Currently proposed endpoints in PSC: the search for reliable surrogate outcome parameters

Cyriel Ponsioen, MD PhD
Department of Gastroenterology & Hepatolog
Amsterdam University Medical Centers
Amsterdam, Netherlands

Outline

- requirements for endpoints

- recommended endpoints in reflection paper
  - disease modifiers only

Sources

HEPATOLOGY

Official Journal of the American Association for the Study of Liver Diseases

SPECIAL ARTICLE

Surrogate Endpoints for Clinical Trials in Primary Sclerosing Cholangitis: Review and Results From an International PSC Study Group Consensus Process

Cyriel Y. Ponsioen,[1] Roger W. Chapman,[2] Olivier Chazouillères,[3] Gideon M. Hirschfield,[4] Tom H. Karlsen,[5] Ansgar W. Lohse,[6] Massimo Pinzani,[7] Erik Schrumpf,[8] Michael Trauner,[9] and Gregory J. Gores[10]

IPSCSG

International PSC Study Group

REVIEW | HEPATOLOGY, VOL. 68, NO. 3, 2018

Design and Endpoints for Clinical Trials in Primary Sclerosing Cholangitis

Cyriel Y. Ponsioen,[1] Keith D. Lindor,[2] Ruby Mehta,[3] and Lara Dimick-Santos[3]

Ponsioen et al. Hepatology 2015; Sep 29
Ponsioen et al. Hepatology 2018; 68: 1174

Sources

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

15 November 2018
EMA/CHMP/299976/2018
Committee for Medicinal Products for Human Use (CHMP)

Reflection paper on regulatory requirements for the development of medicinal products for chronic non-infectious liver diseases (PBC, PSC, NASH).

hierarchy of endpoints

level 1: true clinical efficacy measure

level 2: validated surrogate endpoint ✗

level 3: non-validated reasonable surrogate endpoint ?

level 4: measure of biological activity ✔

Fleming. Health Aff 2005; 24: 67

endpoint requirements

level 1: true clinical efficacy measure ✗

level 2: validated surrogate endpoint ✗

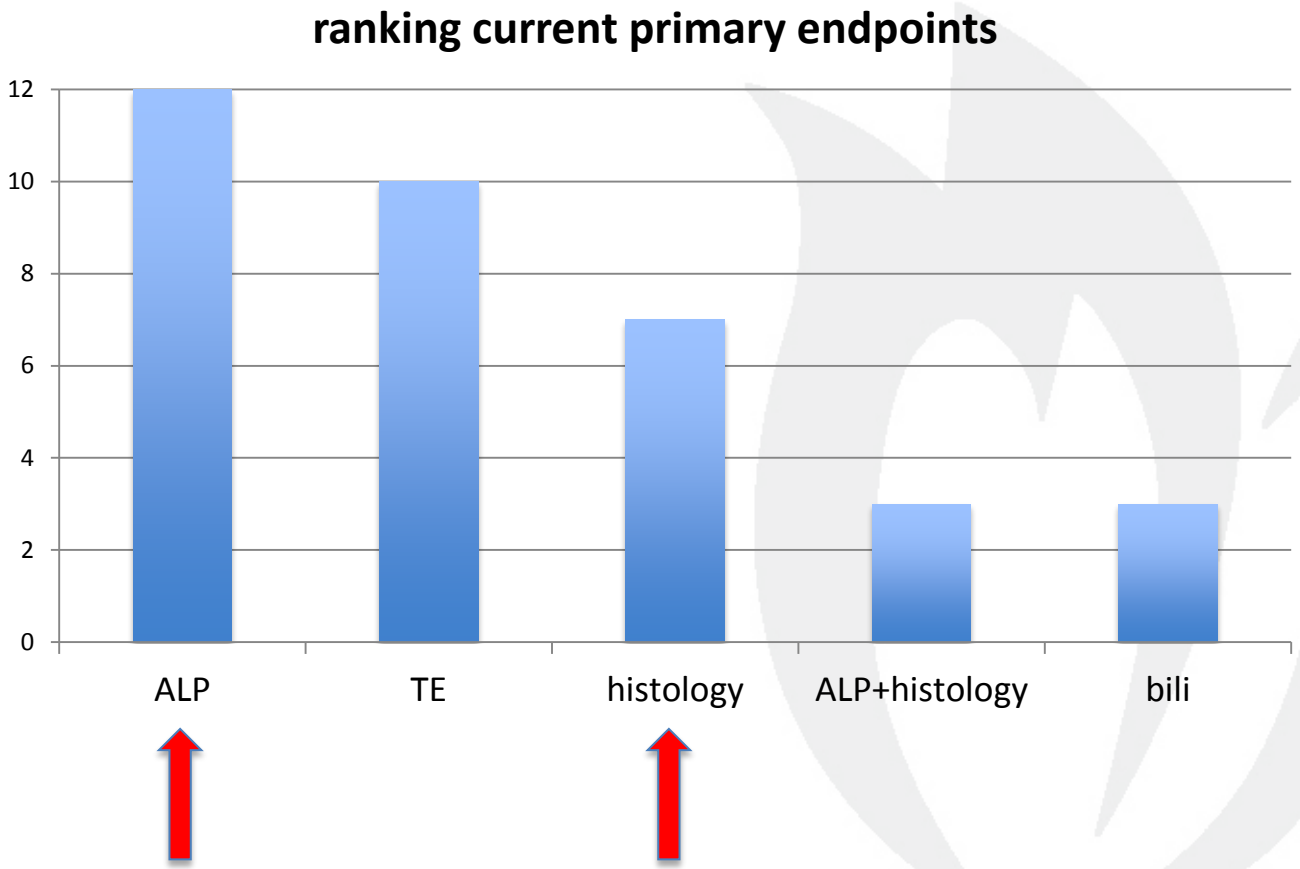
1st author	year	n	type	mth	intervention	endpoints	results
Olsson(7)	1995	84	RCT	36	colchicine vs placebo	1st: clinical 2nd: ALP, bili, ALT, albumin symptoms histology	no changes no difference no changes no change
Lindor(8)	1997	105	RCT	24	UDCA vs placebo	1st: clinical histology 2nd: ALP, bili, AST	no difference NS improvement*
Olsson(5)	2005	219	RCT	60	UDCA vs placebo	1st: clinical 2nd: ALP, ALT, bili	NS NS
Lindor(3)	2009	150	RCT	60	UDCA vs placebo	1st: clinical 2nd: ALP, AST, bili histology	worsening* improvement* n.s **
Muir	2017	234	RCT	24	simtuzumab vs placebo	1st: histology 2nd: clinical ALP	no change no difference no change

level 3: non-validated reasonable surrogate endpoint ✔ ??

level 4: measure of biological activity ✔

requirements for surrogate endpoints

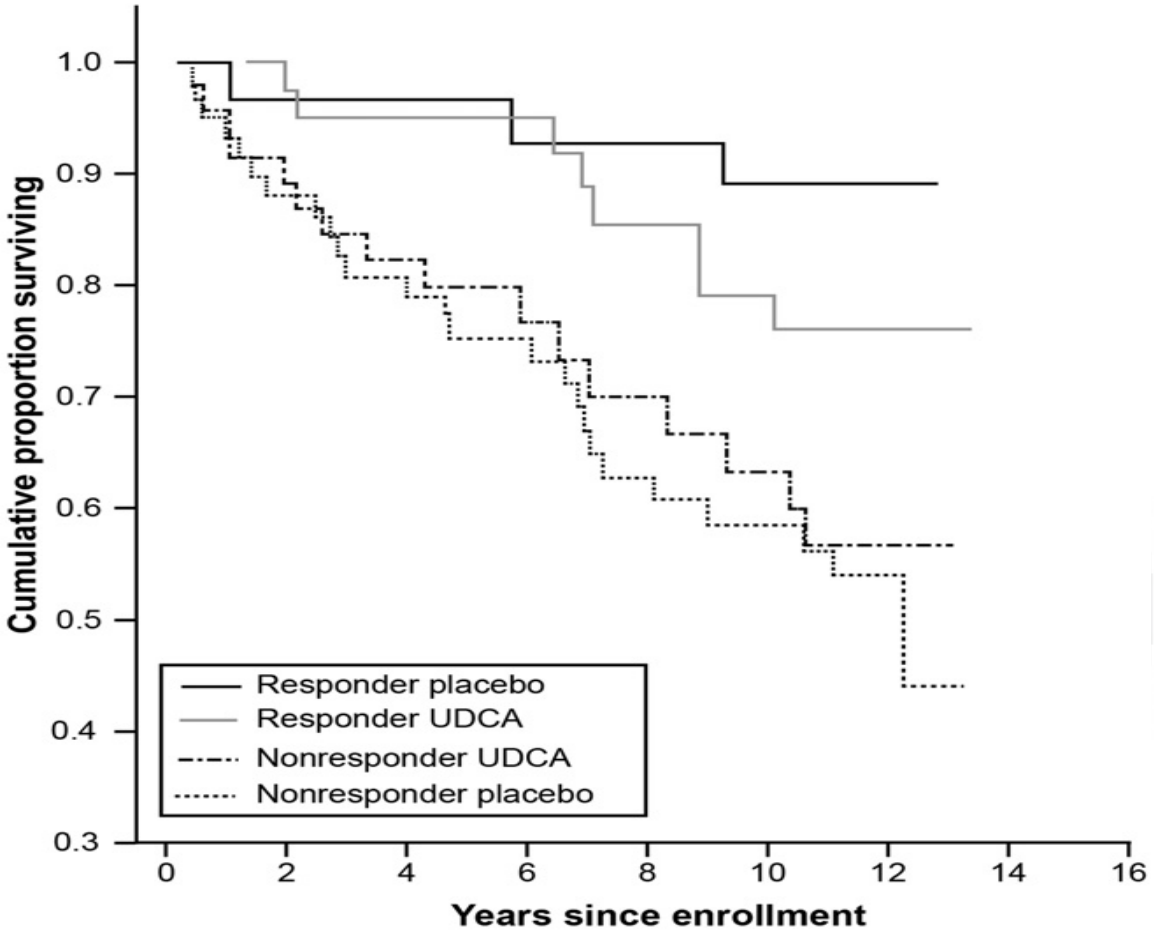
- measurable/interpretable
- sensitive to change
- natural variability contained
- on the pathway to a clinically meaningful endpoint



Ponsioen et al. Hepatology 2015; Sep 29

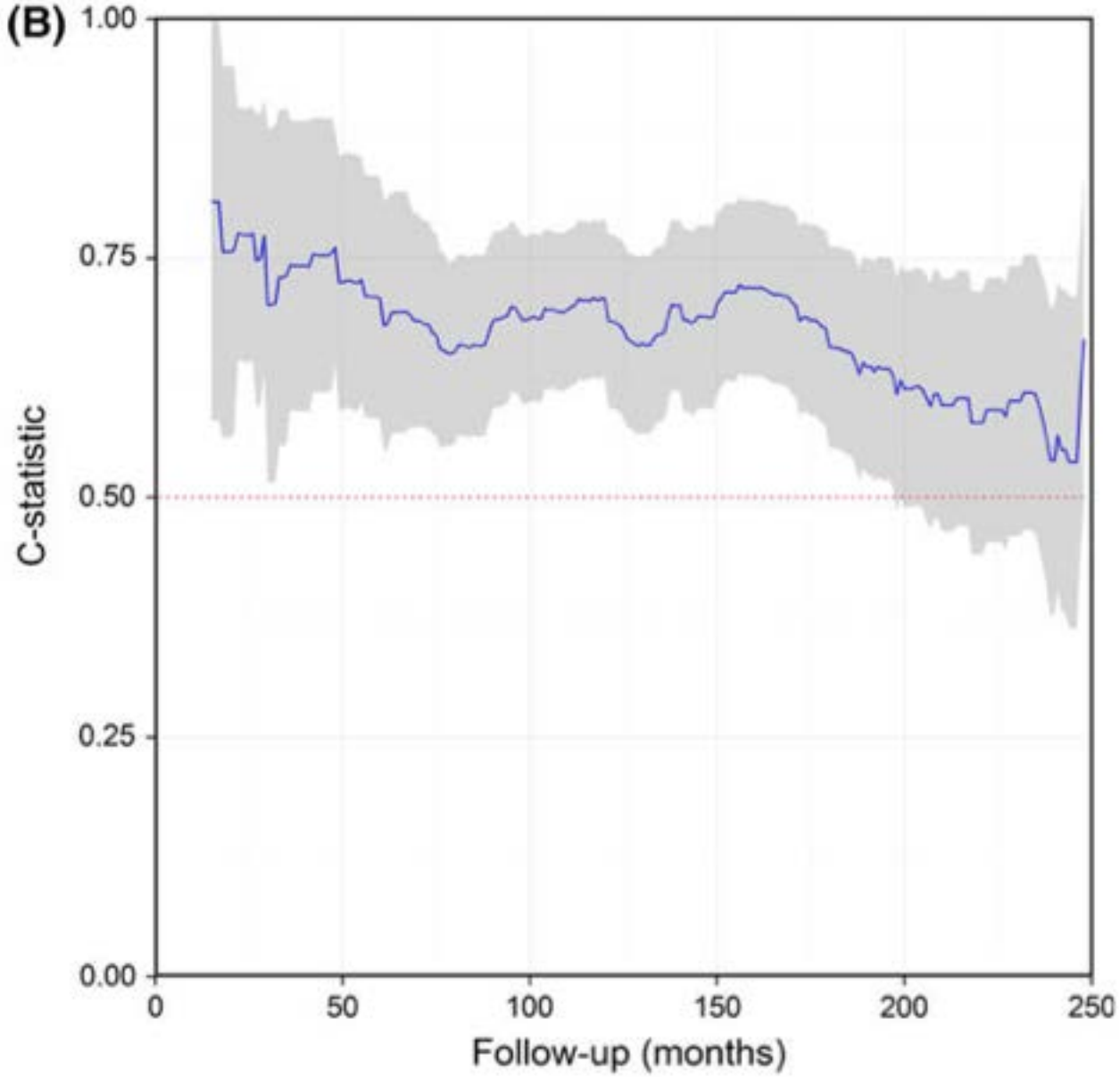
Alkaline Phosphatase

- biochemical hallmark of cholestasis
- used in all studies in past 25 years
- no significant association with clinical outcome observed so far

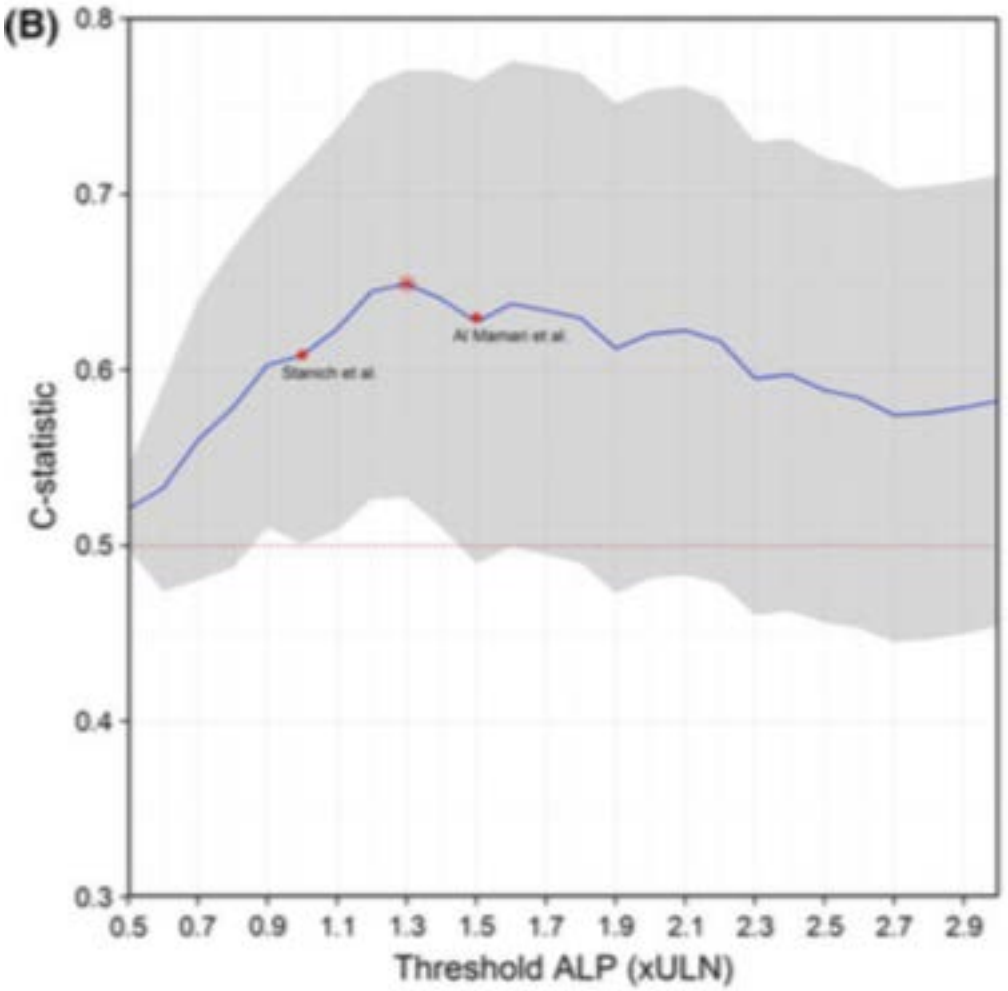


Years	0	2.5	5	7.5	10
Responder placebo	36	32	29	26	25
Responder UDCA	43	40	38	30	27
Nonresponder UDCA	51	41	33	25	23
Nonresponder placebo	65	51	43	33	31

Lindström et al. CGH 2013;11:841



de Vries et al. Liver Int 2016; 36: 1867



de Vries et al. Liver Int 2016; 36: 1867

(surrogate) endpoints requirements

164 1-year intervals in non advanced PSC patients from Amsterdam and Birmingham potentially eligible for a phase II trial

➔

ALP: mean Δ 1-year intervals = 5.7 % (SD=36%)

➔

variability during natural course limited

Endpoints at Week 96	SIM 75 mg (N = 79)	SIM 125 mg (N = 77)	Placebo (N = 78)	P-value 75 mg vs. placebo	P-value 125 mg vs. placebo
Change in hepatic collagen,[1] %					
Mean (SD)	−0.5 (5.78)	0.5 (6.94)	0.0 (4.76)	0.73	0.33
Median (IQR)	−0.6 (−3.3. 0.7)	−1.1 (−2.8, 1.6)	−0.9 (−2.3, 0.9)		
Ishak Fibrosis Stage,[2] % (n)					
Worsening	33.8 (26)	32.8 (22)	44.4 (32)	0.12	0.13
Improvement	32.5 (25)	31.3 (21)	22.2 (16)		
Clinical Events,[3] % (n)					
All	20.3 (16)	18.2 (14)	17.9 (14)	0.74	0.84
Ascending cholangitis	15.2 (12)	14.3 (11)	9.0 (7)	—	—
Change in ALP,[4] U/L					
Mean (SD)	−30 (116.0)	−2 (128.6)	−4 (155.5)	0.29	0.94
Median (IQR)	−12 (−98, 16)	−7 (−50, 38)	4 (−25, 42)		

[1]Based on mixed-effected models for repeated measures.
[2]Based on Cochran-Mantel-Haenszel tests.
[3]Based on stratified log-rank tests.
[4]Based on t-tests.

Muir et al J Hep 2017; 66: S73

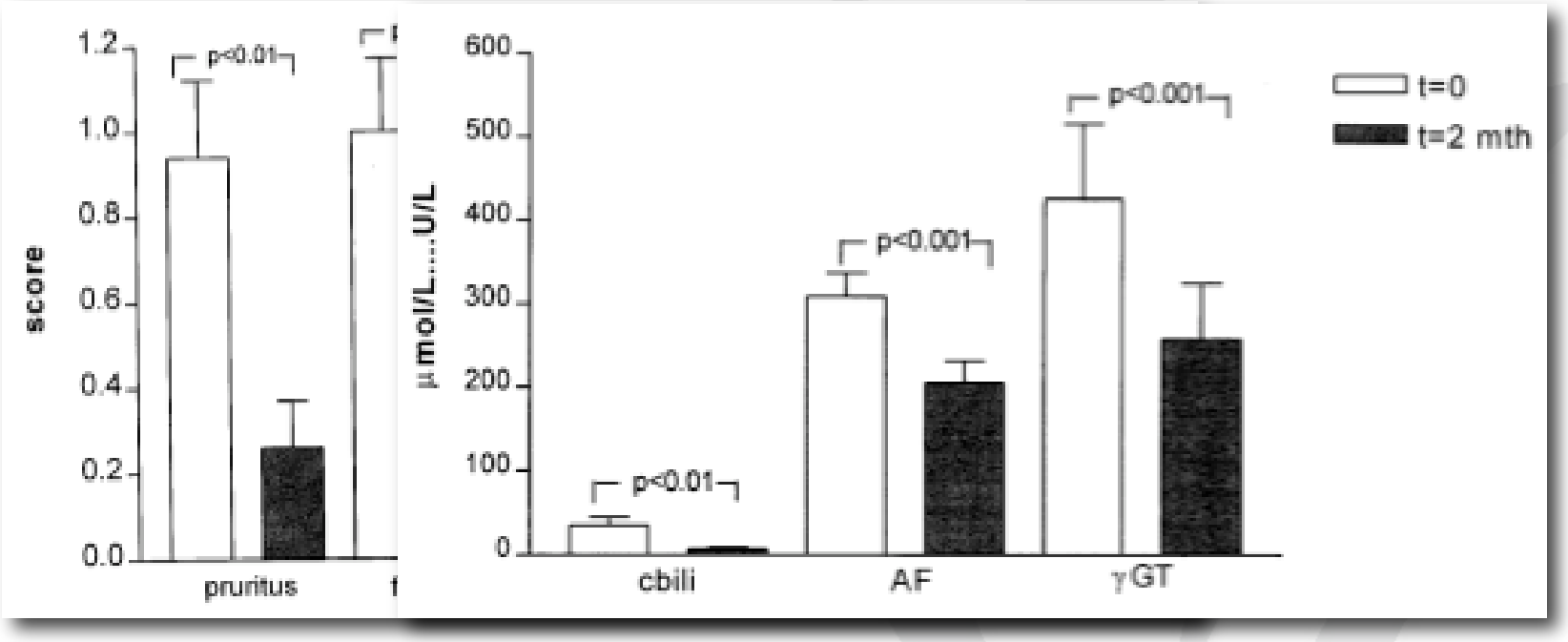
(surrogate) endpoints requirements

- "A Correlate does not a Surrogate Make"
- Criteria for Surrogate Endpoints
  - Measurable/Interpretable
  - Sensitive

Fleming and deMets, 1996

sensitive to meaningful change?

short-term stenting: results



Ponsioen *et al.* Am J Gas 1999; 94: 2403

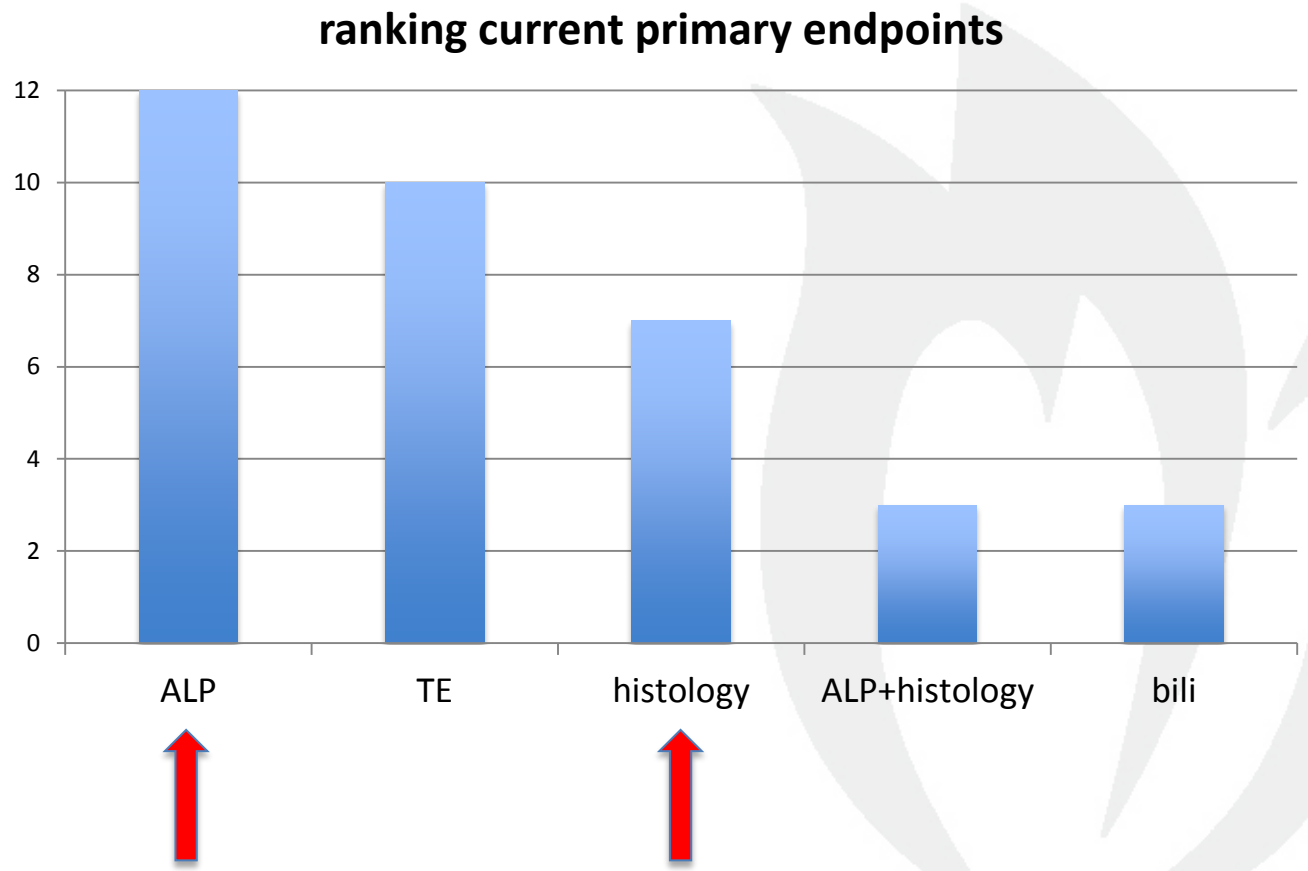
IPSCSG statement 2

Alkaline phosphatase is widely recognized as a clinical measure of cholestasis. Currently, albeit not formally validated, it is regarded as a potential surrogate outcome parameter *[EL 4, RG D].*

Grades of Recommendation

A	consistent level 1 studies
B	consistent level 2 or 3 studies ***or*** extrapolations from level 1 studies
C	level 4 studies ***or*** extrapolations from level 2 or 3 studies
D	level 5 evidence ***or*** troublingly inconsistent or inconclusive studies of any level

"Extrapolations" are where data is used in a situation that has potentially clinically important differences than the original study situation.


ranking current primary endpoints
12
10
8
6
4
2
0
ALP
TE
histology
ALP+histology
bili

liver biopsy

advantages:

- assesses directly disease activity in the target organ
  (face/construct validity)
- has been gold standard to asess liver fibrosis/cirrhosis
  → acceptable to regulatory bodies
- provides human histological material for MOA investigation and safety assessment

disadvantages:

- sampling error *for diagnosis*
- invasive procedure

IPSCSG
International PSC Study Group

Interpretable?

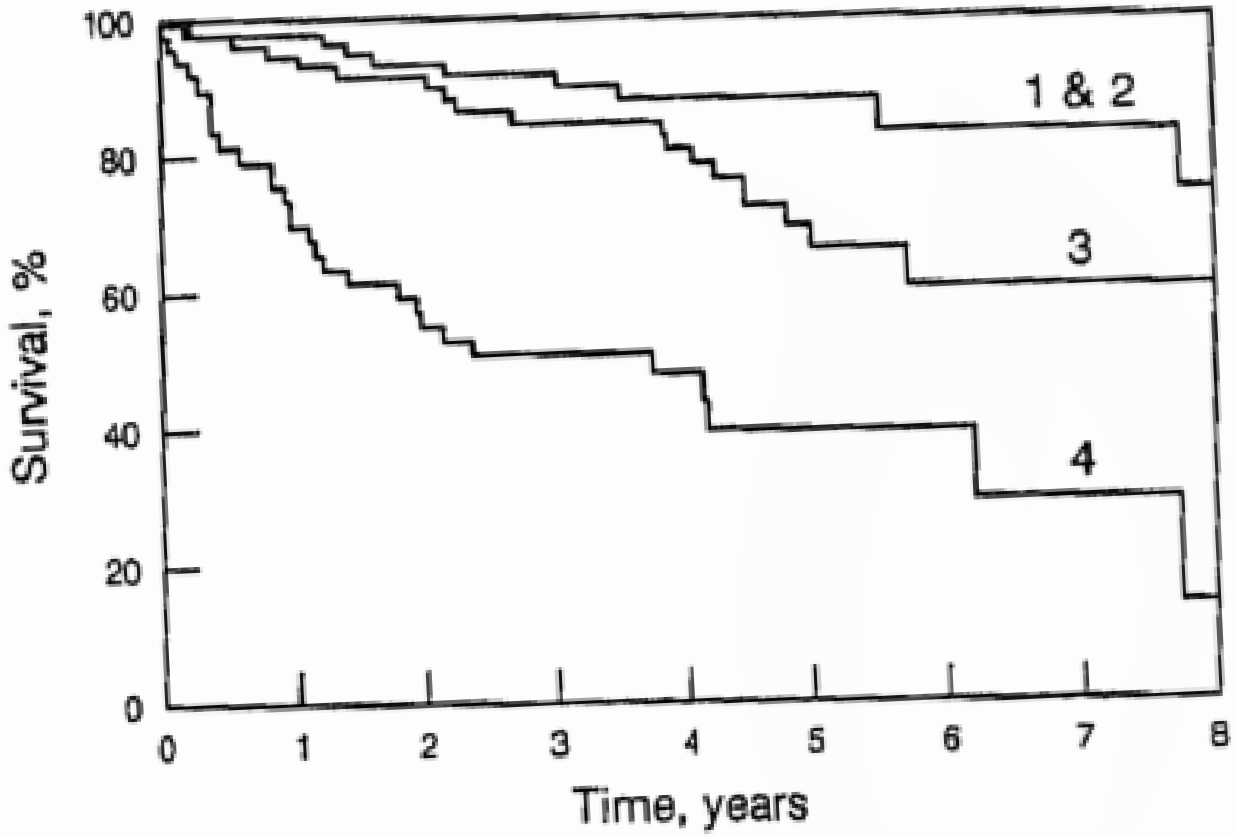
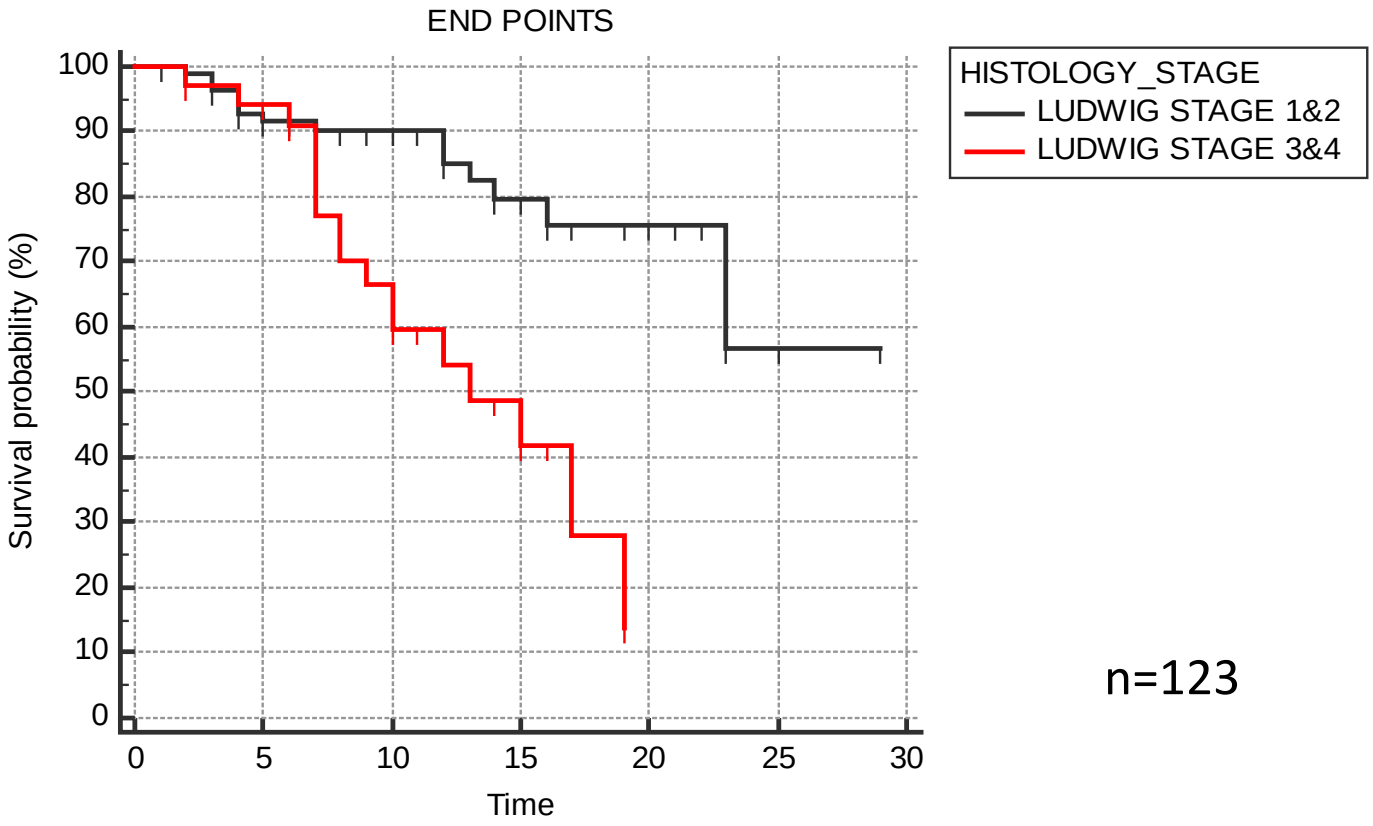


FIG. 2. Kaplan-Meier estimated survival by histologic stage on liver biopsy at study entry.

Wiesner, Hepatology 1989; 10:3=430

Interpretable?



Number at risk

Group: LUDWIG STAGE 1&2						
87	72	40	20	8	1	0
Group: LUDWIG STAGE 3&4						
36	29	12	4	0	0	0

n=123

Interpretable?

Nakanuma staging vs Ishak vs Ludwig

	transplant-free survival		liver transplant	
	HR 95% CI	p	HR 95%CI	p
Nakanuma stage (1-4)	2.58 (1.66-4.03)	<0.001	3.38 (2.01-5.69)	<0.001
Ishak stage (0-6)	1.37 (1.13-1.65)	0.001	1.58 (1.26-1.98)	<0.001
Ludwig stage (0-4)	2.16 (1.28-3.65)	0.004	3.34 (1.74-6.41)	<0.001

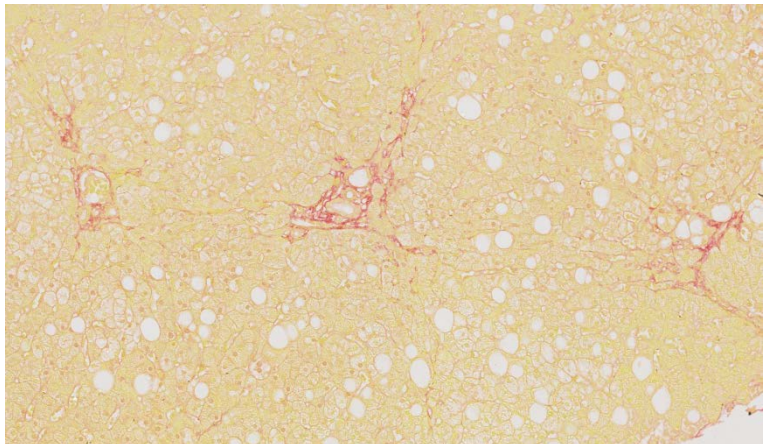
de Vries et al. Hepatology 2017; 65: 907

Measurable?

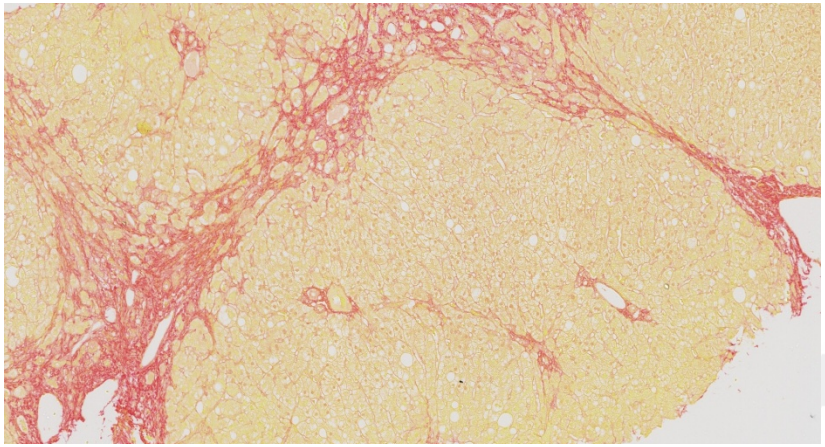
- Interobserver variability kappa's 0.56-0.64
- Six European expert liver pathologists all favoured Nakanuma

WWW.IBD-AMC.NL

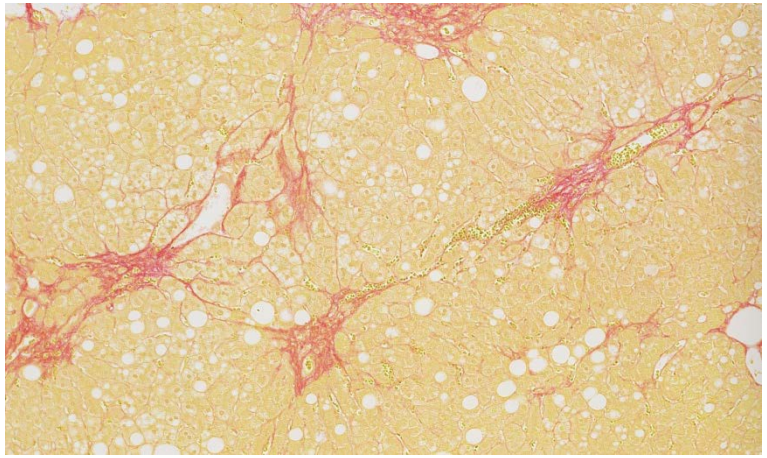
Collagen Proportionate Area (CPA)



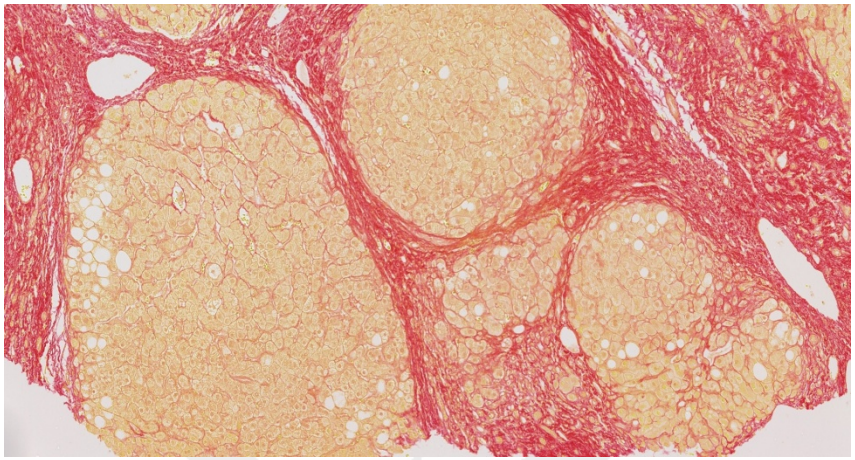
MILD: Ishak 1/2 **CPA: 0.17%**



SEVERE: Ishak 5 **CPA: 11.3%**



MODERATE: Ishak 3/4 **CPA: 5.7%**



VERY SEVERE: Ishak 6 **CPA: 31%**

courtesy of prof Paul Dillon

natural variability?

IPSCSG
International PSC Study Group

Table 1. Observed Transition Between Stages of Disease During the Follow-up

	New Stage				
Old Stage	I	II	III	IV	Total
I	9	2	3	2	16
II	3	15	22	6	46
III	4	10	39	20	73
IV	2	3	8	52	65
Total	18	30	72	80	200

At second biopsy 53% of stage I- III PSC patients had progressed, but 14% of stage III-IV patients now had stage I-II

Angulo et al. Am J Gas 1999; 94: 3310

Endpoints at Week 96	SIM 75 mg (N = 79)	SIM 125 mg (N = 77)	Placebo (N = 78)	P-value 75 mg vs. placebo	P-value 125 mg vs. placebo
Change in hepatic collagen,[1] %					
Mean (SD)	−0.5 (5.78)	0.5 (6.94)	0.0 (4.76)	0.73	0.33
Median (IQR)	−0.6 (−3.3. 0.7)	−1.1 (−2.8, 1.6)	−0.9 (−2.3, 0.9)		
Ishak Fibrosis Stage,[2] % (n)					
Worsening	33.8 (26)	32.8 (22)	44.4 (32)	0.12	0.13
Improvement	32.5 (25)	31.3 (21)	22.2 (16)		
Clinical Events,[3] % (n)					
All	20.3 (16)	18.2 (14)	17.9 (14)	0.74	0.84
Ascending cholangitis	15.2 (12)	14.3 (11)	9.0 (7)	—	—
Change in ALP,[4] U/L					
Mean (SD)	−30 (116.0)	−2 (128.6)	−4 (155.5)	0.29	0.94
Median (IQR)	−12 (−98, 16)	−7 (−50, 38)	4 (−25, 42)		

[1]Based on mixed-effected models for repeated measures.
[2]Based on Cochran-Mantel-Haenszel tests.
[3]Based on stratified log-rank tests.
[4]Based on t-tests.

Muir et al J Hep 2017; 66: S73

sampling variability?

	Total (n = 56)				
Disagreement (grade)	*0*	*1*	*2*	*3*	*4*
Portal lymphocyte infiltration	68	29	3	0	
Portal fibrosis	63	34	2	2	
Periportal fibrosis	71	27	2	0	
Bile duct proliferation	80	14	6	0	
Cholestasis	93	7	0	0	
Steatosis	96	4	0	0	
Stage	73	16	9	2	0

Olsson J Clin Pathol 1995; 48: 933

safety?

IPSCSG
International PSC Study Group

Summary safety of liver biopsy in PSC trials

- No mortality reported in 1851 biopsies in 782 PSC subjects.
- Reported bile leakage in ±0.2%

Literature:

- Mortality: 0.09-0.3/1000 in 68,276 resp. 98,445 biopsies[1,2]
- SAE rate: 0.57% in 2084 Bx (57% under US guidance)[3]
  US guided: 0.5 versus 2.2 % in 836 pts[4]

[1] Piccinino et al. J Hepatol 1986; 2: 165
[2] Poynard et al. Can J Gas 2000; 14: 543
[3] Cadranel et al Hepatology 2000; 32: 477
[4] Lindor et al . Hepatology 1996; 23: 1079

IPSCSG statement 4

Liver histology has the potential to be a robust surrogate endpoint for clinical trials in PSC *[EL2b, RG B].*

Grades of Recommendation

A	consistent level 1 studies
B	consistent level 2 or 3 studies ***or*** extrapolations from level 1 studies
C	level 4 studies ***or*** extrapolations from level 2 or 3 studies
D	level 5 evidence ***or*** troublingly inconsistent or inconclusive studies of any level

"Extrapolations" are where data is used in a situation that has potentially clinically important differences than the original study situation.

Ponsioen et al. Hepatology 2015; Sep 29

ALP + Histology

IPSCSG statement

In the absence of a convincing single surrogate endpoint combining multiple endpoints is considered advisable and should be explored further *[EL 5, RG D].*

phase

- 1/2A: maximize chance of finding potential contrast:
  → composite
- 2B/3: maximize reassurance of beneficial effect:
  → co-primary

Grades of Recommendation

A	consistent level 1 studies
B	consistent level 2 or 3 studies ***or*** extrapolations from level 1 studies
C	level 4 studies ***or*** extrapolations from level 2 or 3 studies
D	level 5 evidence ***or*** troublingly inconsistent or inconclusive studies of any level

"Extrapolations" are where data is used in a situation that has potentially clinically important differences than the original study situation.

Thank you