core outcome set (COS) for use in trials of new treatments for NASH. A COS is a minimum set of outcomes that should be measured and reported in all clinical trials of a specific condition.[21] coreNASH used a multi-stakeholder Delphi consensus process, bringing together patients, clinicians, regulators, payers, health technology assessors, researchers, and drug developers to determine a core set of outcomes to be recommended. The outcomes in the COS measure and differentiate the effectiveness and

Potential conflict of interest: Dr. Nadglowski is employed by the Obesity Action Coalition. Dr. Neuschwander-Tetri consults for, advises, and received grants from Allergan, Bristol-Myers Squibb, Enanta Pharmaceuticals, Inc., GENFIT, Gilead Sciences, Intercept Pharmaceuticals, Madrigal Pharmaceuticals, and NGM Biopharmaceuticals; he consults for and advises Arrowhead Pharmaceuticals, Axcella Health, Boehringer Ingelheim, Durect, Ferring Pharmaceuticals, Fortress Biotech, Inc., Gelesis, HighTide Therapeutics, HistoIndex, Lipocine, MedImmune, Merck, Mundipharma, pH-Pharma, and Siemens; he received grants from Cirius Therapeutics. Dr. Miller's institution received grants from Abbott Molecular, AbbVie (Allergan), Acer Therapeutics, Aligos Therapeutics, Altimmune, AMRA, Amplyx Pharmaceuticals, Antios Therapeutics, Arbutus Biopharma, Arrowhead Pharmaceuticals, Assembly Biosciences, BioMarin Pharmaceutical, Blade Therapeutics, Boeheringer Ingelheim, Bristol-Myers Squib, ChemomAb Ltd., Coherus BioSciences, Covance, Cirius Therapeutics Inc., DDL Diagnostics, EA Pharma Co., Ltd., Echosens, Enanta Pharmaceuticals Inc., ENYO Pharma, E-SCOPICS, Galectin Therapeutics, Genkyotex, Gilead Sciences, GlaxoSmithKline, Glympse Bio, Global Blood Therapeutics, HepQuant, HepaRegenix, HighTide Therapeutics, Horizon Pharma, Immunocore, Indalo Therapeutics, Intercept Pharmaceuticals, Inventiva, Janssen, LabCorp, Eli Lilly and Co, Merck Laboratories, Metacrine, Mirum Pharmaceuticals, Morphic Therapeutic, Nitto Denko Technical Corp, Nordic Bioscience, Novartis, Novo Nordisk Inc., PathAI, Pfizer, PharmaNest, REGENXBIO, Resoundant, Retrophin, Robarts Clinical Trials, Roche Molecular Systems, Sagimet Biosciences, Shire, Springbank Pharma, Takeda, Vir Biotechnology, ViiV Healthcare, Viracor Eurofins, Virion Therapeutics LLC, and WuxiNextCode. Dr. Sanyal consults for and received grants from Conatus Pharmaceuticals, Gilead Sciences, Echosens-Sandhill, Mallinckrodt Pharmaceuticals, Salix Pharmaceuticals, Novartis, Galectin Therapeutics, and Sequana Medical; he consults for and owns stock in GENFIT, HemoShear, Durect, and Indalo Therapeutics; he consults for Immuron, Intercept Pharmaceuticals, Pfizer, Boehringer Ingelheim, Nimbus Therapeutics, Eli Lilly and Co, Merck, Novo Nordisk Inc., Fractyl Laboratories, Inc., Allergan, Chemomab, Afimmune, Teva Pharmaceuticals, Ardelyx, Terns Pharmaceuticals, ENYO Pharma, Bird Rock Bio, Albireo Pharma, Inc., Sanofi, Takeda, Janssen, Zydus Pharmaceuticals, BASF, AMRA, Perspectum Diagnostics, Owl, Poxel, Servier, Second Genome, General Electric, and 89 Bio; he received grants from Bristol-Myers Squibb and royalties from Elsevier and UpToDate Inc.; he owns stock in Exalenz Bioscience, Akarna Therapeutics, and Tiziana Life Sciences and is employed by Sanyal Bio. The other authors have nothing to report.

ARTICLE INFORMATION:

From the [1]Center for Medical Technology Policy, Baltimore, MD, USA; [2]University of California, Berkeley, Berkeley, CA, USA; [3]Obesity Action Coalition, Tampa, FL, USA; [4]Virginia Commonwealth University, Richmond, VA, USA; [5]Saint Louis University, St. Louis, MO, USA.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO:

Elizabeth Clearfield, M.H.S.
Center for Medical Technology Policy
401 East Pratt Street, Suite 631
Baltimore, MD 21202, USA
E-mail: elizabeth.clearfield@cmtpnet.org
Tel.: +1-410-547-2687 ext.121

value of therapy from the perspective of a variety of stakeholders involved; the perspective of patients, health care providers, and payers/HTA agencies was highlighted.

Participants and Methods

The Core Outcomes Measures in Effectiveness Trials (COMET) Initiative's Core Outcome Set-Standards for Development (COS-STAD) recommendations were followed for the development and implementation of this study.[22] This project was registered in COMET's database of COS projects (COMET registration number 1150).

PARTICIPANTS

A multi-stakeholder group from North America and Europe was invited to participate, and participants were categorized into patients/patient advocates, clinicians/researchers, U.S. payers, health technology assessors, industry representatives, or regulators. Participating patients and patient advocates represented a variety of levels of NAFLD disease severity. Diagnoses ranged from nonalcoholic fatty liver through NASH and NASH cirrhosis, a patient who had previously received a liver transplant due to another liver condition, and patient advocates representing patient organizations serving patients across the spectrum of liver disease. We sought a balanced stakeholder group. Each sponsoring life science company working on a NASH product was allocated two votes in the Delphi; however, the total number of industry votes was less than the total number of votes from nonindustry participants. The 2 participants from each company generally included 1 person from the clinical development side and another involved in market access; as these stakeholders represent different time points in the life cycle of a new product, they may also have differing opinions about which outcomes should be included in a COS.

IDENTIFICATION OF OUTCOMES

A list of candidate outcomes for the COS was developed from a targeted literature review and key informant interviews. Three sources were searched for relevant publications: 1) clinical trial records on clinicaltrials.gov, of which completed trials were matched to the corresponding published journal article; 2) the journal *Hepatology* for additional relevant articles discussing NASH outcomes; and 3) Cochrane Reviews on NAFLD and NASH identified in the Cochrane Library. The outcomes from each included study were abstracted as described in Table 1. Whenever possible, clinicaltrials.gov records were linked to a published manuscript (through national clinical trial [NCT] registration number). The literature review was used to compile a list of outcomes that are already in use in liver disease trials. Key informant interviews served as a complement to the targeted review to ensure that the candidate outcome list also reflected outcomes elicited from each interviewee's personal knowledge about and experience with NASH and to understand specific stakeholder perspectives and priorities regarding the most important outcomes to measure as treatments for NASH become available. Outcomes were grouped by domain according to COMET naming taxonomy guidance.[23]

TABLE 1. TARGETED LITERATURE SEARCH ABSTRACTION PLAN

Source	Time Frame for Search	Search Term	Outcomes Reviewed
clinicaltrials.gov	Records first posted in the previous 5 years (06/01/2013-06/13/2018)	NASH - nonalcoholic steatohepatitis	From the trial record*: • current primary outcome measures • current secondary outcome measures • current other outcome measures
Published journal articles	Previous 5 years of archives of the journal *Hepatology*	N/A; hand search of all titles to identify reports of clinical trials and other relevant articles	1. Outcomes reported in the Methods section. 2. Outcomes reported in the Results section (if required).
Cochrane Reviews	All available	NASH, NAFLD	Primary and secondary outcomes reported in the Methods section.

*Data elements required by the International Committee of Medical Journal Editors and the World Health Organization International Clinical Trials Registry Platform.
Abbreviation: Not applicable.

DELPHI VOTING

A modified Delphi consensus process was used to condense and prioritize outcomes on the candidate list. The list was reduced over three rounds of online voting using Qualtrics software. Definitions were provided, and participants rated each outcome on a scale of 1-9, where a score of 1-3 indicates that the outcome is not important to include in the COS, a score of 4-6 indicates that an outcome is important but not critical to include in the COS, and a score of 7-9 indicates that an outcome is critical to include in the COS.[24] Each outcome was rated twice: voters were first asked to consider the outcomes' importance in a core set for trials of treatments for NASH without cirrhosis (fibrosis stage F2-F3), and then the voters were asked to consider the outcomes' importance in a core set for trials of treatments for NASH with cirrhosis (F4). In the first round of voting, participants could suggest additional outcomes they felt should be added to the list. Suggestions were reviewed by the project team and, if relevant clinically or for patient quality of life and not overlapping with already included outcome concepts, added to the candidate list for the next round.

The list of outcomes was reduced according to predetermined consensus criteria (Table 2). Outcomes were eliminated from consideration if <70% of voters rated the outcome from 7 to 9, unless the patient stakeholder group average score was 7 or higher. This consensus criteria was developed based on recommendations in the COMET handbook,[25] in particular, the 70% threshold to retain, but the "patient-important" consideration was a special provision we added to ensure that patient-important and patient-reported outcomes were thoroughly considered. For the final round of voting, the patient-important criterion was dropped and the outcome had to reach high consensus to be included in the COS. After each voting round, participants were given a summary table containing for each outcome the mean overall score for all voters and the mean scores by stakeholder group. In Delphi rounds 2 and 3, participants were reminded of the score they personally selected in the previous round for each outcome plus an indication as to whether their selection was higher or lower than the mean scores for their stakeholder group.

CONSENSUS MEETING

Between Delphi rounds 2 and 3, we hosted an in-person consensus meeting in Baltimore, MD.

Discussion focused on the outcomes that had been retained on the list due to the patient-important criterion. Participants also had the opportunity to "rescue" an outcome that had been eliminated by requesting that it be moved back into the discussion. Voters could also suggest combinations of outcomes or refinements to definitions of outcomes that remained on the list. Proposals presented throughout the day were compiled into an online survey which was presented as a live vote at the end of the meeting.

After the final online voting round, we held a webinar for the voters to review the results. We included live voting during the webinar on four proposals meant to streamline the core sets.

Results

PARTICIPANTS

Fifty-three voters participated in the COS consensus process: 7 patients/patient advocates, 10 clinicians/researchers, 5 U.S. payers, 7 health technology assessors, 22 industry representatives, and 2 regulators (one each from the U.S. Food and Drug Administration [FDA] and European Medicines Agency [EMA]). Stakeholder groups included both U.S.-based and non-U.S.-based participants (Table 3). Participants who served as

TABLE 2. coreNASH CONSENSUS DEFINITIONS

Term	Definition
Consensus to select (high consensus)	An outcome in which ≥70% of all voters rated the outcome with a score of 7, 8, or 9 ("critical importance")
Consensus to select (patient-important)	An outcome in which <70% of all voters rated the outcome 7, 8, or 9, but the stakeholders in the patient group gave the outcome an average rating of ≥7
Consensus to eliminate	An outcome in which <70% of voters rated the outcome 7, 8, or 9, AND the patient group average rating was <7

TABLE 3. coreNASH PARTICIPANTS BY STAKEHOLDER GROUP AND LOCATION

Stakeholder Group	United States Based	Not United States	Total
Patients/patient advocates	3	4	7
Clinicians/researchers	5	5	10
U.S. payers	5	–	5
HTA	1	6	7
Industry	14	8	22
Regulators	1	1	2

*Not United States includes participants from Canada and countries in Europe.

industry representatives are included in Table 3, but it is important to note that although the individuals representing a life science company may be based in a certain country, their viewpoints on outcome selection may reflect priorities in several countries where a trial may be held or a new NASH treatment could be approved.

IDENTIFICATION OF OUTCOMES

We identified 143 clinical trial records on clinicaltrials.gov and six Cochrane Reviews in our targeted literature review. We excluded 46 clinicaltrials.gov records for reasons including studies that were not an intervention to treat NASH, studies that were specifically to observe the pharmacokinetics of a drug, or studies on the natural history of the disease. Journal articles describing clinical trials for NASH treatment were obtained and paired with their clinicaltrials.gov record. In addition, 31 other journal articles that met the inclusion criteria were included and outcomes were abstracted. We reviewed included studies from systematic reviews and searched for additional studies from reference lists of included studies. The outcomes from a total of 168 studies were abstracted; 12 additional journal articles were referenced but not abstracted (e.g., cohort studies of patients with NASH/NAFLD, letters to the editor describing an issue for patients with NASH that could be translated into an outcome or outcomes). Delphi panelists were invited to participate in an interview; the project team hosted interviews with 23 participants (7 patients/patient advocates, 5 clinicians/researchers, 11 industry).

DELPHI VOTING

The initial candidate list included 86 outcomes grouped into the following domains: biomarkers and chemical response, body composition, functioning (including physical functioning, emotional functioning, role functioning, global quality of life, perceived health status, and delivery of care categories), mortality/survival, physiological/clinical, and resource use. After round 1 of the Delphi, we held a webinar and confirmatory subsurvey with the patient/patient advocate group to discuss the outcomes slated to be retained due to the patient-important criteria. A total of 52 outcomes for the NASH without cirrhosis (F2-F3) COS and 82 outcomes for the NASH with cirrhosis (F4) COS moved into round 2 (Fig. 1). This included five new outcomes that were suggested by voters in round 1 and included on both lists (Supporting Information).

After round 2 and the in-person consensus meeting, there were 18 outcomes on the list for NASH without cirrhosis (F2-F3) and 32 outcomes on the list for NASH with cirrhosis (F4). These numbers reflect changes agreed to during our in-person meeting to reintroduce previously eliminated outcomes and combine others. After the completion of round 3 in which the patient-important criterion was dropped and outcomes needed to reach high consensus by the entire Delphi panel to be included, the resulting core sets had two outcomes for NASH without cirrhosis (F2-F3) and nine outcomes for NASH with cirrhosis (F4). The percentage of voters who rated each outcome as critical to include in a core set is listed in Table 4. The core sets determined after round 3 originally included one and three mortality outcomes for NASH without cirrhosis (F2-F3) and NASH with cirrhosis (F4), respectively; the voters agreed to separate the mortality outcomes from the main core set outcomes as the former would be measured in clinical trials as a regulatory requirement (Fig. 2).

Discussion

The decision to develop two separate core sets based on disease progression was made in advance of

TABLE 4. PERCENT OF PARTICIPANTS VOTING 7-9 ON REMAINING OUTCOMES IN DELPHI ROUND 3

	NASH Without Cirrhosis (F2-F3)		NASH With Cirrhosis (F4)	
Outcome	All Voters	Patient Group	All Voters	Patient Group
AST/ALT ratio	–	–	38	71
ALT levels	50	86	30	43
Ability to manage activities of daily living	–	–	62	86
Emotional anxiety	–	–	36	43
General health perception	–	–	44	71
Health-related global quality of life	68	71	76	71
Hepatic-related morbidity	62	43	82	86
Independence	–	–	46	86
Satisfaction with care/ treatment	12	29	–	–
Satisfaction with provider/care team	10	43	–	–
Work absenteeism	–	–	42	71
Blood pressure	18	29	20	14
Cardiovascular risk profile	–	–	30	43
Change in HVPG	–	–	52	57
Cirrhosis	66	57	82	100
Ballooning	66	57	–	–
Inflammation	68	71	58	100
Steatosis	64	71	–	–
Disease activity	78	100	66	57
Evidence of metabolic syndrome	26	29	26	29
Fatigue	56	71	64	71
Fibrosis stage	90	86	80	100
Global liver activity	46	86	50	86
Hepatocellular carcinoma	–	–	84	71
Liver stiffness	50	86	52	71
Clinical need for liver transplant	–	–	70	29
Portal hypertension	–	–	58	57
Pulmonary hypertension	–	–	30	29
Decompensation event	–	–	84	86
MELD score	–	–	82	86
Hospital stay	–	–	74	43
Caregiver burden	–	–	40	57
Mortality outcomes	Overall	Patient group	Overall	Patient group
All-cause mortality	68	43	88	57
Cardiovascular-related mortality	–	–	62	43
Hepatic-related mortality	70	43	88	57
Survival	–	–	76	57

Table shows outcomes remaining on the candidate outcome list for "NASH without cirrhosis" and "NASH with cirrhosis" at the start of round 3 of the Delphi. Outcome lists were identical at the start of round 1. A cell with "–" indicates that the outcome had been eliminated in an earlier voting round for that COS (NASH with or without cirrhosis) and was no longer a candidate to be voted on in round 3. Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; HVPG, hepatic venous pressure gradient.

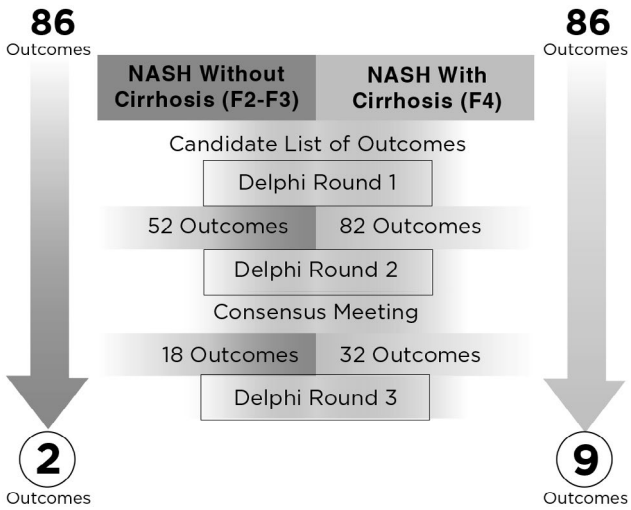

FIG. 1. An overview of the coreNASH Delphi. This was a consensus process to prioritize and condense a candidate outcome list to develop two COSs.

Delphi voting; confirmation to categorize the sets as being for NASH without cirrhosis (with a focus on F2-F3) and NASH with cirrhosis (F4) was agreed during the live vote at our in-person meeting. This split is in alignment with recent draft regulatory guidance that differentiates between noncirrhotic NASH and NASH with cirrhosis[18,20] and asserts that the current focus for treatment in NASH is for patients with F2-F4 fibrosis.[19] The coreNASH voters also agreed that the core set for patients with NASH with cirrhosis should only be used for patients with compensated cirrhosis. Patients with decompensated cirrhosis have a separate set of priorities, and treatments currently in the pipeline target points in the pathway toward decompensation.[26]

The final lists for the two sets represent differences between the two disease stages. The set for NASH without cirrhosis, a time when the course of the disease may be asymptomatic and potentially reversible, designates outcomes that measure the progression or resolution of disease, while the outcomes in the set for NASH with cirrhosis are those relevant to end-stage liver disease and indicators of a transition from compensated to decompensated disease, such as liver cancer, the need for a liver transplant, and decompensation events. The set for NASH with cirrhosis also includes a quality of life measure and a hepatic-related morbidity measure (defined as including complications of cirrhosis and fatigue, anxiety related to disease progression, loss of ability to work), underscoring that NASH at this stage influences functioning in a patient's daily life.

Previous multi-stakeholder work established case definitions and baseline parameters to be used in NASH clinical trials as well as defined terminology to describe changes in NAFLD/NASH in clinical trials relating to the trial endpoints recommended for regulatory review.[15] coreNASH expanded on that existing

	Core Outcome Set	Mortality Outcomes
NASH without cirrhosis (F2-F3)	• Disease stage • Fibrosis stage	• Hepatic-related mortality
NASH with cirrhosis (F4)	• Health-related quality of life • Hepatic-related morbidity • Cirrhosis • Fibrosis stage • Hepatocellular carcinoma • Clinical need for liver transplant • Decompensation event • MELD Score • Hospital stay	• All-cause mortality • Hepatic-related mortality • Survival

FIG. 2. The final coreNASH COS.

foundation by including payers and health technology assessors in a voting process and expanding the charge to consider outcomes that would be useful across the life cycle of a product and those that may be of particular use to postregulatory decision makers. coreNASH also used the Delphi approach, which provides each voter anonymity and an equal say[27] and can elevate the perspective of patients by using specific consensus rules to preserve patient-important outcomes. The range of outcomes preferred by various stakeholders is especially apparent in the core set for NASH with cirrhosis as there are a mix of outcomes that are useful to certain stakeholders as they make decisions about access and coverage of the product. In addition to the previously noted outcomes relevant to patient quality of life, Model for End-Stage Liver Disease (MELD) score and hospital stay are utilization outcomes that are important to payers as a cutoff for approval for a liver transplant and an indication of resource use, respectively. On the other hand, the outcome "clinical need for a liver transplant" is of high importance to clinicians and regulators along with fibrosis stage and decompensation events as markers of progression of disease.

A COS is not meant to limit what should be measured in a clinical trial but rather to serve as a minimum set to allow comparison between treatments and trials. While there are expected benefits from incorporating these outcomes into clinical development programs, the core set may be supplemented with other relevant outcomes. While the coreNASH core sets did not include any biomarkers, recommendations for metabolic and laboratory tests that may be added where capacity allows have been published and delineated for early versus late-phase trials.[14] NASH is a fast-moving area of therapeutic development, and as more research is completed, it is likely that the core set will need to be updated, in particular as biomarkers are discovered or as definitions change. As such, the coreNASH set is a fluid list that we plan to revisit periodically and make updates that respond to regulatory changes and newly identified biomarkers.

The final core sets reflect the expectation that outcome priorities may change as the field moves forward. In the COS for NASH with cirrhosis, the voters included two measures related to liver transplantation: MELD score and "clinical need for liver transplant." MELD score is used to allocate liver transplants and is based on short-term risk of death. It was included on our list as it is used and rated highly by clinicians and payers alike. Meanwhile, the outcome measuring clinical need for liver transplant is less specific and maintains an opening for redefinition as new research becomes available. Documenting the clinical need for a liver transplant rather than whether or not a person has actually received a transplant is a way to categorize those in need of a transplant based on worsening liver disease without being linked to regional organ availability, insurance coverage, severity of obesity and other comorbidities, and availability of necessary before and after operative social support that often determine whether or not a transplant is performed. The inclusion of both MELD score and clinical need for liver transplant outcomes is similar to recent recommendations by disease experts at the American Association for the Study of Liver Diseases/European Association for the Study of the Liver Joint Workshop on Clinical Trial Endpoints in NAFLD.[28] For suggested endpoints for phase 3 trials in those with cirrhotic NASH, they include both MELD progression from <12 to >15 and liver transplant, adding in a footnote that meeting severity of liver disease criteria for consideration for transplant should be considered rather than the number of liver transplants themselves.

For NASH without cirrhosis, the voters selected disease activity and fibrosis stage as the core outcomes. The definition we used for disease activity, "The findings reflective of activation of disease pathways driving the disease toward cirrhosis," is intentionally broad to allow for the possibility of an update to the accepted best measure without a need to change the outcome. Rather than locking into the definition the histologic descriptions of disease activity, although they are currently widely accepted, this broad description allows for the possibility that in the future a biomarker or panel of biomarkers may be validated as a measure of disease activity.

To identify alignment and future opportunity for use of the coreNASH COS, it is useful to review outcomes that are currently being used in NASH trials. The Safety and Efficacy of Selonsertib in Adults With NASH and Bridging (F3) Fibrosis (STELLAR-3; NCT03053050), Safety and Efficacy of Selonsertib in Adults With Compensated Cirrhosis Due to NASH (STELLAR-4; NCT03053063), and Phase 3 Study for the Efficacy and Safety of CVC for the Treatment of Liver Fibrosis in Adults With NASH (AURORA; NCT03028740) trials all report on the same primary endpoint, which is the proportion of participants who achieved a ≥1-stage improvement in fibrosis without worsening of NASH. The Phase 3 Study to Evaluate the Efficacy

and Safety of Elafibranor Versus Placebo in Patients With NASH (RESOLVE-IT; NCT02704403) used the inverse primary endpoint, evaluating NASH resolution without worsening of fibrosis. Still another study, the Randomized Global Phase 3 Study to Evaluate the Impact on NASH With Fibrosis of Obeticholic Acid Treatment (REGENERATE; NCT02548351), included both NASH resolution and fibrosis improvement as coprimary endpoints. Proposed terminology to describe "improvement," "resolution," and "no worsening" was recently published by the Liver Forum Case Definitions Working Group.[17] The primary endpoints used in these trials are in alignment with recommendations by the FDA and EMA in their recent guidance,[18,19] and if the studies had used the coreNASH core sets, all but STELLAR-4 would have used our set for NASH without cirrhosis. The outcomes in those trials compare with the outcomes in the coreNASH set for NASH without cirrhosis as each outcome we recommend can be derived from the measurements used. Improvement in NASH and worsening of NASH are described in these trials as measurements using the NAFLD activity score, which we currently recommend as the measure of disease activity to be used for the core set. There are gaps, however, between our recommended outcomes for more advanced disease. The coreNASH set for NASH with cirrhosis includes some patient-important, payer-important, and HTA-important outcomes that were not included in the STELLAR-4 trial. Endpoints for NASH trials across the development spectrum have been recommended by a joint group of disease experts, and there is good alignment with what the expert group recommended compared with the recommendation of the coreNASH multi-stakeholder group, which included both experts and nonexperts.[28]

There are limitations in the consensus process. Some COS developers undertake a full systematic review before beginning an outcome prioritization exercise, but these can significantly affect timelines. In order to expedite this project and provide a COS for timely inclusion in upcoming trials, we performed a targeted literature review. The candidate outcome list was generated in a systematic manner by exploring NASH/NAFLD-registered trials and making use of already completed systematic reviews. This search for trials was supplemented by searching for relevant articles from the past 5 years in the journal *Hepatology*. The literature in *Hepatology* was intended to represent the current thinking about outcomes in the field, but it is possible that a trial that had been published in another journal with an important outcome not identified in our search could have been missed by using this method. However, the goal of this review was not to exhaustively identify every individual NASH trial but rather to compile a comprehensive list of currently used outcomes that might be prioritized in a COS. In using the described method, we reached saturation quickly and we had several checkpoints throughout the process where participants could alert us to missing outcomes or make suggestions for outcomes to be added.

In this exercise, special weight was given to the view of the patients during the first two rounds of voting, but it is possible that a different set of patients would vote differently or would affect the vote of the entire Delphi panel by advocating for different outcomes. In addition, the approach engages nonexpert patients with experts. Challenges of this approach have been reported by COS developers; in particular, patients may struggle with prioritizing outcomes and tend to rate everything as important.[29] Because of this design and despite orientation sessions to facilitate constructive participation, some of the outcomes identified as potentially "patient important" for the in-person meeting may represent nonexpert misunderstandings rather than truly important outcomes. The meeting is designed to sort out these priorities and bring clarity to true patient (and other stakeholder) priorities; however, participating experts may end up dismissing nonexpert priorities in the final vote where outcomes cannot achieve the "high consensus" threshold without expert buy-in. In this project, some outcomes that were clearly patient important did not reach the threshold criterion to be included. In both COS, >70% of patients called fatigue a critical outcome to include, but the overall votes were not high enough to keep this on the list. Also, in the final round of voting for the list for NASH with cirrhosis, three outcomes were of high importance to patients but were not seen as important by the rest of the group (activities of daily living, general health perception, and independence; Table 4). For these reasons, while not part of the core set, we suggest that clinical trial designers seeking to include patient-important outcomes consider including these additional outcomes, especially fatigue, which in our meeting was discussed in especially compelling terms by the patients.

Additionally, there are inherent logistical challenges with an in-person consensus meeting. Not every member of the Delphi panel is able to be present, and among those who attend, there is a possibility that

when face-to-face with the larger group, some participants may be less vocal than others. We attempted to address these concerns by inviting those who could not be physically present to participate over the phone and employing a trained moderator to lead the discussions and attempt to provide equity during conversations. The final vote was completed online after the in-person meeting, resuming the anonymity of the Delphi method to capture this final assessment.

Another limitation of this approach is that the resulting core set does not clarify when included outcomes should be measured. For the candidate outcome list, we did not differentiate between short-term outcomes and long-term outcomes. Some long-term outcomes were included in our COS (hepatocellular carcinoma, need for liver transplantation, decompensation events), but it may not be possible to measure some over the course of a clinical trial. In this case, it is recommended that these outcomes are included in follow-up trials and post-market studies. Some of the phase 3 trials have a protocol in two parts, first examining histopathologic improvement and the second including a long-term outcome assessment. In a trial where it is not possible to measure long-term outcomes, fibrosis stage, which was included in both of our core sets, may serve as a surrogate as it has been reported that NASH associated with fibrosis increases progression to liver-related hard outcomes, such as liver transplantation, hepatocellular carcinoma, cirrhosis, and death.[30-32]

coreNASH achieved multi-stakeholder consensus on a set of core outcomes to be used in clinical trials for interventions for NASH. By defining and uniformly implementing a set of outcomes to be included across trials and products, benefits can be seen at several key points in the life cycle of the product. Such benefits include: 1) as a complement to the standard requirements for regulatory decision making, a COS ensures that outcomes are defined, collected, and reported in all clinical trials; 2) for health care payers and HTA organizations, a COS increases predictability and consistency in technology assessments that inform payer policies for coverage and reimbursement; 3) for on-market use, a COS establishes recommendations for outcomes that are meaningful to the quality of life and functioning of patients.[33] Use of the core set in clinical development programs will ensure that the evidence on these outcomes of importance to postregulatory decision makers is collected, reported, and used to make decisions about getting the appropriate treatments to patients with NASH of varying disease stages.

Acknowledgment: We thank the coreNASH Advisory Committee members and the coreNASH Delphi Panel (see Appendix A).

Appendix A

THE coreNASH PANEL[1]

William Baldyga (The Liver Forum); Maria-Magdalena Balp (Novartis Pharma AG); Catherine Brun-Strang (Sanofi); Robyn Carson (Allergan); Bosco D'Sa (Allergan); Wayne Eskridge (Fatty Liver Foundation); Kurinchi Gurusamy (University College London); Lars Hansen (AstraZeneca); Morten Hansen (Novo Nordisk, Inc.); Katy Harrison (National Institute for Health and Care Excellence); Andras Harsanyi (Eötvös Loránd University); Anders Hvelplund (Novo Nordisk, Inc.); Zoltán Kaló (Syreaon Research Institute and Eötvös Loránd University); Marko Korenjak (European Liver Patients Association [ELPA]); EJ Milne (Coventry University); Euan McLeod (Pfizer); Michelle Mocarski (Novo Nordisk, Inc.); Joachim Musaeus (European Medicines Agency [EMA]); Fady Ntanios (Pfizer); Mazen Noureddin (Cedars Sinai Medical Center); Daniel Ollendorf (Institute for Clinical and Economic Review); Jan Oscarsson (AstraZeneca); Stephen Rossi (NGM Biopharmaceuticals); Vlad Ratziu (Hôpital Pitié Salpêtrière); Yaron Rotman (National Institute of Diabetes and Digestive and Kidney Diseases [NIDDK]); Leonardo Ruiz-Casas (HCD Economics); Lewis Sandy (UnitedHealth Group); Jörn M. Schattenberg (Metabolic Liver Research Program, University Medical Center Mainz, Germany); Mohammad Shadab Siddiqui (Virginia Commonwealth University School of Medicine); Gregory Steinberg (McMaster University); Ethan Weiss (University of California San Francisco); Keith White (Intercept Pharmaceuticals); José Willemse (Dutch Liver Patients Association); Kathy Wright (Rotherham Metropolitan Borough Council); Michael Zemel (NuSirt Biopharma).

[1] The full Delphi panel consisted of 53 voters. All participants were contacted and given the opportunity to be listed in the manuscript as a panelist. Those who responded are included in Appendix A. Those who did not respond or those who indicated that they did not want to be included were not listed.

REFERENCES

1) Estes C, Razavi H, Loomba R, Younossi Z, Sanyal AJ. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. Hepatology 2018;67:123-133.
2) Lazo M, Clark JM. The epidemiology of nonalcoholic fatty liver disease: a global perspective. Semin Liver Dis 2008;28:339-350.
3) **Goldberg D**, **Ditah IC**, Saeian K, Lalehzari M, Aronsohn A, Gorospe EC, et al. Changes in the prevalence of hepatitis C virus infection, nonalcoholic steatohepatitis, and alcoholic liver disease among patients with cirrhosis or liver failure on the waitlist for liver transplantation. Gastroenterology 2017;152:1090-1099.e1.
4) Vernon G, Baranova A, Younossi M. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther 2011;34:274-285.
5) **Fazel Y**, **Koenig AB**, Sayiner M, Goodman ZD, Younossi ZM. Epidemiology and natural history of non-alcoholic fatty liver disease. Metabolism 2016;65:1017-1025.
6) Connolly JJ, Ooka K, Lim JK. Future pharmacotherapy for non-alcoholic steatohepatitis (NASH): review of phase 2 and 3 trials. J Clin Transl Hepatol 2018;6:264-275.
7) Rotman Y, Sanyal AJ. Current and upcoming pharmacotherapy for non-alcoholic fatty liver disease. Gut 2017;66:180-190.
8) Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. Nat Med 2018;24:908-922.
9) Sanyal A, Charles ED, Neuschwander-Tetri BA, Loomba R, Harrison SA, Abdelmalek MF, et al. Pegbelfermin (BMS-986036), a PEGylated fibroblast growth factor 21 analogue, in patients with non-alcoholic steatohepatitis: a randomised, double-blind, placebo-controlled, phase 2a trial. Lancet 2019;392:2705-2717.
10) Loomba R, Lawitz E, Mantry PS, Jayakumar S, Caldwell SH, Arnold H, et al. The ASK1 inhibitor selonsertib in patients with nonalcoholic steatohepatitis: a randomized, phase 2 trial. Hepatology 2018;67:549-559.Erratum in: Hepatology 2018;67:2063.
11) **Friedman SL**, **Ratziu V**, Harrison SA, Abdelmalek MF, Aithal GP, Caballeria J, et al. A randomized, placebo-controlled trial of cenicriviroc for treatment of nonalcoholic steatohepatitis with fibrosis. Hepatology 2018;67:1754-1767.
12) Ratziu V, Harrison SA, Francque S, Bedossa P, Lehert P, Serfaty L, et al.;GOLDEN-505 Investigator Study Group. Elafibranor, an agonist of the peroxisome proliferator-activated receptor-alpha and -delta, induces resolution of nonalcoholic steatohepatitis without fibrosis worsening. Gastroenterology 2016;150:1147-1159.e5. Erratum in: Gastroenterology 2017;152:2084.
13) Neuschwander-Tetri BA, Loomba R, Sanyal AJ, Lavine JE, Van Natta ML, Abdelmalek MF, et al.; NASH Clinical Research Network. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. Lancet 2015;385:956-965.Erratum in: Lancet 2015;385:946.
14) Hameed B, Terrault NA, Gill RM, Loomba R, Chalasani N, Hoofnagle JH, et al.; NASH CRN. Clinical and metabolic effects associated with weight changes and obeticholic acid in non-alcoholic steatohepatitis. Aliment Pharmacol Ther 2018;47:645-656.
15) **Siddiqui MS**, **Harrison SA**, Abdelmalek MF, Anstee QM, Bedossa P, Castera L, et al.; Liver Forum Case Definitions Working Group. Case definitions for inclusion and analysis of endpoints in clinical trials for nonalcoholic steatohepatitis through the lens of regulatory science. Hepatology 2018;67:2001-2012.
16) Patel YA, Imperial JC, Muir AJ, Anstee QM, DeBrota D, Dimick-Santos L, et al.; Liver Forum's Data Standardization Working Group. Baseline parameters in clinical trials for nonalcoholic steatohepatitis: recommendations from the Liver Forum. Gastroenterology 2017;153:621-625.e7.
17) Cheung A, Neuschwander-Tetri BA, Kleiner DE, Schabel E, Rinella M, Harrison S, et al.; Liver Forum Case Definitions Working Group. Defining improvement in nonalcoholic steatohepatitis for treatment trial endpoints: recommendations from the Liver Forum. Hepatology 2019;70:1841-1855.
18) U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Noncirrhotic nonalcoholic steatohepatitis with liver fibrosis: developing drugs for treatment. Guidance for industry. Draft guidance. https://www.fda.gov/media/119044/download. Published December 2018. Accessed January 2021.
19) European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP). Reflection paper on regulatory requirements for the development of medicinal products for chronic non-infectious liver diseases (PBC, PSC, NASH). https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-regulatory-requirements-development-medicinal-products-chronic-non-infectious-liver_en.pdf. Published November 19, 2018. Accessed June 2020.
20) U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER).Nonalcoholic steatohepatitis with compensated cirrhosis: developing drugs for treatment guidance for industry. https://www.fda.gov/media/127738/download. Published June 2019. Accessed January 2021.
21) Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials 2012;13:132.
22) Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core outcome set-STAndards for Development: the COS-STAD recommendations. PLoS Med 2017;14:e1002447.
23) Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. J Clin Epidemiol 2018;96:84-92.
24) Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. Trials 2017;18(Suppl. 3):280.
25) Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 2011;64:395-400.
26) Banini BA, Sanyal AJ. Current and future pharmacologic treatment of nonalcoholic steatohepatitis. Curr Opin Gastroenterol 2017;33:134-141.
27) Yousuf MI. Using experts' opinions through Delphi technique. PARE 2007;12. https://scholarworks.umass.edu/pare/vol12/iss1/4/?utm_source=scholarworks.umass.edu%2Fpare%2Fvol12%2Fiss1%2F4&utm_medium=PDF&utm_campaign=PDFCoverPages.
28) **Rinella ME**, **Tacke F**, **Sanyal AJ**, **Anstee QM**. Report on the AASLD/EASL joint workshop on clinical trial endpoints in NAFLD. Hepatology 2019;70:1424-1436.
29) Gargon E, Williamson PR, Young B. Improving core outcome set development: qualitative interviews with developers provided pointers to inform guidance. J Clin Epidemiol 2017;86:140-152.
30) Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. Gastroenterology 2015;149:389-397.e10.

31) Dulai PS, Singh S, Patel J, Soni M, Prokop LJ, Younossi Z, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. Hepatology 2017;65:1557-1565.
32) **Hagstrom H**, **Nasr P**, Ekstedt M, Hammar U, Stal P, Hultcrantz R, et al. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. J Hepatol 2017;67:1265-1273.
33) Skinner MW, Clearfield E, Iorio A, Tunis SR. Comparing outcomes across clinical trials: core outcome set for hemophilia gene therapy as a model for other diseases. DIA Global Forum 2017;9:16. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjX3eHz_p3uAhVEp1kKHSTnA7oQFjAAegQIARAC&url=https%3A%2F%2Fnews.wfh.org%2Fwp-content%2Fuploads%2FComparing-Outcomes-Across-Clinical-Trials.pdf&usg=AOvVaw2u-Q32FAjIqZlGP9MWas2z.

Author names in bold designate shared co-first authorship.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep4.1678/suppinfo.