Challenges and Recommendations in Using Biomarker/Surrogate Endpoints for the Accelerated Approval

Peter Mesenbrink PhD, Executive Director Biostatistics

ASA Biopharmaceuticals Section Regulatory-Industry Statistics Workshop

September 23-25, 2020

NOVARTIS | Reimagining Medicine

Agenda

- Why is NASH challenging
- Are histological endpoints in NASH a reliable surrogate endpoint?
- Study-to-study variability in placebo effect

Challenges in clinical drug development in NASH

- Currently there are no approved therapies for the treatment of NASH or NASH cirrhosis
- There are no validated biomarkers of disease diagnosis or disease progression
- The patient population is heterogeneous with many pre-existing co-morbidities (Type 2 Diabetes, Hypertension, Hypercholesterolemia/Hyperlipidemia, High BMI)
- The standard of care being lifestyle management can lead to many intercurrent events in the estimand used in estimating the treatment effect
- Phase 2b studies conducted before large Phase 3 studies often not large enough to rule out results observed are not due to chance
- Treatment response in monotherapy to date is low (The one intervention to meet its primary objective for histology in Phase 3 had a clinical response rate of 23%)

NOVARTIS | Reimagining Medicine

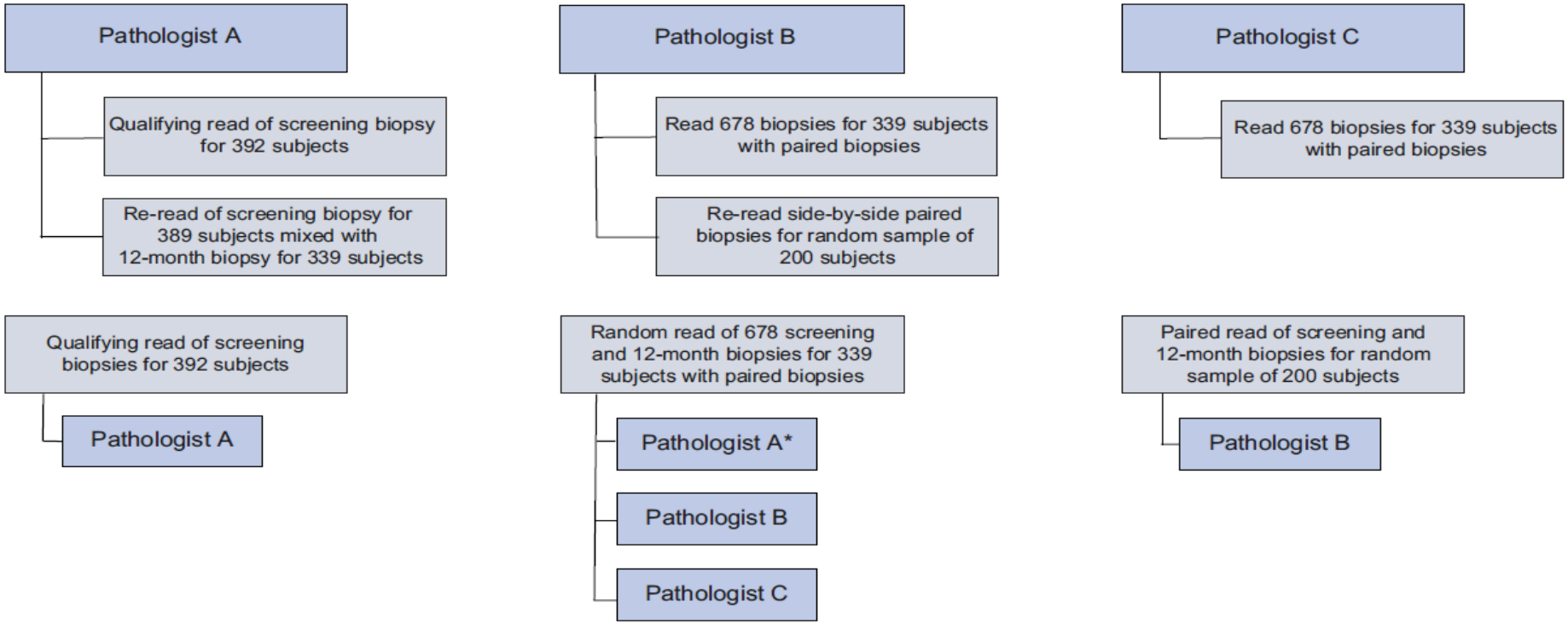
Are changes in histology reliable surrogates in NASH?

- In most NASH Phase 2b/3 clinical trials, a single pathologist is responsible for reading the biopsy samples that are collected to determine patient eligibility in the study and in calculating most of the key efficacy endpoints
- Screening/baseline biopsy is usually read twice
  - Once to determine patient eligibility
  - A second time as a pair with the end of study biopsy (the pathologist generally does not know the order of the biopsies when performing the paired readings)
- Were all patients really eligible to enter the study based on NAS score and fibrosis stage?
- Are patients who were classified as responders really non-responders?
- Are patients who were classified as non-responders really responders?

Study performed to evaluate inter-rater and intra-rater variability for assessing histological change in NASH

- A re-read of the biopsies by 2 independent histologists was performed for the EMMINENCE study (Davison et. Al. JHep 2020)
- MSDC-0602K which was evaluated at 3 dose levels vs. placebo in the study (N=392) is an insulin sensitizer which demonstrated significant reduction at the 2 high doses with respect to
  - Glucose
  - Insulin
  - Liver enyzmes
  - NAS
  - Did not show improvement in primary and secondary histological endpoints (Harrison 2019)

Will different pathologists see different results?



How were readers evaluated?

- Inter-reader reliability was estimated by comparing all possible pairs of the hepatopathologists with respect to
  - NASH scores (ballooning, inflammation, steatosis, fibrosis)
  - NAS total scores
  - NASH diagnosis at baseline
  - Improvement fibrosis without worsening of NASH
  - NASH resolution without worsening of fibrosis
- Present both weighted (using linear weights) and unweighted kappa coefficients
- Present percentage agreement observed and percentage agreement expected by chance

Inter-reader reliability on 678 biopsies from 339 patients with NASH (1/2)

NASH CRN Score	Inter-reader comparison	% Agree	% Agree expected by chance	Unweighted Kappa (95% CI)	Weighted Kappa (95% CI)
Ballooning	Pathologist A vs. B	62.83	33.64	0.440 (0.386, 0.494)	0.543 (0.494, 0.592)
	Pathologist A vs. C	64.60	37.28	0.436 (0.382, 0.490)	0.523 (0.474, 0.571)
	Pathologist B vs. C	60.18	35.17	0.386 (0.332, 0.439)	0.486 (0.439, 0.533)
	Overall agreement	45.58			
Inflammation	Pathologist A vs. B	57.96	42.99	0.263 (0.204, 0.321)	0.323 (0.267, 0.378)
	Pathologist A vs. C	65.34	52.00	0.278 (0.209, 0.346)	0.322 (0.257, 0.386)
	Pathologist B vs. C	57.82	42.49	0.267 (0.209, 0.324)	0.338 (0.284, 0.392)
	Overall agreement	42.33			
Steatosis	Pathologist A vs. B	56.34	28.59	0.389 (0.338, 0.439)	0.543 (0.500, 0.587)
	Pathologist A vs. C	67.40	30.15	0.533 (0.484, 0.583)	0.650 (0.609, 0.691)
	Pathologist B vs. C	66.22	30.54	0.514 (0.464, 0.564)	0.635 (0.593, 0.678)
	Overall agreement	63.32			

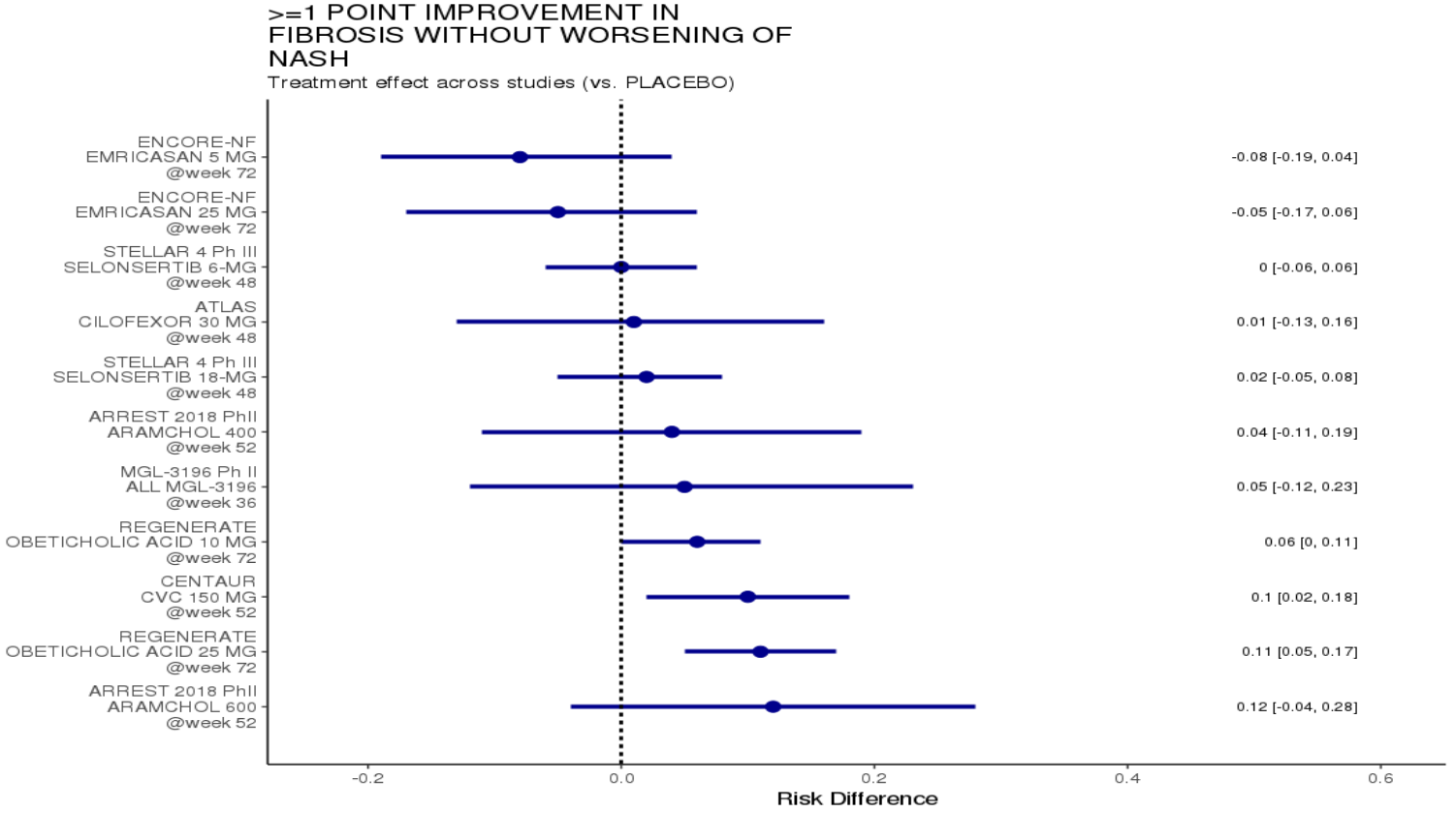
Inter-reader reliability on 678 biopsies from 339 patients with NASH and paired biopsies (2/2)

NASH CRN Score	Inter-reader comparison	% Agree	% Agree expected by chance	Unweighted Kappa (95% CI)	Weighted Kappa (95% CI)
Fibrosis	Pathologist A vs. B	57.23	27.39	0.411 (0.363, 0.459)	0.592 (0.552, 0.632)
	Pathologist A vs. C	42.18	26.37	0.215 (0.173, 0.256)	0.383 (0.346, 0.419)
	Pathologist B vs. C	53.39	29.33	0.341 (0.294, 0.387)	0.477 (0.439, 0.515)
	Overall agreement	31.27	28.91		
NASH diagnosis	Pathologist A vs. B	80.53	60.62	0.506 (0.434, 0.577)	
	Pathologist A vs. C	81.42	71.29	0.353 (0.267, 0.438)	
	Pathologist B vs. C	76.84	65.61	0.327 (0.254, 0.399)	
	Overall agreement	69.47	66.60		
Improvement in fibrosis with no worsening of NASH	Pathologist A vs. B	76.11	60.70	0.392 (0.286, 0.497)	
	Pathologist A vs. C	75.81	65.49	0.299 (0.184, 0.413)	
	Pathologist B vs. C	78.76	63.94	0.411 (0.301, 0.521)	
	Overall agreement	65.78			

Re-read changing interpretation of histology results

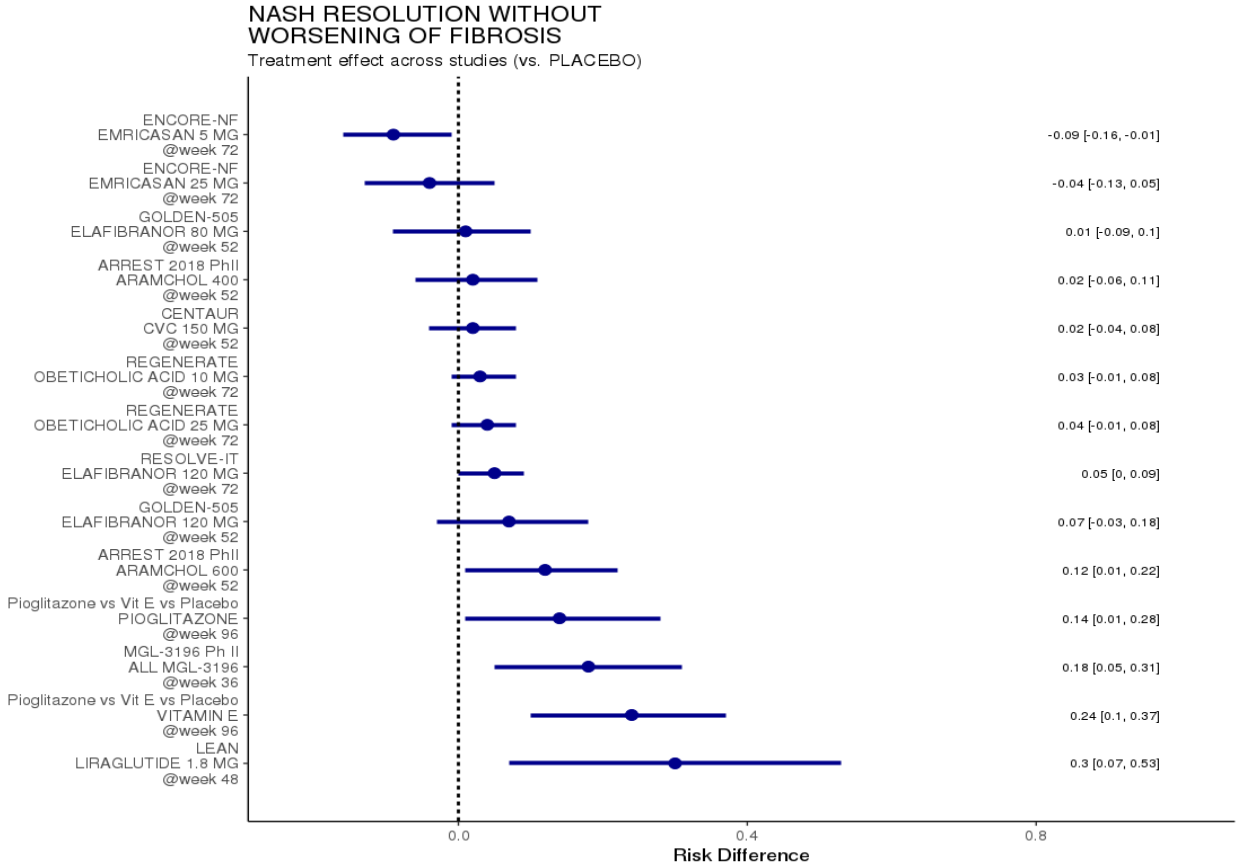
- When the original reader was asked to re-read the same baseline biopsies mixed with follow-up ones, the re-reads showed that 16% of patients were determined to have “met” the primary endpoint of the EMMINENCE study from the re-read
- Why is this relevant?

Effect sizes for endpoints is small (1/2)



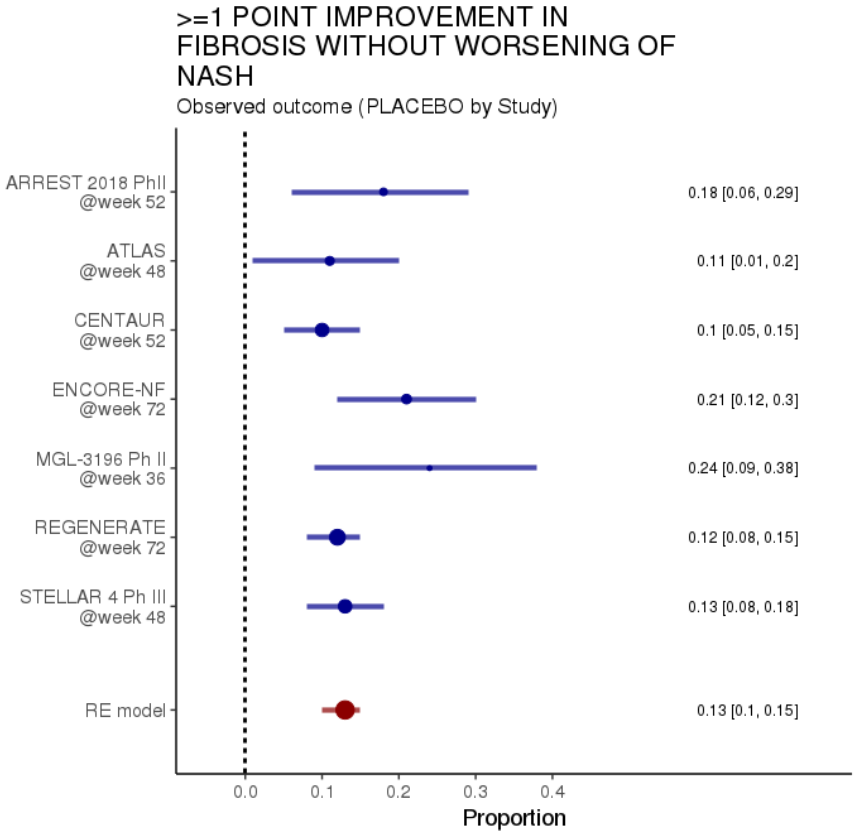
NOVARTIS | Reimagining Medicine

Effect sizes for endpoints is small (2/2)



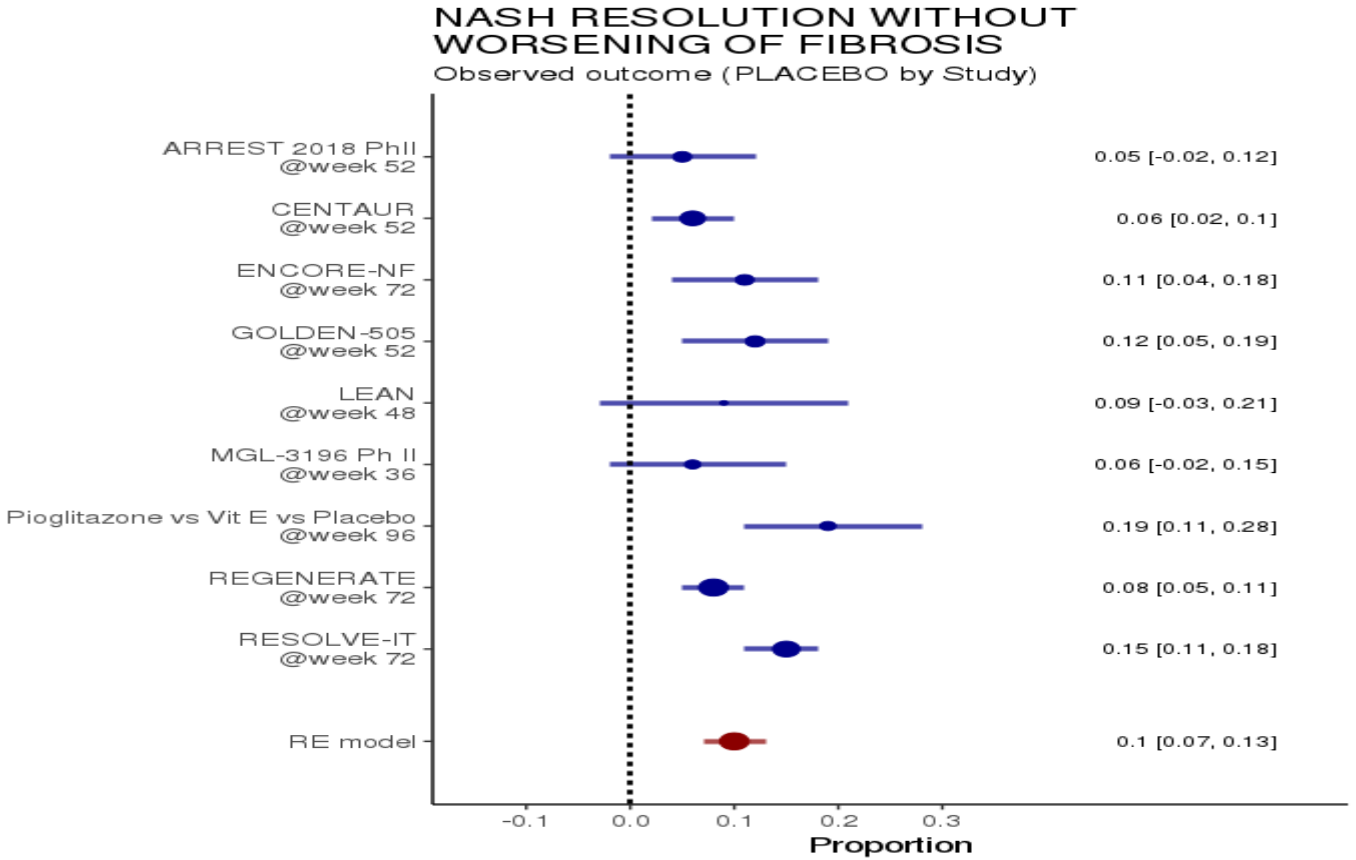
NOVARTIS | Reimagining Medicine

Highly variable placebo effect (1/2)



NOVARTIS | Reimagining Medicine

Highly variable placebo effect (2/2)



NOVARTIS | Reimagining Medicine

Where do we go from here?

- Continue to support research through interventional and non-interventional studies to develop improved non-invasive surrogate endpoints that are predictive of clinical outcomes
- The NASH community is very active to find better endpoints to allow for better determinations of treatment effectiveness
  - Non-Invasive Biomarkers of Metabolic Liver Disease (NIMBLE) in US
  - Liver Investigation: Testing Marker Utility in Steatohepatitis (LITMUS) in EU
  - Target-NASH (N=15000) 5-year non-interventional longitudinal study US/EU
- Agree to an adaptive platform trial structure that would allow investigational treatments being evaluated to be assessed under the same set of rules (EU-Pearl)
- Agree to committees of 3 readers in the review of biopsies from platform trials

Thank you

NOVARTIS | Reimagining Medicine