THE FORUM FOR COLLABORATIVE RESEARCH

LIVER FORUM 10

Summary of Proceedings

Thursday, September 19, 2019
Friday, September 20, 2019
Bethesda, MD

1608 Rhode Island Avenue NW
Washington, DC 20036
www.forumresearch.org

THE FORUM
For Collaborative Research℠

Berkeley Public Health

CONTENTS

DAY 1, SESSION IV: LIVER FORUM UPDATES

Liver Forum Project Updates

Presenter: Katherine Barradas, Forum for Collaborative Research

Slides: https://bit.ly/3qd8mEV

Liver Forum Overview

- Platform for ongoing, continuous, multi-stakeholder dialogue to identify barriers, prioritize research, and identify solutions to accelerate therapeutic development for NAFLD/NASH.
- Provides a neutral, independent, safe space for discussion and deliberation across stakeholder groups.
- Focus on developing consensus, increasing synergy and collaboration, and reducing duplication and uncertainty.
- Ongoing working group activity throughout the year and anchored by larger project events
- Dependent on the active and engaged participation of members
- Discussions during meeting proceedings are not for attribution, and participants speak as individuals and express views that may not represent those of their organizations.

Recent Activities and Accomplishments

- Responses submitted to regulatory guidance documents
    - EMA Reflection paper: Regulatory Requirements for the Development of Medicinal Products for Chronic Non-Infectious Liver Diseases (PBC, PSC, NASH)
    - FDA draft guidance: Nonalcoholic Steatohepatitis with Compensated Cirrhosis: Developing Drugs for Treatment
- Manuscripts
    - Submitted: Standard of Care Lifestyle Management Working Group Manuscript
        - Standardization of Diet and Exercise in Clinical Trials of NAFLD-NASH: Recommendations from the Liver Forum
        - Oliver Glass, Claudia Filozof, Mazen Noureddin, Mark Berner-Hansen, Elmer Schabel, Katherine Barradas, Jörn M. Schattenberg, Veronica Miller, Sven Francque, Manal F. Abdelmalek, for the Liver Forum Standard of Care Working Group
    - Accepted: Pediatric Issues Working Group Manuscript
        - Factors to Consider in Development of Drugs for Pediatric Nonalcoholic Fatty Liver Disease
        - Miriam B. Vos, Lara Dimick-Santos, Ruby Mehta, Stephanie O. Omokaro, Johannes Taminiau, Elmer Schabel, David E. Kleiner, Peter Szitanyi, Piotr Socha, Jeffrey B. Schwimmer, Stephanie Noviello, Debra G. Silberg, Richard Torstenson, Veronica Miller, Joel E. Lavine, on behalf of the Liver Forum Pediatric Issues Working Group
        - https://doi.org/10.1053/j.gastro.2019.08.048

Working Group Status and Associated Manuscripts

- Active/Open
    - Start of Care Comorbidity Management
        - Manuscript under development
    - Pediatric Issues
        - Factors to Consider in Development of Drugs for Pediatric Nonalcoholic Fatty Liver Disease: https://doi.org/10.1053/j.gastro.2019.08.048
    - NASH Cirrhosis
        - Manuscript under development

- Active/Closed
  - Estimands in NASH Clinical Trials
- To be Assessed
  - Standard of Care Lifestyle Management
    - Manuscript submitted for publication
- Completed
  - Data Standardization
    - Baseline Parameters in Clinical Trials for Nonalcoholic Steatohepatitis: Recommendations From the Liver Forum: https://doi.org/10.1053/j.gastro.2017.07.024
  - Case Definitions
    - Case Definitions for Inclusion and Analysis of Endpoints in Clinical Trials for Nonalcoholic Steatohepatitis Through the Lens of Regulatory Science: https://doi.org/10.1002/hep.29607
    - Defining Improvement in Nonalcoholic Steatohepatitis for Treatment Trial Endpoints: Recommendations from the Liver Forum: https://doi.org/10.1002/hep.30672

DAY 1, SESSION V: BIOMARKER DEVELOPMENT

Opportunities in Biomarker Development

Presenter: Christopher Leptak, U.S. Food and Drug Administration

Slides: https://bit.ly/2XzKIq4

FDA Regulatory Approach to Biomarkers

- BEST: Biomarkers, Endpoints, and Other Tools
  - Glossary of terminology and uses of biomarkers and endpoints in basic biomedical research, medical product development, and clinical care
  - Developed by NIH-FDA Biomarker Working Group
  - http://www.ncbi.nlm.nih.gov/books/NBK326791/
- Biomarker: a defined characteristic that is measured as an 1) indicator of normal or pathogenic biological processes or 2) response to an intervention.
  - Broadly defined, with multiple biomarker types including molecular, histologic, radiographic, and physiologic. (i.e., serum protein, change in tumor size by imaging study, algorithm for QT determination on ECG)
  - Characteristic is not a clinical assessment of how a patient feels, functions, or survives (contrasted with Clinical Outcome Assessments)
  - Although a biomarker may be used by clinical or basic science research communities, regulatory acceptance focuses on a drug development context that is supported by data for that context.
- Biomarker Classes
  - Susceptibility/Risk: Indicates potential for developing disease or medical condition in an individual who does not currently have clinically apparent disease or the medical condition
  - Diagnostic: Detects or confirms the presence of a disease or condition of interest or to identify individuals with a subset of the disease
  - Monitoring: Assesses status, through serial measurement, of a disease or medical condition including degree or extent of disease

  - Prognostic: Identifies likelihood of a clinical event, disease recurrence or progression, in patients who have the disease or medical condition of interest in the absence of a therapeutic intervention
  - Predictive: Identifies patients who are more likely to experience a favorable or unfavorable effect from a specific treatment
  - Pharmacodynamic/Response: Indicates that a biological response has occurred in a patient who has received a therapeutic intervention. May become clinical trial endpoints and for a very small subset, surrogate endpoints.
  - Safety: Indicates the likelihood, presence, or extent of toxicity to a therapeutic intervention when measured before or after that intervention
- Considerations for Biomarker Utility
  - Context of Use (COU): 1) BEST biomarker category and 2) how the biomarker impacts the clinical trial or drug development program.
    - What question is the biomarker intended to address?
      - Inclusion/exclusion criteria for prognostic or predictive enrichment?
      - Alter treatment allocation based on biomarker status?
      - Result in cessation of a patient's participation in a clinical trial because of safety concern?
      - Result in adaptation of the clinical trial design?
      - Establish proof of concept for patient population of interest?
      - Support clinical dose selection for first in human or Phase 3 studies?
      - Evaluate treatment response (e.g. pharmacodynamic effect)?
      - Support regulatory acceptability of a surrogate endpoint for accelerated or traditional approval?
    - Example: "Total Kidney Volume, measured at baseline, is a prognostic enrichment biomarker to select patients with ADPKD at high risk for a progressive decline in renal function (defined as a confirmed 30% decline in the patient's estimated glomerular filtration rate (eGFR)) for inclusion in interventional clinical trials. This biomarker may be used in combination with the patient's age and baseline eGFR as an enrichment factor in these trials."[1]

Biomarker Integration into Drug Development

- Three biomarker development pathways available and often parallel efforts are ongoing within multiple pathways.
- Drug Approval Process
  - Individual drug sponsors are either using established biomarkers or putting forward novel ones and engaging in conversation with agency
  - Most common way in which biomarkers are used, but not necessarily the way that novel biomarkers get established for regulatory purposes
  - Most of the biomarkers that are in use in this process are the long established ones.
- Scientific Community Consensus
  - Information in the public domain: recommendations from professional societies, publications in scientific journals, etc
  - Consensus over time, finding that data about the biomarker is consistent
  - Challenge is often lack of primary data available in public domain
- Biomarker Qualification Program
  - Mission is to work with external stakeholders to develop biomarkers as drug development tools

[1] Qualification of Biomarker Total Kidney Volume in Studies for Treatment of Autosomal Dominant Polycystic Kidney Disease Draft Guidance for Industry. 2016. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/qualification-biomarker-total-kidney-volume-studies-treatment-autosomal-dominant-polycystic-kidney

- Opportunity to pool resources, share costs, and engage outside experts and stakeholder groups.
- The outcome is a public guidance
- Group effort and data sharing may be a challenge

Biomarker Development and Qualification

- Enablers for biomarker development
  - Data standards (CDISC)
    - Common definitions of disease
  - Data quality
  - Data reproducibility
  - Data sharing
  - Assay/imaging pre-analytic standardization
  - Assay/imaging protocols/ SOPs
  - Evaluating impact on clinical trial elements
- Analytical assay and clinical validation considerations
  - The specific context of use for a biomarker will drive the extent of evidence needed for qualification
  - Analytical and clinical validation often occurring in parallel for novel biomarkers
  - Analytical validation
    - Establishes performance and acceptance characteristics of the biomarker assay
  - Clinical validation
    - Establishes that the biomarker acceptably identifies, measures, or predicts the concept of interest
- Biomarker Qualification
  - 21st Century Cures Act and PDUFA VI has placed FDA as an active participant in the drug development process, and facilitates early conversations about ideas and direction of a development program.
  - Three step process:
    - Letter of Intent (LOI): concise document that describes the DDT, a relevant drug development need, and a proposed COU (only one COU). Includes brief scientific rationale to support the DDT and COU.
      - Feasibility assessment of proposal will include information to support that measurement of the novel DDT is, in fact, possible.
    - Qualification Plan (QP): description of available relevant data, knowledge gaps, proposed data collection plans, and analysis plans. Full study protocols and analytic plans should be included.
      - Project development plan from concept to information to be developed/provided to support the DDT's COU. For biomarkers, to determine clinical utility and clinical validation, important to know that the analytical validation has been completed and information submitted to QP.
    - Full Qualification Package (FQP): final stage which includes descriptions of all studies, analyses, and results related to DDT and COU. Evidence should include full study protocols and reports, statistical or quantitative analysis plans, summary data.
      - Review of data to support the clinical validation of the DDT for the COU
  - Transparent process
  - Setting and implementing reasonable timeframes for submission review and decisions
- Transparency Provisions

    - All interested parties know what tools are in development, stage of development, and FDA determinations including rationale
    - Information about the submission and FDA's determination including recommendations are posted on DDT website
- Acceptance of biomarker into qualification
    - Acceptance decision for each submission (LOI, QP, FQP) based upon scientific merit:
        - Does the proposal address an impactful drug development need?
        - Is there enough information to suggest a likelihood of success?
        - What is the feasibility of the proposed analytical biomarker measurement approach?
    - Prioritization of review of submissions based upon:
        - "the severity, rarity, or prevalence of the disease or condition targeted by the drug development tool and the availability or lack of alternative treatments for such disease or condition; and
        - the identification by the Secretary or by biomedical research consortia and other expert stakeholders, of such drug development tool and its proposed context of use as a public health priority"
- Composite biomarkers
    - Terminology: composite, panel, multi-modal, score, index, etc.
        - Start by listing out the individual biomarker components that make up the composite
        - Understand that the list will likely evolve and change over time
    - Strategies:
        - Option: With a single COU, explore each biomarker/measurement methodology individually, and then build into composite once value of each component has been demonstrated
            - Works well with a few (3-5) components, but not as well for 10+ components due to challenges with studying each individually
        - Option: With a single COU, start with a composite of the most promising candidates and then refine over time
    - Measurement Scenarios:
        - Each member of the composite is measured with a separate platform. Each platform has to be independently validated. The readouts are then "manually" transformed into a composite
            - Single composite readout/interpretation = cut points (ex. Mild, Moderate, Severe)
            - Group assessment: still have cut points for "biomarker positivity" but interpretation simplified (ex., if any 2/7 positive, then composite is "positive")
        - Individual biomarkers of the composite are measured by a single device
            - Does the device readout individual values or combine them?
            - If combined into a score, the algorithm which generates the score needs to be provided (how components are weighted, etc)
- Internal Review: Three Tiers
    - DDT program assessment and recommendations
        - Work with requestor to clarify DDT, COU, project proposal, ensure administrative components are completed
        - Provide tool-specific recommendations based on past and ongoing projects
    - Discipline-specific SME assessment and recommendations
        - Includes OND division management participation
        - Evaluation based on regulatory precedent, current disease-specific challenges, level of impact on drug development programs

    - CDER DDT committee assessment, recommendations, and decision
        - Opportunity for broad senior CDER input early and throughout in the process
        - Works towards greater consistency across therapeutic areas and divisions

Biomarker Resources

- List of Qualified Biomarkers: https://www.fda.gov/drugs/biomarker-qualification-program/list-qualified-biomarkers
- Biomarker Qualification Submissions: https://www.fda.gov/drugs/biomarker-qualification-program/biomarker-qualification-submissions
- Table of Surrogate Endpoints: https://www.fda.gov/drugs/development-resources/table-surrogate-endpoints-were-basis-drug-approval-or-licensure
- Surrogate Endpoint Resources for Drug and Biologic Development: https://www.fda.gov/drugs/development-resources/surrogate-endpoint-resources-drug-and-biologic-development
- PDUFA VI: https://www.fda.gov/media/99140/download

NASH Consortia Update: LITMUS

Presenter: Quentin Anstee, Newcastle University

Overview

- NAFLD/NASH is a moving target with substantial inter-patient variation in disease natural history, rate of disease progression, and outcome
- Steatohepatitis is the biological driver of disease progression and fibrogenesis, and fibrosis is the best predictor of long-term outcomes.
- Reliance on liver biopsy
    - Significant proportion of the population have NAFLD, but only a minority progress to advanced liver disease or morbidity/mortality
    - Biopsy is a useful diagnostic, prognostic, and monitoring tool. It can also be used to assess pharmacodynamics/ response as a likely surrogate.
- Problems with liver biopsy
    - Sampling issues, patient tolerability, expensive
    - Challenging both in drug development as well as clinical practice
- Common biomarker needs for clinical practice and drug development
    - Is the diagnosis NASH
    - How active or advanced is the disease
    - Is disease stable, progressing, regressing
    - Is treatment or intervention working

LITMUS – Liver Investigation: Testing Marker Utility in Steatohepatitis

- Funded through EU IMI, 2017-2022, public-private co-funding model
- Strong and collaborative relationships with FNIH NIMBLE Consortium, FDA, and EMA
- 53 partners: 29 academic, 23 industry, 1 professional society
- Recruiting in 14 countries internationally
- Objectives: to develop, robustly validate and advance towards regulatory qualification biomarkers that diagnose, risk stratify and/or monitor NAFLD/NASH progression and fibrosis stage for use in drug discovery.

Recent Progress

- European NAFLD Registry: recruiting in 14 countries, with samples sent to a centralized biobank.

    - Collecting dataset on individuals, clinical data, imaging data, patient reported outcome measures, and biological samples.
    - All centers operating according to a single trial protocol and have a harmonized laboratory manual for all biological sample handling, and range of questionnaires.
    - ~7,000 individuals in the registry representing the spectrum of disease.
- WP2: Methodological Evaluation and Data Synthesis
    - Prepared detailed data management plans
    - Report on minimally acceptable performance criteria for biomarkers
    - Series of detailed systematic reviews summarizing the existing evidence and utility for a range of biomarkers
    - Data analysis is underway from the first batch of cases from the Metacohort
- WP4: Central Laboratories
    - Validation reports and quality control checks for first prioritized biomarker assays
        - Precision, accuracy, sensitivity, linearity, CLSI validation for some
    - Full range of biomarker assays set up and assay performance completed
- WP7: Qualification, Exploitation, and Dissemination
    - Regulatory interactions with EMA and FDA
    - Submission of letters of intent and briefing books

NASH Consortia Updates: NIMBLE

Presenter: Arun Sanyal, Virginia Commonwealth University

Slides: https://bit.ly/3nC1PC7

NIMBLE Overview:

- Public-private partnership established by the FNIH, including stakeholders from academia, industry, and the NIH, in order to advance non-invasive tests in the assessment of NAFLD.
    - Stakeholders include 12 funding companies, 9 academic centers, FDA, NIH, and biomarker companies
- Project Goal: leverage state of the art contemporary scientific tools to qualify strategically relevant biomarkers to enable timely development of NASH therapeutics
- Biomarker-related questions from clinician's perspective:
    - Is NAFLD/NASH likely to develop?
    - Is NAFLD present?
    - Is the patient likely to die from NASH?
    - What intervention is needed?
    - Is the disease trajectory changing?
- "What is the risk of liver-related outcome" is highest priority target
    - Very high impact and critical to determine who requires drug/surgical/endoscopic intervention.
- "Is the disease trajectory changing, with or without intervention) is 2nd highest priority
    - High impact and is needed to determine when to intervene, assess disease progression/ regression and impact of therapy.

Key Milestones Since Launch

- Contracting with academic centers, key collaborators, CROs, vendors
- Protocol and study design development
- Regulatory submissions
- Governance structure, COIs, project management
    - COI policy: different levels of conflict of interest for academic investigators, industry investigators, and key collaborators.
        - Firewalls in place to ensure integrity of data
    - Steering committee

        - Global Liver Institute has joined as patient representative to inform patient perspective into the design and execution of the project.
        - FDA representatives have joined to provide guidance from regulatory standpoint.
            - The representatives will not be involved in the FDA review
- Overall project plan has been approved
    - Stage 1: retrospective analysis and method studies
    - Stage 2: prospective study including circulating, functional, and imaging markers

Recent Progress

- Imaging Work Stream
    - Working to harmonize how data is read out across multiple imaging platforms to understand how results from one platform relate to results from another.
        - Ultrasound, MRI, VCTE
    - Working to finalize protocol with SAP, and contract with sites and FNIH
    - ICF approval
    - Recruiting patients
- Circulating Markers Work Stream
    - Circulating biomarkers flagged for review and inclusion based on literature review
    - Protocol drafted
    - CROs and vendors identified and going through contracting
    - Submission of draft LOI to FDA
    - Critical steps
        - Methodological issues: sample collection, storage and transport, analysis, quantification and internal/external controls, data reporting
        - Study design: populations, standardization of collection of meta-data, analytic issues (determination of disease activity, separation from F0 or F1 vs higher stages, separation of F4 from lower stages).
- Regulatory Submission
    - Submitted LOI to FDA, received feedback and working on revision.

DAY 2, SESSION I: NASH COMBINATION THERAPY

Combination Approaches for Pre-Cirrhotic NASH

Presenter: Brent Neuschwander-Tetri, Saint Louis University
Slides: https://bit.ly/37HNgZ9

Pathogenesis of NASH

- Substrate overload lipotoxic liver injury: free fatty acids promote the generation of toxic lipids
    - Common question: what about 2-hit hypothesis? (steatosis + oxidant stress, lipid peroxidation, injury)
        - Degree of steatosis does not correlate with NASH severity or outcomes
        - Oxidant stress occurs but not yet shown to play a role in NASH
    - Fatty acids come from two major sources:
        - De novo lipogenesis, ~5-25%
        - Adipose tissue and release of fatty acids through lipolysis, 75-95%
            - Adipose insulin resistance
        - Minor sources: autophagy, membrane turnover, lipoprotein remnant uptake
    - Burn substrate prior to reaching liver though exercise, brown adipose tissue, other metabolically active tissues throughout body; or, liver can oxidize some through mitochondrial beta-oxidation.

- If not burned off, re-esterify into glycerol and make triglyceride, typically pumped out of the liver and incorporated into very low density lipoprotein (VLDL)- complex process requiring enough amino acids to make ApoB 100 protein.
    - Protein starvation leads to fat buildup in the liver as there is not enough amino acid supply to make ApoB 100 and get the fat out of the liver. Also need phosphatidylcholines and choline source.
    - If cannot get out of the liver, triglycerides accumulate as lipid droplets, turn over through lipolysis and releases fatty acids back into the liver
        - Likely where PNPLA3 (most common genetic polymorphism that is associated with progression of fatty liver disease) plays a role in regulating where and when that happens.
- Lipids set off hepatocellular injury and wound repair response that stimulates inflammation and leads to the phenotype of NASH, fibrosis, and HCC
    - Modifiers: gut microbiome, cholesterol, uric acid, adipokines
    - Hypothesis that inflammation contributes to the injury, but unproven

Targets of Combination Therapy Trials

- Fructose, glucose, adipose tissue: diet, bariatric surgery, FGF21, MCH1Ri
- Muscle, brown adipose tissue: exercise, TGR5, FGF19/21, GLP-1, PPAR-δ, uncouplers, SGLT2i
- De novo lipogenesis: FXR, FGF19, ACYLi, ACCi, FASNi, SCDi
- Adipose tissue -> circulating free fatty acids: PPAR-γ, anti-imflamm
- Free fatty acids -> triglyceride: DGAT2i
- Hepatic injury -> fibrosis: Galectin-3i, Integrin-I, RAASi
- Inflammation: CCR2/5i, PDEi, ASK1i, Inflammasome-i
- Hepatic injury: Caspase-I, hedgehog
- Mitochondrial beta-oxidation: PPARα, MPCi, THR-β, uncouplers
- Modifiers: probiotics, TLRi, Anti-LPS, Statins

Examples

- FXR + ACCi + ASK1i
    - FXR will globally down-regulate de novo lipogenesis by down regulating SREBP1c, the transcription factor that up-regulates those enzymes. Will also potentially have an effect on energy disposal and brown adipose tissue.
    - ACC inhibitor- inhibiting specific enzyme involved in de novo lipogenesis
    - Addition of ASK1i to impair inflammation
- FGF19/21 + ACCi + PPAR-γ (hypothetical)
    - Could use FGF19 and FGF21 together to increase energy disposal, ACCi to inhibit de novo lipogenesis, and PPAR-γ to inhibit release of fatty acids out of adipose tissue
- Peptide polyagonists[2]
    - Target GLP1 receptor, gastric inhibitory peptide receptor, or glucagon receptor
        - Some can hit all three receptors
        - Since there are multiple pathways, if you hit one, likely the others compensate for this. By hitting multiple receptors, hopefully there is a greater change of beneficially changing metabolism.
- FGF19 + FGF21

[2] Sloop KW, et al. Beyond Glucagon-like Peptide-1: Is G-Protein Coupled Receptor Polypharmacology the Path Forward to Treating Metabolic Diseases? ACS Pharmacology & Translational Science. 2018;1(1):3-11.

    - FGF19 released from ileum in response to FXR being stimulated with bile acids
        - Hits FGFR4 receptor: decreases gluconeogenesis in the liver, decreases bile acid synthesis, increases FGF21 expression, decreases food intake, increases energy disposal
    - FGF21 released by liver in response to ER stress (ketogenic diet, ChREBP, fructose, alcohol)
        - Hits FGFR1c, FGFR2, FGFR3 receptors: in adipose tissue this increases glucose uptake, browning, and energy expenditure
    - Hypothetical combination- possible to get positive effects of each
- Hypothetical de-escalating NASH combination therapy (future)
    - For NASH with advanced fibrosis, targeting 3 processes:
        - Underlying metabolic disease (metabolic correction) → indefinitely
        - Hepatocyte stress/inflammation (anti-NASH) → 1 year?
        - Stellate cell activation (anti-fibrotic) → 3-6 months?

Summary

- Multiple pathways are involved in NASH pathogenesis, thus, combination therapy will likely have a role
- Many combinations are in develop for NASH and metabolic disease
    - Need to focus on rational mechanisms rather than convenience
- Understanding of NASH pathogenesis is fairly robust- despite there being details to work out, over all, a lot is known. As a field, the conversation should not be 'NASH is a disease no one really understands', because it undersells the knowledge that is in fact known, and leads decision makers who allocate resources to think those involved in research do not have handle on their understanding of the disease.

Combination Approaches for Cirrhotic NASH

Presenter: Julia Wattacheril, Columbia University
Slides: https://bit.ly/2Ktbrl7

Cirrhotic NAFLD

- Often think of fibrosis stage 4 as one category of advanced cirrhosis; however, there are stages of cirrhosis, and some are reversible.
- Within stage 4, there are 4a, 4b, 4c
    - 4a: fibrous septa, a lot of parenchyma remaining, significant amount of substrate
    - 4b: diminishing levels of intrahepatic fluid, thickened fibrous septa
    - 4c: more durable fibrotic response
- Stages correlate with outcomes, with 4A and B, which are thought to be reversible, correlating with better clinical outcomes.
    - 4c is the population where significant mortality is seen from end stage complications including decompensation, hepatocellular carcinoma, liver related death, and liver related events.
- Compensated Cirrhosis
    - Minimal portal hypertension
        - HVPG: 5-10mmHg
        - Very low risk of decompensation
        - Less liver fibrosis
        - Increased intrahepatic resistance
        - Treatment of underlying mechanism may prevent CSPH
    - Clinically significant portal hypertension (CSPH)
        - HVPG: >10mmHg
            - Appearance of varices

        - 4 times higher risk of decompensation
        - More fibrosis
        - Increased splanchnic blood flow
        - Treatment may decrease time to decompensation, but risk still exists

Targeting Disease Activity and Fibrosis

- Targeting metabolic substrate, NASH, and fibrosis will all depend on what that liver is able to tolerate from a safety perspective. And also how much substrate, remaining parenchyma, remaining lipid, there is to act on.
- Anti-NASH with anti-fibrotic potential: ASK1 inhibitor, PPARs, FXR
    - Potentially crossing both pathways, both mechanisms, with the same drug.
    - Combining with anti-fibrotic mechanism: pan caspase inhibitors, LOXL2 inhibitors, and Gal-3 inhibitors, depending on the subtype of cirrhosis that you're targeting.
- Example from Simtuzumab trial
    - Simtuzumab: monoclonal antibody directed at LOXL2
    - LOXL2: secreted, copper dependent amino oxidase, contributing to fibrogenesis by cross-linking collagen and elastin
    - In murine models, LOXL2 stabilizes fibrotic matrix
        - Inhibition shown to decrease liver fibrosis
    - Natural history study of 475 patients, with median follow-up >30 months[3]
        - Clinical events: ascites, encephalopathy, development of new varices, EVH, CPT ≥ 2-point increase, and/or MELD ≥15, death
        - Liver-related clinical events occurred in 19% (50/258) of patients with cirrhosis.
        - Correlation of ELF and HVPG - HVPG typically used in studies of patients with cirrhosis in order to stratify patients, to help better characterize the intrahepatic substrate and also safety.
- Example from Belapectin trial
    - Gal-3 is a lectin protein, binds galactose residues on glycoproteins
        - Increased in NASH, liver fibrosis, cirrhosis
        - Preclinical knockout models: resistant to development of NASH, fibrosis
    - GR-MD-02 = complex carbohydrate drug
        - Inhibits gal-3
        - Improves histopathology of NASH and reverses fibrosis in animal models
        - Phase 1 studies demonstrated safety, tolerability in NASH with advanced fibrosis, as well as reduction in portal pressure
    - Perfusion of the liver in PK studies show that there were varying concentrations at each dosage, demonstrating importance of measuring PK in the compensated cirrhotic patient population, where there's an alternative blood flow that's potentially increasing in terms of splanchnic blood flow, and overall cardiac compensation.
- Example from Emricasan trial
    - Caspase-mediated apoptosis has been observed with chronic liver disease (viral, metabolic)
    - Accumulation of apoptotic cells and subcellular fragments like microvesicles contain biologically active particles
    - Caspase cleaves cytokeratin-18 (CK-18)
        - Cleaved CK18 (cCK18) is a biomarker associated with inflammation in different etiologies of chronic liver disease (HCV, NAFLD, NASH)
    - Inhibition of caspase activity may decrease apoptosis and associated microparticles

[3] Sanyal AJ, et al. The natural history of advanced fibrosis due to nonalcoholic steatohepatitis: Data from the Simtuzumab trials. Hepatology. 2019;70(6):1913-27.

    - Population was patients with NASH cirrhosis and baseline HVPG ≥12 mmHg – high risk population
        - Primary endpoint was HVPG at 24 weeks

Safety in Advanced Liver Disease

- ACC inhibition- increased circulating triglycerides (seen with ACC inhibition)
    - Combined with FXR agonist, with effects on LDL, concern for increased atherogenesis
- Vitamin E + Immunomodulators – unknown MOA with an antioxidant in addition to anti-fibrotic pathway target – concern for extrahepatic carcinogenesis
- Malignancy concerns with more advanced fibrosis
- Potentiation of off-target effects

Considerations for Endpoints in Compensated Cirrhosis

- Not the regulatory endpoints included in recent FDA draft guidance
- Proof of concept
    - Improvement in disease activity (NAS)
    - No worsening of fibrosis
    - No worsening of HVPG
- Meaningful Benefit (longer term)
    - Reversal of fibrosis
        - >1 stage improvement
    - Time to progression to CSPH
    - Time to liver-related clinical events

DAY 2, SESSION II: REGULATORY CONSIDERATIONS

NASH Regulatory Update: EMA

Presenter: Elmer Schabel, BfArM/ European Medicines Agency
Slides: https://bit.ly/3rp3LAQ

EMA Reflection Paper

- Development strategy with “surrogate” endpoints at intermediate time-points and confirmatory approach post-licensing possible due to unmet medical need; placebo-control recommended.
- Patient population: Either non-cirrhotic (fibrosis stage 2 and 3) or cirrhotic population; NASH diagnosis by histology (activity), or (cirrhotics) appropriate support of NASH with other factors
- Non-cirrhotic population endpoint: Co-primary of histological evaluations of biopsies: “resolution of NASH without worsening of fibrosis” and “improvement of fibrosis and no worsening of NASH”.
    - Needs to be confirmed by hard endpoint data.
- Endpoints based on MoA: Possible, but if both co-primary EPs cannot be targeted, additional support needs to be presented (e.g. 2-stage fibrosis improvement in anti-fibrotics).
- Cirrhotic population endpoint: reversal of cirrhosis. Potentially needs additional support and long-term outcome data. Other endpoints possible (e.g. MELD score improvement; HVPG improvement).
- Cirrhotic population with previous decompensation: Hard outcomes recommended (Decompensation events, LTx, death).
- Fixed duration of trials mentioned (although flexibility allowed)
- Combination treatment: Suitable for non-responder populations and those with high risk

Stakeholder Meeting, December 2018

- Main points of discussion: (presented at previous Liver Forum 9)
  - The request for co-primary endpoints is a too demanding requirement and one of the components may be important enough from the patient's perspective
    - The request for co-primary endpoints is not adequate for certain MoA
    - The disharmonization of FDA and EMA requirements should be avoided
    - May unnecessarily prolong trial duration, and may take false conclusions
  - Requirements for combination therapy too strict
  - Need for development of PRO tools and inclusion of symptoms into trials not sufficiently addressed
  - Extent of CV safety documentation not clear
  - Pediatric issues: presence of genetic factors, differences to adult disease, need for different endpoints

Comments on Reflection Paper

- Total number of comments received: 19
- Total number of comments with regard to NASH parts: 15
- Stakeholder classification with comments on NASH:
  - 9 Industry (single company)
  - 1 Industry (association)
  - 2 Scientific organization/Learned society
  - 1 EU National regulatory agency
  - 1 Patient Advocacy Group/Organization
  - 1 Multi-stakeholder Organization
- Comment Areas:
  - General:
    - Separate into 2 (or even 3) guidance documents (1 comment)
    - Separate the three disease entities more clearly (2 comments)
    - Separate chapters for non-cirrhotic and cirrhotic populations (1 comment)
    - Separate non-compensated and compensated cirrhotic chapters more clearly (1 comment)
    - Clarify the terms "early and late clinical trials" (2 comments)
    - Should be mentioned that agents regarded to act on the cause of the disease are more suitable than anti-fibrotics (1 comment)
    - Medical need in NASH should be discussed, also more from a CV perspective (1 comment)
    - Classification of fibrosis stages be adapted to AASLD classification (1 comment)
    - Estimand chapters should be more precise (1 comment)
    - Too much emphasis on the need for hard endpoints (1 comment)
    - Non-invasive diagnosis methods favored/should be more promoted (1 comment)
  - Disease characterization:
    - Oversimplification of pathophysiology should be avoided (1 comment)
  - Inclusion criteria:
    - Consider concomitant medication (and add-on medication during the trial) potentially influencing disease outcome (4 comments)
    - NASH diagnosis requirements based on NAS-requirements (NAS>5 and NAS>4 with additional requirements) too strict (3 comments)
    - Requirement for previously failed dietary treatment should be deleted (1 comment) or clarified/modified (2 comments)
    - Cirrhosis diagnosis should not be histological (but clinical) (2 comments)

        - Criteria for presence of features of the metabolic syndrome should be inclusion criteria (2 comments)
        - Simplify inclusion criteria and separate cirrhotic and non-cirrhotic (1 comment)
        - NASH-cirrhosis diagnosis should be more flexible (1 comment)
        - Severity of cirrhosis should be classified according to established criteria (e.g. Child Pugh) (1comment)
    - Trial design/Endpoints:
        - The request for co-primary evaluation of the two composites too strict (11 comments)
        - The request for 2-stage improvement of fibrosis in anti-fibrotics is too strict (6 comments)
        - Symptoms (PROs) and evaluation of QoL should be included (6 comments)
        - The study duration should be given as “flexible” only (5 comments)
        - Combination therapy requirements too strict (not only 2nd line and “at risk populations”) (5 comments)
        - Requirement for histology in “early” trials should be deleted (4 comments)
        - Include “manifestation of T2DM” (1 comment) and CV MACE (3-component) events in the “hard endpoints” (2 comments)
        - Replace requirement for “MELD>14” with “MELD>15” (2 comments)
        - Reversal of cirrhosis should be classified as “hard endpoint” (1 comment) or be recognized intermediate endpoint without restrictions (3 comments)
        - Manifestation of cirrhosis should be defined as “hard endpoint” (1 comment)
        - Mention the use of special design features (e.g. adaptive design, extrapolation of placebo control) for development of combination treatments (1 comment)
        - Suitability of MELD for a patient population with CV disease be checked (1 comment)
        - Allow more flexibility with regard to definition of “resolution of NASH” especially with regard to the ballooning criteria (2 comments) or the steatosis criterion (1 comment)
    - Safety:
        - Clarify/State that CV outcome trials are not required for documentation of safety (3 comments)
        - It should be mentioned that treatments have no detrimental effect on other aspects of the metabolic syndrome (T2DM, obesity, serum lipids)
    - Children:
        - Age cut-off at 10 years proposed (1 comment)
        - Biomarkers should be primary endpoints (1 comment)
        - Histology as endpoint should not be mandatory (1 comment)
        - Histology as endpoint may be needed (1 comment)
        - Different histological features to be taken account of, different scoring system likely be needed (1 comment)
        - Young children (age 6-10) may not be candidates for pharmacological treatment (1 comment)
        - Studies in children (<12 years) should be deferred until more natural history data are available (1 comment)

NASH Regulatory Update: FDA

Presenter: Yao-Yao Zhu, U.S. Food and Drug Administration
Slides: https://bit.ly/2WFZtHq

NAFLD/ NASH Submission Trends

- 2012: <5 submissions for NAFLD/NASH
- 2019: >40 submissions for NAFLD/NASH
- Development Programs
    - Mostly commercial INDs
    - Few research INDs
    - Most in phase 1 and 2; some phase 3
        - Few phase 3 completions
- Investigational Treatment
    - Repurposing or review of previously approved/studied agents
        - T2DM agents, anti-hyperlipidemia, weight loss
        - Many comorbidities resulting in confounders for safety
            - May need to reassess benefit and risk when considering old drugs for new indications.
    - Combination therapy
- Failed Phase 3 Trials
    - Variability of histological readings
    - Adequacy of the surrogates
    - Non-invasive biomarkers
- Potential efficacy endpoints

NASH Regulatory Guidance Documents

- Noncirrhotic NASH with Liver Fibrosis: Developing Drugs for Treatment. December 2018[4]
- NASH with Compensated Cirrhosis: Developing Drugs for Treatment. June 2019[5]
    - Reviews phase 2 and 3 programs
    - Includes eligibility criteria, study design, efficacy endpoints & safety monitoring
- Comments: 9 responses submitted by industry, academic institutions and organizations
    - Primarily focused on efficacy endpoints, surrogates or clinical benefit, and eligibility criteria
- Challenges with Accelerated Approval
    - Phase 4 trials are required to verify and describe the clinical benefit of a drug; however, there are serious challenges in completion and obtaining the necessary efficacy data.
        - Difficult enrollment and retention once the product is approved for market
        - Difficult to keep a prolonged blinded study for many years
    - Continue to recommend double blind placebo controlled trials; however, it may be possible to discuss historical control.
        - High quality, detailed natural history studies are essential and should be started early in drug development

Study Populations

- Sub-populations for compensated NASH cirrhosis population
    - Early cirrhosis without CSPH (only mildly elevated HVPG)
    - Cirrhosis with CSPH (varices, thrombocytopenia)
- Classification may provide opportunity to enrich trials for more advanced disease, and therefore may shorten the trial period to achieve decompensation endpoint
- Clinical benefit endpoint: development of varices requiring treatment, in patients without varices at baseline
    - Would need to be based on appropriate definitions and agreed methods of detection

[4] https://www.fda.gov/media/119044/download
[5] https://www.fda.gov/media/127738/download

- Early cirrhosis without CSPH
  - No varices
  - Need to define cut-offs for HVPG, platelet count, INR, TB, albumin
- Compensated cirrhosis with CSPH
  - Clinical based definitions
    - Presence of varices
    - HVPG based on selected thresholds
    - Platelet count based on selected thresholds
    - Albumin
  - Child-Pugh-Turcotte
    - TB<2
    - INR<1.7
  - DILIN
    - TB 2
    - INR 1.5
- Composite clinical endpoints:
  - Current: death, liver transplant, decompensation events (varices bleeding, HE, ascites), MELD >15 in patients with MELD<12 at baseline
  - New: development of varices requiring treatment? (banding or pharmacological)
  - Prospective statistical planning for single component drivers of composite endpoints
    - Need to ensure all aspects of the disease will be positively impacted by the drug

Baseline Assessments

- Are baseline assessments needed to measure efficacy of an endpoint?
  - Is a baseline biopsy necessary?
- Generally, it is possible to assess treatment difference between randomized groups on an endpoint without baseline measurements
  - It is not possible to assess and compare improvement in biopsy-based outcome measures
- Lack of baseline measurement may increase uncertainty regarding enrolled population
  - Current NASH/ liver fibrosis biomarkers are not accurate in identifying/differentiating non-cirrhotic NASH fibrosis stages 2 or 3, or early cirrhosis
  - Variability in liver biopsy
  - May require large sample size to detect treatment effect

Endpoints and Biomarkers

- Weight loss as a potential surrogate?
- ALT, ELF (Enhanced Liver Fibrosis), TE (transient elastography), and other Biomarkers
- Pediatric population considerations
  - Progression to diabetes (may be challenging to dissect the relationship to NASH given the prolonged delay to NASH outcomes and complex physiology interplay)

DAY 2, SESSION III: PARALLEL BREAKOUT SESSIONS

NASH Cirrhosis Working Group

COMPENSATED NASH, RISK STRATIFICATION

Presenter: Mazen Noureddin, Cedars-Sinai Medical Center
Slides: https://bit.ly/3amxILY

Review of Comorbidities

- NAFLD→ NASH→ Cirrhosis→ CSPH→ decompensation events
  - Should progression of cirrhosis be considered differently for NASH?
  - NASH is a multi-system disease, increasing evidence that NAFLD can lead to cardiovascular and kidney events (cardiac dysfunction, congestive heart failure, cardiac arrhythmias, chronic kidney disease)
- Meta-analysis [6] looking at the prevalence of comorbidities with NASH
  - 82% have obesity
  - 44% T2DM
  - 72% hyperlipidemia/ dyslipidemia
  - 68% hypertension
  - 71% metabolic syndrome
    - Implications for NASH clinical trials, as patients have high prevalence of these comorbidities and they have an impact on the disease, but only really focusing on the liver and only partially paying attention to the other organs.
- Cardiovascular disease
  - There is a good amount of literature suggesting that NAFLD is associated with cardiac dysfunctions – including endothelial dysfunction, increased carotid artery intima thickness, stiffness of the arteries, coronary artery disease, aortic valve sclerosis, arrhythmias, diastolic dysfunction
  - High association between NAFLD and diastolic dysfunction
  - Relaxation of the left ventricle correlated with the NAS score – there is likely an association between the two, but still evolving whether NAFLD is a cause of this
  - Meta-analysis[7] of many patients demonstrating that NAFLD is associated with fatal and non-fatal cardiovascular events – increased odds ratio, and risk appeared to increase with greater severity of NAFLD.
  - Angulo et al.[8] demonstrated that CVD is the most common cause of death/ liver transplantation in NAFLD/ NASH—one of the first studies to show that patients died from cardiovascular events (38%), non-liver cancers (19%), other (18%), cirrhosis complications (8%) and infections (8%)
- Kidney Disease
  - Some evidence suggesting that presence and severity of NAFLD is associated with an increased prevalence of CKD - need more evidence.
  - 20-55% NAFLD patients have CKD, compared to 5-30% of the general population
  - Presence and severity of NAFLD predicts the development of incident CKD independent of traditional cardiorenal risk factors
  - Meta-analysis[9] of 96,000+ patients, 34% had NAFLD, and 4,653 had CKD stage ≥3
  - Patients with NAFLD had a significantly higher risk of incident CKD than those without NAFLD (increase of 1.37)
  - Patients with more 'severe' NAFLD according to ultrasound and non-invasive fibrosis markers were more likely to develop incident CKD (increase of 1.50)
  - The leading cause for simultaneous liver-kidney transplantation is NASH
  - In the U.S., over 10% of adult population (and more than 25% in 65 years and older) have CKD

[6] Younossi ZM, et al. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology. 2016;64(1):73-84. https://doi.org/10.1002/hep.28431
[7] Targher G, et al. Non-alcoholic fatty liver disease and risk of incident cardiovascular disease: A meta-analysis. Journal of Hepatology. 2016;65(3):589-600. https://doi.org/10.1016/j.jhep.2016.05.013
[8] Angulo 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(2):389-97.e10. https://doi.org/10.1053/j.gastro.2015.04.043
[9] Mantovani A, et al. Nonalcoholic fatty liver disease increases risk of incident chronic kidney disease: A systematic review and meta-analysis. Metabolism. 2018;79:64-76. https://doi.org/10.1016/j.metabol.2017.11.003

    - NAFLD and CKD share risk factors
    - Hepatorenal syndrome can develop in patients with cirrhosis with portal hypertension
- Cancer
    - Paper from Kim et al.[10] which followed 25,000 patients over 7 years found that there were increased malignancies (other than HCC) – including colorectal cancer and breast cancer.
    - Article in press by Allen et al.[11] from the Mayo Clinic Group which followed 14,000 patients, of which 4,700 had NAFLD, and found that there is increased risk of malignancy liver cancer, uterine cancer, stomach cancer, pancreatic cancer, and colon cancer.
    - Article found that when controlled for, obesity did not have an increased risk of incident cancer a little controversial, should be replicated
- Other Risk Factors
    - NASH is a whole-system problem, and beyond the 'disease', patients are very often fatigued and depressed
    - Diabetes, polycystic ovary syndrome, sleep apnea, hypopituitarism, hypogonadism, psoriasis
    - Not really being considered in current NASH clinical trials

Considerations for Clinical Trials

- Kidney Disease
    - Patients enter trials with proteinuria, CKD, or end-stage renal failure
        - Medications can affect glomerular filtration rate (GFR)
        - Creatine is used in the GFR formula, but it is not usually measured accurately in obese patients, and this is even worse in patients with cirrhosis, and more so in decompensated cirrhosis.
        - For example when patients develop ascites and need to be treated with diuretics
- Cardiovascular Disease
    - Patients enter trials with endothelial dysfunction, increased arterial stiffness and elevated coronary calcium scores, diastolic dysfunction, history of CVD or MI, cirrhosis cardiomyopathy
    - Most trials do not assess cardiac function or look at stiffness or diastolic dysfunction
    - Some medications may worsen lipid profile (not necessarily a problem but need to understand this better)
    - Some trials exclude patients with a history of CVD/ MI (many NASH patients fall into this category, and would likely use a drug when approved)

Examples of Risk Stratification from other Systemic Diseases

- Obesity
    - Kings Criteria: airway, BMI, CV risk, diabetes, economic complications, functional limitation, gonadal dysfunction, health status, image
        - Each criteria put into a stage 0-3
        - Commonalities to NAFLD—add kidney and malignancies, and sarcopenia

[10] Kim G, et al. Association between non-alcoholic fatty liver disease and cancer incidence rate. Journal of Hepatology. 2018;68(1):140-6. https://doi.org/10.1016/j.jhep.2017.09.012
[11] Allen AM, et al. The risk of incident extrahepatic cancers is higher in non-alcoholic fatty liver disease than obesity - A longitudinal cohort study. Journal of Hepatology. In Press. https://doi.org/10.1016/j.jhep.2019.08.018

    - Metabolomic study by Cirulli et al.[12] stratifying patients with different types of metabolome, demonstrating increased insulin resistance and increased cardiac events in those with obese metabolome compared to those with healthy metabolome
- Diabetes
    - Study by Ahlqvist et al[13] looking at type 2 diabetes long-term complications, especially with kidney and heart. Stratified by 6 variables: glutamate decarboxylase antibodies, age at diagnosis, BMI, HbA1C, homoeostatic model assessment of B-cell function, insulin resistance.
    - Insulin resistance cluster had more chronic kidney disease, but had the same medications as other clusters – is this fair to be lumped together with the other T2DM patients?
    - Risk stratification based on complications – the data demonstrated that clusters had different courses in terms of time to develop complications such as end-stage renal disease, CVD, CKD

DECOMPENSATED NASH CIRRHOSIS, RISK STRATIFICATION

Presenter: Jasmohan Bajaj, Virginia Commonwealth University
Slides: https://bit.ly/3p1BVsB

Defining Decompensation Events

- We are defining decompensation by 3 complications: ascites, variceal hemorrhage (VH), and hepatic encephalopathy (HE).
    - These complications occur frequently and often together.
    - Other complications such as hepatorenal syndrome and portal hypertension may occur, but are not as frequent
- Ascites
    - Case definition of ascites:
        - Clinically overt based on physical examination
        - Free fluid in abdomen on imaging (ultrasound, CT, MRI, etc.)
    - Considerations and recommendations for clinical trials:
        - Prior to initiating a NASH cirrhosis clinical trial, it is recommended to obtain a baseline ultrasound to assess the presence of ascites.
        - For decompensated trials, consider a “treatment requirement” with diuretics because it may strengthen the certainty of ascites and decompensation.
        - Hepatic hydrothorax could also be considered in the absence of ascites and after exclusion of other causes of pleural effusion as an “ascites equivalent”
    - Grey zone (no specific guidance) and recommendations:
        - Perihepatic ascites only on imaging
        - Previous episode of transient ascites related to a precipitant (excess salt, VH, etc.) now resolved
        - It is also recommended to exclude these patients from phase 2 trials.
            - However, it may be beneficial to include a subpopulation of these patients in phase 3 studies.
            - Such patients should be analyzed separately, and their enrollment should be designed with the regulatory authorities at the planning stages.

[12] Cirulli ET, et al. Profound Perturbation of the Metabolome in Obesity Is Associated with Health Risk. Cell Metabolism. 2019;29(2):488-500. https://doi.org/10.1016/j.cmet.2018.09.022
[13] Ahlqvist E, et al. Clusters provide a better holistic view of type 2 diabetes than simple clinical features. The Lancet Diabetes & Endocrinology. 2019;7(9):668-9. https://doi.org/10.1016/S2213-8587(19)30257-8

        - It is very important to clearly define ascites in all forms to ensure there are no doubts at the end of a trial regarding the population.
    - Patients with normal synthetic function and no asicites have a long time course to clinical events- enriching the patient population using presence of perihepatic ascites or previous transient ascites could increase the likelihood of outcomes with a 1-2 year timeframe.
    - Re-compensation: refers to when a patient has experienced a decompensation event but has since been stable for a long period of time. Though stable, this patient still has different characteristics, different underlying situation than a patient that has never decompensated.
        - This has become more frequent with the advent of successful Hepatitis C treatment.
    - The diuretics used to treat ascites are spironolactone equivalence or furosemide equivalence in the normal dosage level. This treatment is used to define whether someone who has cirrhosis related ascites will experience decompensation
        - Others, like hydrochlorothiazide are not usually used for cirrhosis

- Variceal hemorrhage
    - Case Definition:
        - Upper GI hemorrhage that required hospitalization and on endoscopy showed any of the following:
            - Varix spurting blood
            - Varix with overlying clot or white nipple
            - Only varices and no other lesion
    - Considerations and recommendations for clinical trials:
        - Acute (not chronic) bleeding from portal hypertensive gastropathy that required hospitalization may be considered a “VH equivalent”
        - Recommend waiting for a period of 3 months or more for stability prior to enrollment in NASH decompensated clinical trial
    - Grey zone and recommendations
        - Previous (>1-2 years) episode of documented VH that required hospitalization and has not developed re-bleeding
            - Could be still on a stable dose of NSBB
        - Chronic bleeding from portal hypertensive gastropathy
        - Recommended inclusion of a subpopulation of TIPS in phase 3 trials. It may be an option upon discussion with the regulatory authorities, and depends on the outcome of the trial, mechanism of drug action and duration since TIPS.
- Hepatic encephalopathy (HE)
    - Case Definition:
        - Overt (≥grade 2) HE per AASLD/EASL guidelines
    - Considerations and recommendations for clinical trials
        - Consider “requiring treatment” as evidence of chronic decompensation
        - Consider “requiring hospitalization” as stronger evidence of definitive HE
        - Recommend investigator perform a thorough chart review to investigate the initial diagnosis, although this is often missed and is essential for the diagnosis.
    - Grey zone and recommendations
        - Previous transient episode of overt HE related to a precipitant (infection, VH, metabolic, etc.) now resolved, not requiring treatment
        - Covert HE (minimal or grade 1) with no prior history of overt HE currently on treatment.
            - May be excluded from phase 2 studies and limited in phase 3.
        - HE occurring primarily due to porto-systemic cause (e.g., occluded shunt)

- Notes:
    - Gray Zone: Not everyone is in ascites, bleeding, or has hepatic encephalopathy all the time. The grey zone refers to the people who have had these conditions a long time ago and have completely resolved or have had such significant control that they were able to discontinue therapy or therapy has kept their condition under control.
    - Specific trials may include or exclude patients fulfilling grey zone criteria from compensated trials but these patients should be analyzed as a separate subgroup analyses or stratified at randomization, especially if they are a large component of the total population.

Decompensated Cirrhosis: Stratification for Clinical Trials

- Early Decompensation
    - Population:
        - Patients with a history or presence of a single decompensating event (ascites, VH, or HE), but are well controlled on a specific therapy
        - Consider (or not) patients in the grey zone
    - Stratification:
        - Grey zone or not
        - CP A vs early B
        - Type of decompensation event (i.e., ascites vs other)
        - Other comorbidities (CKD, CHF, CAD, etc)
        - MELD (lower vs higher)
    - Primary endpoint:
        - Second decompensation event, further decompensation or death
    - Outcomes:
        - Clinical (primary)
            - Development of a second type of decompensation event
            - Further decompensation (refractory ascites or refractory HE, SBP, HRS)
            - Critical illness requiring hospitalization
            - Death (all-cause mortality)
            - Recompensation?
        - Surrogate (candidate)
            - Progression in MELD
        - Exploratory:
            - Improvement in functional test
            - Additional emerging biomarkers
        - Safety;
            - HCC (not an outcome but confounding variable – consider as competing event unless trial specifically for HCC
- Advanced Decompensation
    - Population:
        - Patients with history or presence of two or more decompensating events (ascites, VH, HE)
    - Stratification:
        - Grey zone or not
        - CP B vs CP C
        - Other comorbidities (CKD, CHF, CAD)
        - MELD (lower vs higher)
        - Presence/absence of ascites
    - Primary endpoint:
        - Death

- Outcomes:
  - Clinical (primary)
    - Death (all-cause mortality)
  - Clinical (secondary)
    - Further decompensation (refractory ascites or refractory HE, SBP, HRS)
    - Critical illness requiring hospitalization
  - Surrogate (candidate)
    - Progression in MELD
  - Safety:
    - HCC

Pediatric Issues in NASH Working Group

DEFINING RESPONSE IN PEDIATRIC NAFLD – SURROGATE BIOMARKERS

Presenter: Miriam Vos, Emory University

Slides: https://bit.ly/34oOfv3

- Definitions
  - Surrogate endpoint: a clinical trial endpoint used as a substitute for a direct measure of how a patient feels, functions, or survives. A surrogate endpoint does not measure the clinical benefit of primary interest in and of itself, but rather is expected to predict that clinical benefit.
    - FDA Approved Surrogate Endpoints
    - Example of FDA Approved Surrogate Endpoint for Lipodystrophy: combination of serum HbA1c, fasting glucose, and triglycerides.
      - Interesting that the surrogate is a combination of features, and is a systematic approach. This could be applicable to NAFLD.
  - Biomarker: 1.) A defined characteristic that is measured as an indicator of normal or pathogenic biological processes or 2.) a response to an intervention
    - Response Biomarker: a biomarker used to show that a biological response has occurred in an individual who has been exposed to a medical product or an environmental agent.
    - Response can be non-progression, or it can be reversal
- Critical Questions
  - Will biomarkers approved/validated/qualified for adults with NASH be applicable to children?
  - What are the most important needs for pediatric biomarkers? Diagnostic? Response?
- Current Recommendations
  - Developed by the Liver Forum Pediatric Working Group and published in the December 2019 issue of Gastroenterology[14].
    - For early phase studies, reduction of elevated serum ALT is a reasonable primary outcome
    - While steatosis can be measured accurately with MRI, there's inadequate data to support that steatosis reduction will lead to clinically meaningful benefit.

[14] Vos MB, et al. Factors to consider in development of drugs for pediatric nonalcoholic fatty liver disease. Gastroenterology. 2019;157(6):1448-56. https://doi.org/10.1053/j.gastro.2019.08.048

        - Endpoints "reasonably likely to predict clinical outcomes" by the regulatory authorities for adults are as follows, and pediatric trials may use similar endpoints in those with NASH:
            - FDA: Biopsy based resolution of steatohepatitis and no worsening of fibrosis OR at least one-point improvement in fibrosis with no worsening of steatosis, ballooning or inflammation.
            - EMA: Biopsy based resolution of steatohepatitis and no worsening of fibrosis AND at least one-point improvement in fibrosis with no worsening of steatosis, ballooning or inflammation.
                - This may be an area that needs additional questioning and thinking- pediatric pathology is very different than adults; therefore, having the same histologic outcome definitions for pediatric and adult patients may not be reasonable.
- ALT
    - Secondary analysis from the TONIC trial (placebo, Metformin, Vitamin E) regrouped by improved histology, stable histology, and worsened histology. Analyzed by mean ALT compared with NASH and mean ALT compared with fibrosis
        - NASH- patients with improved NASH had a dramatic drop in mean ALT compared with patients who remained stable or who progressed
        - Fibrosis- patients with improved fibrosis also had a decrease in mean ALT compared with patients who remained stable or who progressed.
        - This was also seen when looking at mean ALT % change from baseline.
            - Stronger data than seen in adults showing that ALT reflects histology change
            - Mean ALT at baseline is also higher than seen in adults (120-140 U/L)
    - Analysis of CYNCH trial grouped by responders and non-responders (response defined as decrease in NAS of ≥ 2 at 52 weeks)
        - Change in ALT, AST, and GGT were all significantly different between responders and non-responders.
        - May already have surrogate marker(s) that are specific to children – need to carefully review and comment on pediatric data and the implications for differences in trial designs compared with adults.
- GGT
    - 2006 study[15] looking at predictors of NAFLD and obese children – strongest predictors of disease were GGT and ALT.
    - Need more longitudinal data that can model GGT change with histology change.
- Hepatic Fat
    - Fat can be very high in children
    - Comparison of MRI PDFF measurements of liver fat with histological steatosis grade shows that MRI can accurately quantify fat and match up with the histological grade of steatosis in pediatric patients[16].
        - Change in fat by MRI PDFF predicts change in histological findings (improve, stable, worsen)

[15] Sartorio A, et al. Predictors of non-alcoholic fatty liver disease in obese children. European Journal of Clinical Nutrition. 2007;61:877-83. https://doi.org/10.1038/sj.ejcn.1602588
[16] Middleton MS, et al. Diagnostic accuracy of magnetic resonance imaging hepatic proton density fat fraction in pediatric nonalcoholic fatty liver disease. Hepatology. 2018;67(3):858-72. https://doi.org/10.1002/hep.29596

        - However, change in fat by MRI PDFF was not related to changes in lobular or portal inflammation scores, hepatocellular ballooning scores, or fibrosis scores.
    - 8-week study[17] of low free sugar diet vs. standard of care; meals were provided to family
        - Primary outcome was liver fat- the mean decrease in hepatic steatosis from baseline to week 8 was significantly greater for the intervention diet group compared with the control group.
        - Also looked at ALT, AST, and GGT: reduction in fat was strongly associated with a reduction in ALT and GGT
- Fibrosis
    - Insufficient longitudinal data correlated with histology exists at this time
- Histology
    - Response in histology is currently based on NAS score or NASH; however, the lack of ballooning in pediatric patients is an issue and it is unclear if the current NAS score captures pediatrics adequately enough.
    - Need studies that compare histology to pediatric clinical status; long term natural history studies with baseline surrogates and clinical outcomes
- Phenotypes
    - What is the relationship of progression to these phenotypes?
        - Prepubertal, pubertal and post pubertal (adult)
        - Insulin resistant, prediabetic, diabetic
        - Dyslipidemic, normolipidemic
        - Lean, overweight, obese
        - Low ALT, mid-range and very high (>250)
        - No fibrosis, early fibrosis, advanced fibrosis

TYPE 2 DIABETES AS CLINICAL ENDPOINT IN PEDIATRIC NASH

Presenter: Stavra Xanthakos, Cincinnati Children's Hospital
Slides: https://bit.ly/3alQTFy

- Overlap between type 2 diabetes and NAFLD in youth
    - In the U.S. The proportion of diabetes in adolescents that is type 2 has increased from 4% in 1994, to 33% in 2008-2009.
    - Between 2001 and 2009, there were significant increases in type 2 diabetes prevalence in the U.S. for both sexes, across all ages 10-19, in non-Hispanic white, Hispanic, and African American youth[18].
        - African American, Hispanic, Asian Pacific Islander, and American Indian youth have a greater proportion of type 2 diabetes vs. type 1 diabetes, compared with non-Hispanic white youth.
        - African American youth have a high risk of type 2 diabetes, but a low risk of NAFLD.
    - Between 2003 and 2013 there has been a global increase in type 2 diabetes in youth and young adults, driven by increases in Southeast Asia and Western Pacific.

[17] Schwimmer JB, et al. Effect of a low free sugar diet vs usual diet on nonalcoholic fatty liver disease in adolescent boys: A randomized clinical trial. JAMA. 2019;321(3):256-65. https://doi.org/10.1001/jama.2018.20579
[18] Dabelea D, et al. Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. JAMA. 2014;311(17):1778-86. https://doi.org/10.1001/jama.2014.3201

  - In study[519] of children with type 2 diabetes, of the patients that had liver enzymes available (42%), 45% had ALT > ULN for lab reference range (40 IU/L at the time of the study)
  - Despite normal ALT, adult patients with type 2 diabetes can have NAFLD – 50% prevalence of NAFLD observed in 2015 study[20].
  - Study[21] looking at severity of liver disease on histology found that having prediabetes or type 2 diabetes increased risk of having NASH.
    - Those with type 2 diabetes had 3 times higher risk for definite NASH
  - Amongst adolescents undergoing bariatric surgery[22], 20% had NASH, 39% had NAFLD, and 41% had no NAFLD
    - The most significant predictors of having more severe liver disease was ALT elevation.
    - Fasting glucose elevation was also associated with higher odds of having more severe liver disease.
    - Diabetes and ALT were the only risk factors associated with having fibrosis.
  - Over a 1-2 year period, children in placebo arms receiving standard lifestyle counseling, ½ had improved NASH or fibrosis, while ⅓ had worsened NASH or fibrosis.
    - Disease progression was significantly related to worsening of HbA1C.
    - Preliminary data showing type 2 diabetes developed in 8%
- Differences between youth vs. adult-onset T2D
  - Those with adolescent onset type 2 diabetes have significantly lower insulin sensitivity than adults, as well as much higher insulin secretion rates.
  - Adolescents tend to have very rapid progression, comorbidities of hypertension, dyslipidemia and microalbuminuria over three to four years of follow-up.
  - Adolescents have significant increases in carotid thickness over relatively short period of follow-up.
  - By 15-20 years[23] after diagnosis, half of patients with youth-onset type 2 diabetes have developed clinically significant nephropathy, neuropathy, retinopathy, or other major complications.
  - Despite being on treatment (Metformin; Metformin + Lifestyle; Metformin + Rosiglitazone) youth with type 2 diabetes had declines in insulin sensitivity and beta cell function.
- Relevance to pediatric clinical trials in NASH and T2DM
  - How will having both youth-onset NASH and type 2 diabetes affect responses to treatments?
    - Most type 2 diabetes trials exclude patients with ALT > 2.5 to 3x ULN, and there is a lack of liver imaging in these trials.

---

[19] Nadeau KJ, et al/ Type 2 Diabetes in children is frequently associated with elevated alanine aminotransferase. Journal of Pediatric Gastroenterology and Nutrition. 2005;41(1):94-8. https://doi.org/10.1097/01.mpg.0000164698.03164.e5

[20] Portillo-Sanchez P, et al. High prevalence of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus and normal plasma aminotransferase levels. The Journal of Clinical Endocrinology & Metabolism. 2015;100(6):2231-6. https://doi.org/10.1210/jc.2015-1966

[21] Newton KP, et al. Prevalence of prediabetes and type 2 diabetes in children with nonalcoholic fatty liver disease. JAMA Pediatrics. 2016;170(10):e161971. https://www.doi.org/10.1001/jamapediatrics.2016.1971

[22] Xanthakos SA, et al. High prevalence of nonalcoholic fatty liver disease in adolescents undergoing bariatric surgery. Gastroenterology. 2015;149(3):623-34. https://doi.org/10.1053/j.gastro.2015.05.039

[23] Dart AB, et al. Earlier onset of complications in youth with type 2 diabetes. Diabetes Care. 2014;37(2):436-43. https://doi.org/10.2337/dc13-0954

        - Earlier NASH trials excluded all patients with diabetes, and still poorly controlled type 2 diabetes is an exclusion (ex. A1C >9%)
    - Due to the overlap between NAFLD and type 2 diabetes, it will be important to develop treatments that are of mutual benefit to those that are affected by both diseases.
    - Correctly classifying type 2 diabetes at trial entry and follow-up
        - Capture duration, medications, etc.
    - Additional cross-talk is needed in pediatric NASH and type 2 diabetes trials.

Biomarker Development: Diagnostics

Presenter: John Sninsky, Independent Consultant
Slides: https://bit.ly/3oZmS2z

State of biomarker development

- The last five years has seen a shift in appreciation, development of guidelines, and transparency for the diagnostic community, for the required rigor and associated quality of evidence that diagnostic biomarkers need to have
- A paper by Ioannidis et al.[24] concluded that the current biomarker pipeline is too prone to failures and consideration of clinical needs should become a starting point
    - More stringent methodology is needed if these biomarkers are going to meaningfully contribute

Common Missteps in Diagnostic Studies

- Performance of test in Discovery set only (overfit test performance)
- Use 'normal' samples as comparator rather than differential diagnosis samples (exaggerated performance)
- Dissimilar Discovery, Validation and Clinical Use sets (inaccurate estimate of performance) or distribution of samples
- Mixture of Discovery and Validation sets (inaccurate estimate of performance, overfit; solely statistical cross-validation insufficient)
- Lack pre-specified clinical/statistical analysis plan (introduction of bias)
- Convenience or opportunistic samples (solely retrospective; not representative; inaccurate performance)
- Single center study rather than multi-center study (test robustness)
- Poorly validated analytical performance (inaccurate performance, robustness, transferability)
- Does not consider implications of pre-analytical variation of biomarker
- Samples tested with different versions of test (inaccurate performance)
- Small sample sets (likely bias and chance; lack generalizability)
- Provide clinical validity but not clinical utility (questionable reimbursement)
- Lacks attention to PPV or NPV for indication of test (actionability)
- Cost effectiveness not modeled (questionable reimbursement)
- Statistical analysis only includes ROC, or sensitivity and specificity (test performance but not patient performance)
- Lack actionable outcomes (what will clinician or patient do differently with information)
- Does not compare performance relative to single or combined routinely used tests or information (independence relative to presently used information)

Sea Change in Clinical Diagnostics

[24] Ioannidis JPA, et al. Waste, Leaks, and Failures in the Biomarker Pipeline. Clinical Chemistry. 2017;63(5):963-972. https://doi.org/10.1373/clinchem.2016.254649

- Diagnostics and overall understanding of the underlying nature of disease in NASH has changed substantially over the last 5 to 10 years
- A formal phased development of diagnostic tests, similar to drug development, has been adopted: analytical validation, clinical validation, and clinical utility
    - The cost effectiveness of the intervention is also evaluated
- High quality evidence needs to be provided by test service or test kit providers
- Clinical utility is now required for reimbursement instead of only clinical validity as in the past
- Evidence is now understood to be a continuum
    - It is important to ensure diagnostics are "fit for purpose" and the amount invested into a diagnostic or biomarker test must be relative to the value being returned

Biomarker Regulatory Oversight and Reimbursement

- Although diagnostics make up 3% of healthcare expenditures, they inform 65% of how spending is directed[25]
- There is a concern that medical insights are not being translated in a timely manner
- Translational, clinical development, and regulatory sciences are evolving at a rapid pace
- Accelerated translation of discoveries into practice of medicine requires a 'directed path' instead of an 'exploratory walk'
- High quality, evidence-supported 'clinical-grade' biomarker assays require substantial investment
- If clinical-grade assays are not value priced, innovation from government and private industry will be stifled

Biomarker Discovery and Biomarker Translation

- Discovery and translation are equally valuable
- Biomarker Discovery: the exploratory walk
    - Biomarker or biology-centric
    - Promises key insights into fundamental underlying pathophysiology
    - Plethora of biologically plausible biomarkers
    - Benefits from a deep understanding of biology
    - Correlations and group diagnostic metrics suffice
- Biomarker Translation for Clinical Practice (directed path)
    - Clinical question-centric
    - Promises improved patient management
    - Few biomarkers merit prioritization
    - Benefits from translation and diagnostic development path knowledge
    - Predictive values are most important for individual patients

Regulatory and Reimbursement Considerations

- Two Paths for Regulatory Oversight in U.S.
    - Clinical Laboratory Improvement Amendments of 1988 (CLIA)
        - Approves the laboratory with clinical validation
        - Regulated by CMS
    - FDA In-Vitro Diagnostics
        - Approves a commercial kit that is used by laboratories
        - Either 510k or PMA
    - The 510(k) Paradigm Continues to Evolve
        - Over the last decide, additional 510(k)s have been developed.
        - Types of 510(k)'s include traditional, special, abbreviated, and de novo

[25] Rohr U, et al. The Value of In Vitro Diagnostic Testing in Medical Practice: A Status Report. PLOS ONE. 2016;11(3). https://doi.org/10.1371/journal.pone.0149856

- Reimbursement for CMS
    - Requires the submission of a technical and clinical dossier
    - Depending on the quality of evidence and nature of the unmet need, reimbursement may be approved between 1 and 3 years
    - The path is informed by analytical validation, clinical validation, and clinical utility

Reproducibility and Utility of Biomarker Studies

- Why Most Published Research Biomarker Studies Are Not Reproducible- adapted from Ionnidis 2005[26]
    - Corollary 1: The smaller the studies conducted in a scientific field, the less likely the research findings are to be true (reproducible).
    - Corollary 2: The smaller the effect sizes in a scientific field, the less likely the research findings are to be true.
    - Corollary 3: The greater the flexibility in designs, definitions, outcomes, and analytical modes in a scientific field, the less likely the research findings are to be true.
    - Corollary 4: The assay does not address a clear unmet actionable diagnostic need.
    - Corollary 5: The study does not accurately reflect the eventual intended use population.
    - Corollary 6: The level of evidence is insufficient to be used in a clinical setting with confidence.
- Types of Reproducibility
    - Reproducibility of methods: the ability to understand or repeat as exactly as possible the experimental and computational procedures
    - Reproducibility of results: the ability to produce corroborating results in a new study, having followed the same experimental methods
    - Reproducibility of inferences: the making of knowledge claims of similar strength from a study replication

Studying Biomarkers

- Understand timeframes
    - Different biomarkers have value in distinct time frames
    - Important to understand biological variation of a biomarker
    - Biological variation may be due to temporary ‘homeostatic disruption’
    - Biomarkers for managing treatment are a compelling unmet need
    - Statistical tools vary across types of biomarkers
- Identify the right question
- Understand the needed evidence
- Commit to high quality studies
    - Intended use drives evidentiary studies
        - The assay and specimen must be described
        - The target population must be identified
        - The results, the nature of the results, the threshold, and the indication must be identified
- Three steps for diagnostic test development:
    - Analytical validation: how well the test predicts the presence or absence of the marker or measures it accurately.
    - Clinical validity: how well the marker being analyzed is related to the presence, absence, or risk of specific disease.

[26] Ioannidis JPA. Why Most Published Research Findings Are False. PLoS Medicine. 2005;2(8). https://doi.org/10.1371/journal.pmed.0020124

    - Clinical utility: can the biomarker provide clinically relevant information about diagnosis, treatment, management, or prevention of a disease that will be helpful to a patient, healthcare provider, or family member.
        - Is it actionable in terms of going forward with patients?
    - The process also involves cost effectiveness analysis, which is the comparative analysis of two or more alternative interventions in terms of their health and economic consequences
- Actionability: results that guide decision making, it is an evolving concept and varies with patient, clinician, guideline committee, and payors
    - Contextual for the stage of disease (early vs advanced)
    - Guidelines and FDA approved drug labels formally define accepted criteria
    - Actionability is not binary but is supported with a continuum of evidence
    - Fit-for-purpose (or matched) benefit – risk of managed patient group
    - Framework for Clinical Utility
        - Who should be tested and under what circumstance?
        - What does the test tell us that we did not know?
        - Against what comparator is the test measured?
        - Can we act on the information provided by the test?
        - Will we act on the information provided by the test?
        - What is the effectiveness of the action?
        - Does the outcome of action change in a way in which we find value?
    - Clinical-grade vs Research-grade: Assays
        - 'Biomarkers' are not validated, or approved. 'Biomarker assays' are validated.
            - Clinical-grade assays are much more than just testing clinical samples
            - Clinical-grade assays have to be of highest quality because they inform critical patient management decisions

New Appreciation of Study Designs

- Randomized controlled trials can have compromised value
    - Include only narrowly defined, less ill patients (general validity in question)
    - Difficult to find time and funding for all trials desired
    - Not 'real world' studies
- Registries bring value to evidence collection
    - Permits collection of real-world data to complement and extend RCT data
    - Facilitates collection of comprehensive and unbiased data on diagnostic tests to enhance the available body of evidence for informed patient management decisions
    - Provides insights into short and long-term outcomes
    - Allows health systems, clinicians, and patients to work together to create a setting for generating evidence in practice
- Levels of Evidence
    - Similarity of inclusionary and exclusionary criteria (homogenous vs heterogeneous) across tested sample sets including intended use population
    - Number of patients and events in each sample set
    - Expected 'effect size' of tested diagnostic
    - Expected number of events (prevalence)
    - Single center versus multi-center collection
    - Study Design used (retrospective (selection criteria), chronological, prospective, prospective-retrospective, single-arm with historical control, etc.)
    - Study Objectives—Non-inferiority vs. Superiority vs. Equivalence
    - Critical that pre-specified statistical analysis plans be used for validation

Statistical Considerations

- Metrics for Test Performance
  - Prioritize individual classification over group averages
  - No single statistical measure provides sufficient insight
  - Predictive values (NPV and PPV) are more important than sensitivity and specificity (clinically relevant)
  - ROC curves are informative but not directly clinically relevant
  - Multivariate analysis with standard measures are critical
  - Methods based on risk stratification have recently been proposed to compare models
  - Bayesian models for diagnostic test performance provide key insights (conditional probabilities; likelihood ratios)
  - Explore integration of conventional factors and molecular biomarkers
- Redefined Statistical Threshold
  - Proposal gaining momentum to set the statistical threshold at 0.005[27]
  - More focus on effect sizes and confidence intervals, treating the P value as a continuous measure
  - Proposal should not be used to reject publications of novel findings with $0.005 < P < 0.05$ properly labelled as suggestive evidence
  - Failing to reject the null hypothesis does not mean accepting the null hypothesis
- Receiver Operator Curves (ROC)
  - The ROC curve is developed knowing what the disease distribution is and what the normal distribution, or differential diagnosis is
  - The ROC curve measures how much overlap there is between the two
  - Classifying the accuracy of a diagnosis test:
    - 0.90-1.00 = excellent
    - 0.80-.90 = good
    - 0.,70-.80 = fair
    - 0.60-.70 = poor
    - 0.50-.60 = likely random
      - Permits the selection of cut points for dichotomous or binary categorization
  - AUC-ROC is not a directly clinically relevant diagnostic metric
    - Paucity of data compromises confidence of result
    - ROC plots false positives (1-specificity) versus true positives (sensitivity) for every possible cutoff including regions not clinically relevant
    - Requires highly accurate and related reference method to be informative
    - A test with high sensitivity may have an identical or similar AUC to a test with high specificity
    - Binary interpretation compromised (“Dichotomania”)
    - Weights false positives and false negatives equally
    - Does not address predictive values critical to ruling-in and ruling-out a diagnosis
    - Insensitive to changes in absolute risk of tests compared
  - Dichotomania
    - Disadvantages of dichotomous threshold
      - Information loss
      - Smaller difference between negative and positive groups
      - Threshold significantly impacted by population distribution
      - Intended use rarely represents a step function
      - Less flexibility for intended use

[27] Benjamin DJ, et al. Redefine statistical significance. Nature Human Behavior. 2018;2(1):6-10. https://doi.org/10.1038/s41562-017-0189-z

        - Practical use considers subjects at threshold differently anyway
        - Critically dependent on ground truth accuracy of reference
    - Dichotomous Test Comparison
        - Extremes of dichotomous tests agree with each other a large fraction of time
        - Dichotomous test comparisons are more discordant at thresholds
        - Raises question of ground truth
- Prevalence and Predictive Value
    - The prevalence of intended use may vary from sample set tested
    - Predictive values change with the prevalence of disease
        - As the population prevalence increases, positive predictive value increases
        - As population prevalence increases, negative predictive value decreases
    - The results of a study may not apply to all situations if there are different prevalence rates between discovery and validation studies, or development and clinical practice populations.
    - Predictive value, or, the probability that the patient has the disease, is typically more important to a doctor and patient than sensitivity and specificity

Key Design Issues in Definitive Validation

- Size (and events) of Training and Validation sets
- Training and Validation sets need to be similar (e.g. prevalence, covariates, outcomes, co-morbidities, etc.)
- Study population needs to be same as intended clinical application
    - Sufficiently general; multiple institutions
- Marker well-defined in advance
    - Validation separate from Discovery
    - Locked assay (assays, analytes, model, and thresholds)
        - It is important to lock the assays to ensure the results will be reproducible
    - Same assay used to demonstrate Clinical Validity
- Pre-specified minimally acceptable performance criteria to be met
    - Describe justification
- Individual classification is critical, not group differences
- Anticipated/desirable performance drives sample size calculations

Critiquing and Critically Considering Biomarker Papers

- Are individual clinical validation training and test sets independent and matched with each other as well as with the intended use population?
- Is there a chance that bias or chance was introduced into sample sets being compared?
- Was the assay specifically locked (e.g. analyte(s), weighting, transform and thresholds) before validation testing?
- Was rigorous analytical validation of assay performed and published in peer-reviewed journal?
- Was a pre-specified statistical analysis plan put in place?
- What was the level of evidence collected (e.g. convenience, retrospective, prospective, single-center, multi-center, etc.)?
- Was a commonly accepted reference test used for comparison?
- Was potential of inaccuracy in reference test considered in analysis?
- Was test performance compared to and combined with conventional covariates for standard-of-care?