PET interpretation issues: experience in NHL with 5PS

Emmanuel Itti, MD PhD

H. Mondor Hospital, AP-HP
Paris-Est University, Créteil, FR

Menton April 8th, 2010

Methods

- 49 IVS patients from 4 PET centers
  (Créteil n=15; Dijon n=14; Cuneo n=11; Rouen n=9)
- PET/CT at baseline and 2 cycles
- Interpretation by 3 observers using the 5PS
- Transfers/readings on Positoscope workstations
- Inter-observer agreement (Kappa)
- Quantification with ΔSUV (66% cut-off)

5-point scale
weighted Kappa (Cohen)

Observer A	Créteil
Observer B	Dijon

Observer B	Observer A 1	2	3	4	5	
1	4	1	2	0	1	(16,3%)
2	3	4	0	0	0	(14,3%)
3	0	1	6	0	0	(14,3%)
4	0	1	1	7	2	(22,4%)
5	1	0	0	2	13	(32,7%)
	(16,3%)	(14,3%)	(18,4%)	(18,4%)	(32,7%)	

Weighted Kappa	0,744
Standard error (Kw'=0)	0,143
Standard error (Kw'#0)	

Observer A	Créteil
Observer B	Cuneo

Observer B	Observer A 1	2	3	4	5	
1	5	0	2	0	2	(18,4%)
2	3	5	2	1	2	(26,5%)
3	0	2	5	5	1	(26,5%)
4	0	0	0	3	7	(20,4%)
5	0	0	0	0	4	(8,2%)
	(16,3%)	(14,3%)	(18,4%)	(18,4%)	(32,7%)	

Weighted Kappa	0,568
Standard error (Kw'=0)	0,126
Standard error (Kw'#0)	

Observer A	Dijon
Observer B	Cuneo

Observer B	Observer A 1	2	3	4	5	
1	5	2	1	0	1	(18,4%)
2	2	4	3	1	3	(26,5%)
3	1	1	3	8	0	(26,5%)
4	0	0	0	2	8	(20,4%)
5	0	0	0	0	4	(8,2%)
	(16,3%)	(14,3%)	(14,3%)	(22,4%)	(32,7%)	

Weighted Kappa	0,604
Standard error (Kw'=0)	0,125
Standard error (Kw'#0)	0,099

Landis and Koch scale	
< 0	no agreement
0.00 – 0.20	slight
0.21 – 0.40	fair
0.41 – 0.60	moderate
0.61 – 0.80	substantial
0.81 – 1.00	almost perfect

Observer A	Créteil
Observer B	Dijon

Observer B	Observer A		
	0	1	
0	12	3	(30,6%)
1	3	31	(69,4%)
	(30,6%)	(69,4%)	

Kappa	0,712
Standard error	0,110
95% CI	0,496 to 0,928

5-point scale
binary (cut-off ≥3, MBP)
Kappa (Cohen)

Observer A	Créteil
Observer B	Cuneo

Observer B	Observer A		
	0	1	
0	13	9	(44,9%)
1	2	25	(55,1%)
	(30,6%)	(69,4%)	

Kappa	0,533
Standard error	0,124
95% CI	0,289 to 0,776

Landis and Koch scale	
< 0	no agreement
0.00 – 0.20	slight
0.21 – 0.40	fair
0.41 – 0.60	moderate
0.61 – 0.80	substantial
0.81 – 1.00	almost perfect

Overall Kappa (Fleiss)
(3 obs.) $\kappa = 0.58$

Observer A	Dijon
Observer B	Cuneo

Observer B	Observer A		
	0	1	
0	13	9	(44,9%)
1	2	25	(55,1%)
	(30,6%)	(69,4%)	

Kappa	0,533
Standard error	0,124
95% CI	0,289 to 0,776

Observer A	Créteil
Observer B	Dijon

Observer B	Observer A		
	0	1	
0	21	1	(44,9%)
1	3	24	(55,1%)
	(49,0%)	(51,0%)	

Kappa	0,836
Standard error	0,078
95% CI	0,683 to 0,990

5-point scale
binary (cut-off ≥4, liver)
Kappa (Cohen)

Observer A	Créteil
Observer B	Cuneo

Observer B	Observer A		
	0	1	
0	24	11	(71,4%)
1	0	14	(28,6%)
	(49,0%)	(51,0%)	

Kappa	0,555
Standard error	0,118
95% CI	0,323 to 0,787

Landis and Koch scale	
< 0	no agreement
0.00 – 0.20	slight
0.21 – 0.40	fair
0.41 – 0.60	moderate
0.61 – 0.80	substantial
0.81 – 1.00	almost perfect

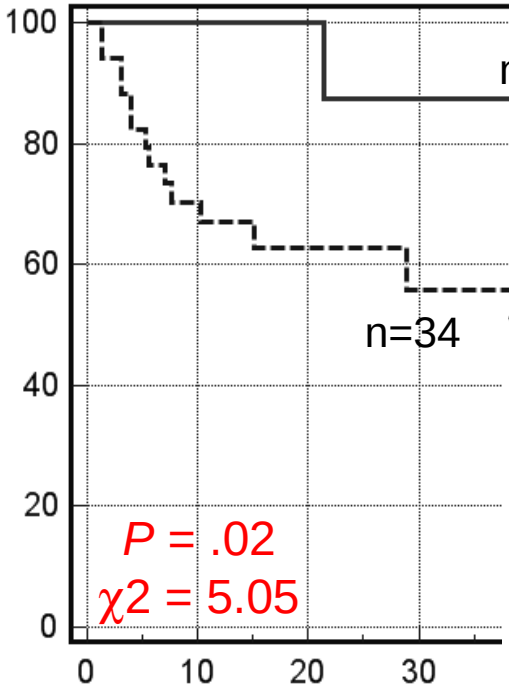
Overall Kappa (Fleiss)
(3 obs.) $\kappa = 0.61$

Observer A	Dijon
Observer B	Cuneo

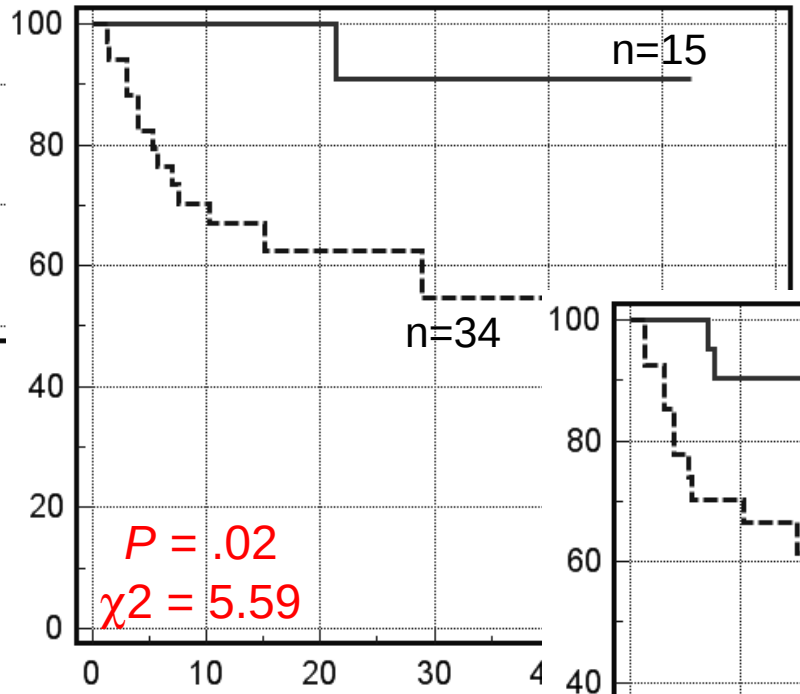
Observer B	Observer A		
	0	1	
0	22	13	(71,4%)
1	0	14	(28,6%)
	(44,9%)	(55,1%)	

Kappa	0,492
Standard error	0,121
95% CI	0,255 to 0,728

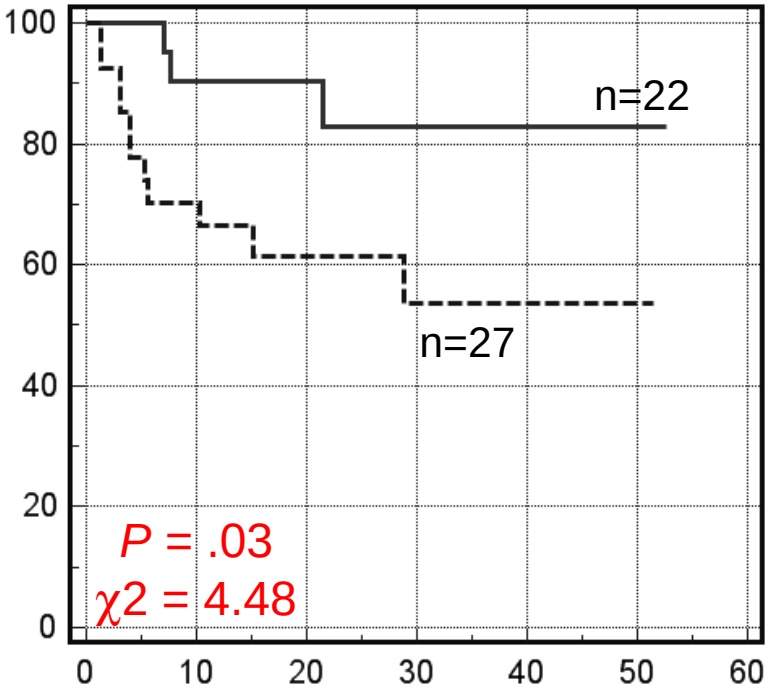
5-point scale (cut-off ≥3, MBP) Event-free survival



Créteil :
2-y EFS : 88% vs. 63%



Dijon :
2-y EFS : 91% vs. 63%

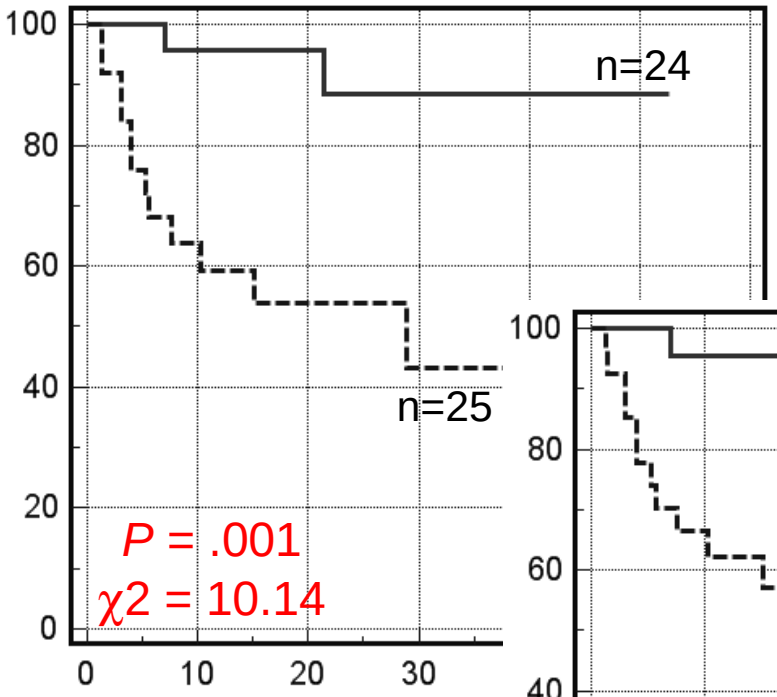


Cuneo :
2-y EFS : 83% vs. 61%

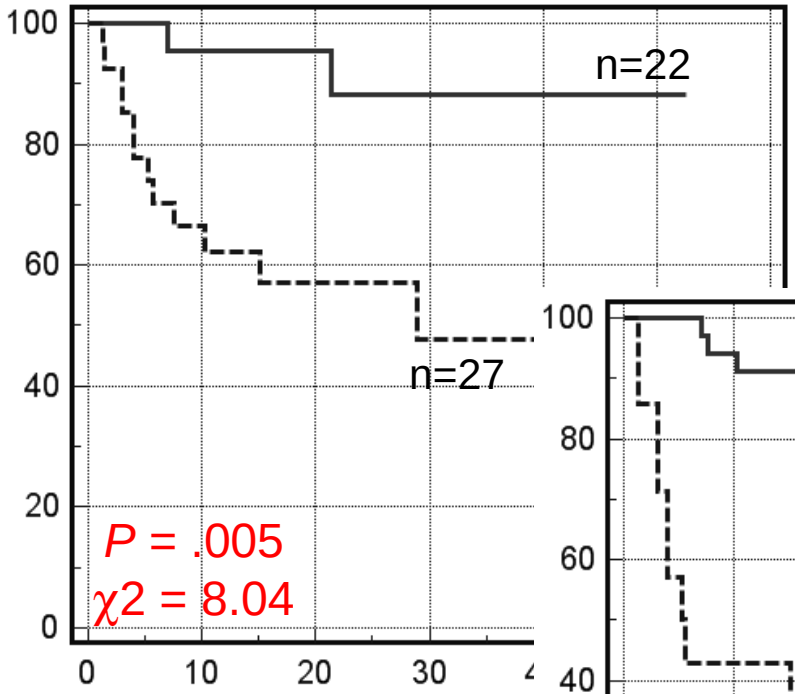
# of events = 14
Median f-u = 25 mo

→ Generates false-positives

5-point scale
(cut-off ≥4, liver)
Event-free survival

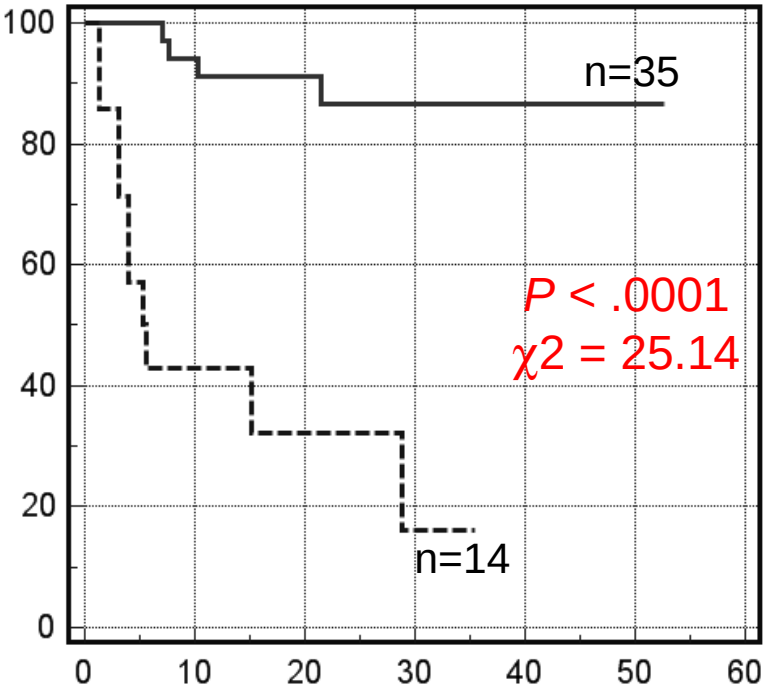


Créteil :
2-y EFS : 89% vs. 54%



Dijon :
2-y EFS : 88% vs. 57%

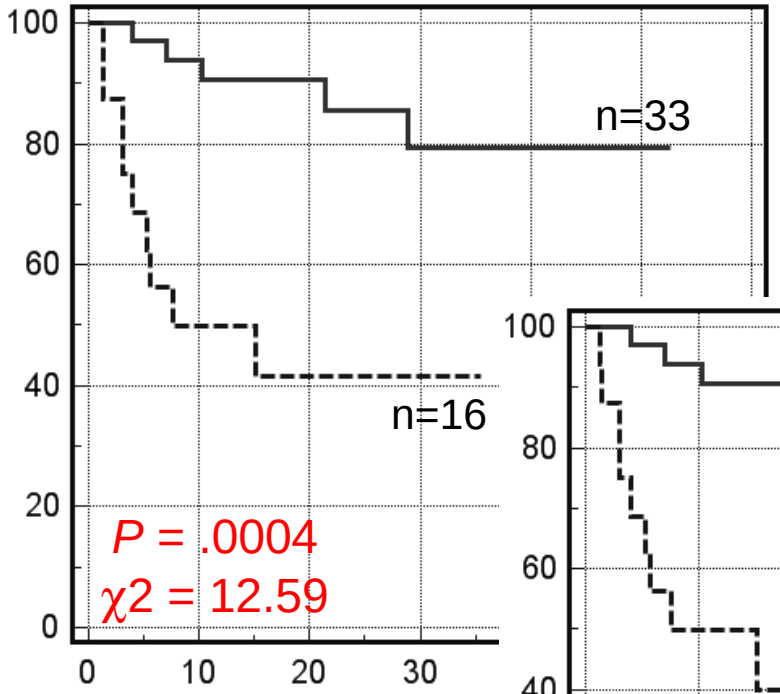
# of events = 14
Median f-u = 25 mo



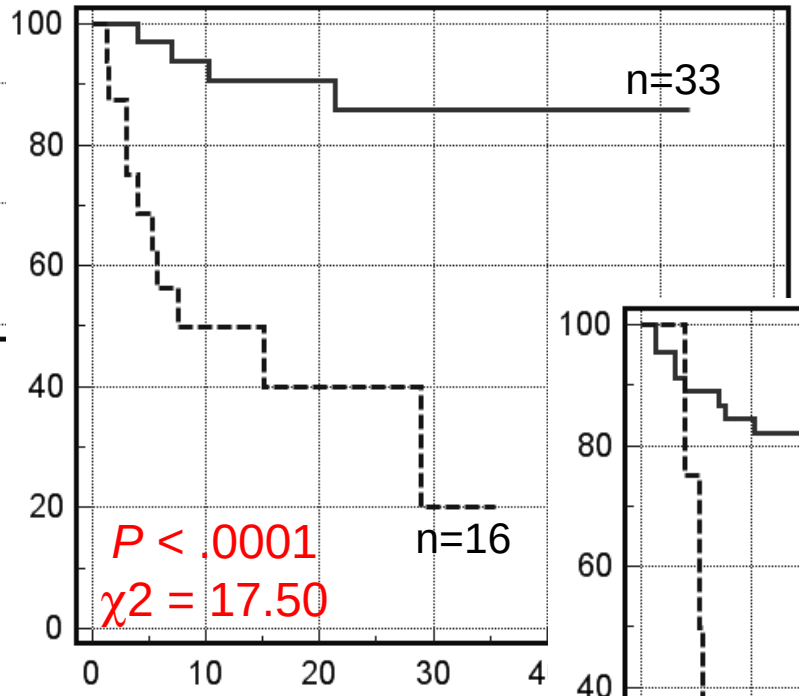
Cuneo :
2-y EFS : 87% vs. 32%

→ Reduction of false-positives
→ Cuneo's interpretation ++

5-point scale
(cut-off ≥5, >>liver)
Event-free survival

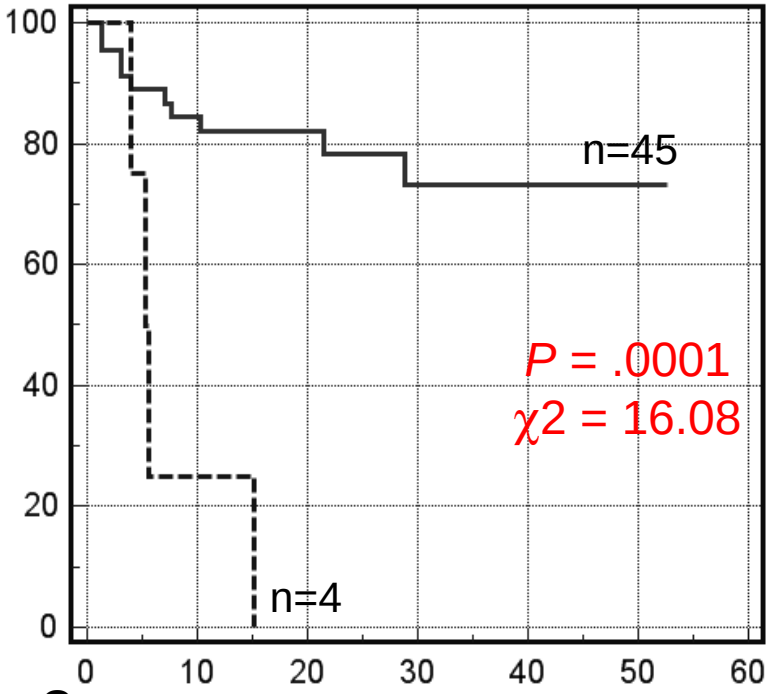


Créteil :
2-y EFS : 86% vs. 42%



Dijon :
2-y EFS : 86% vs. 40%

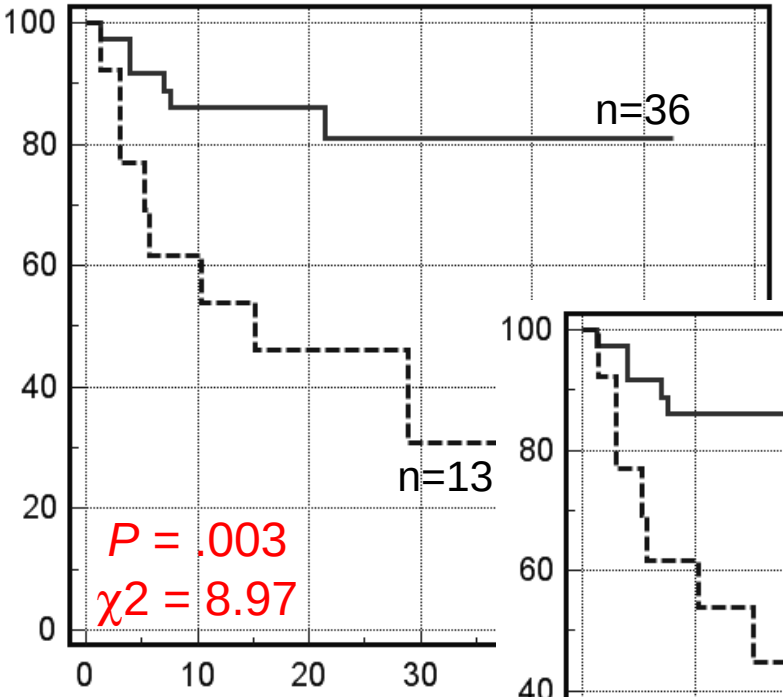
# of events = 14
Median f-u = 25 mo



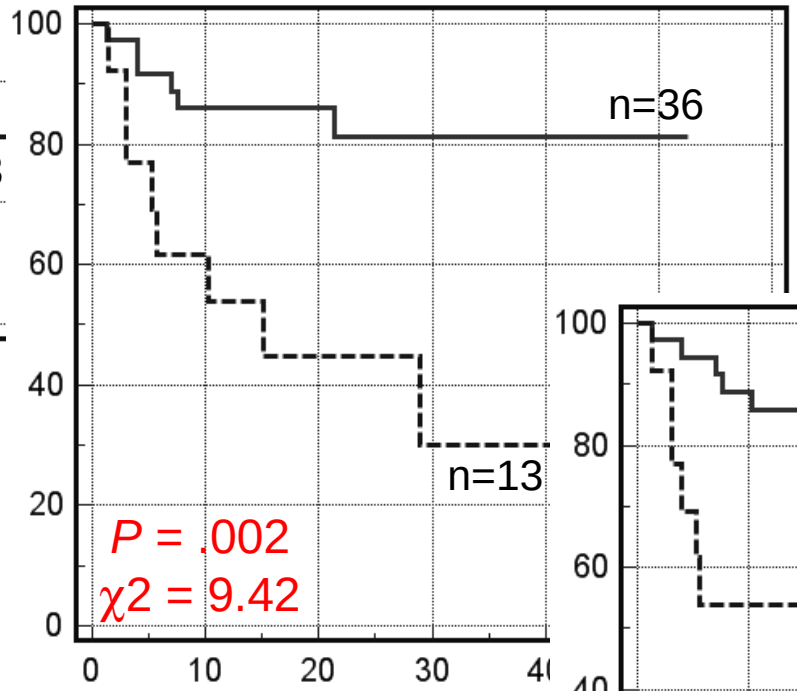
Cuneo :
2-y EFS : 78% vs. 0%

→ Créteil-Dijon's interpretations ++
→ Cuneo : generates false-negatives

Quantification
ΔSUV (cut-off >66%)
Event-free survival

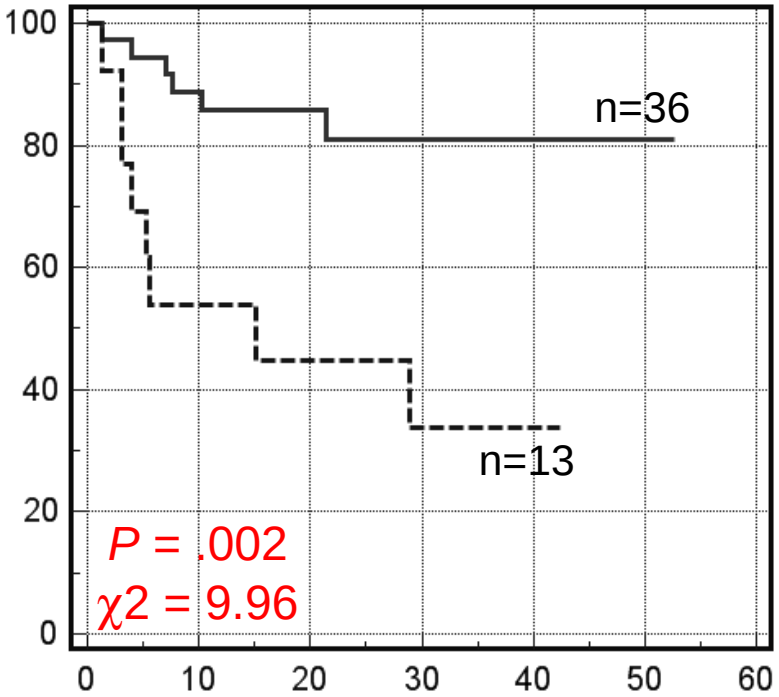


Créteil :
2-y EFS : 81% vs. 46%



Dijon :
2-y EFS : 81% vs. 45%

# of events = 14
Median f-u = 25 mo



Cuneo :
2-y EFS : 81% vs. 45%

→ Better agreement between observers

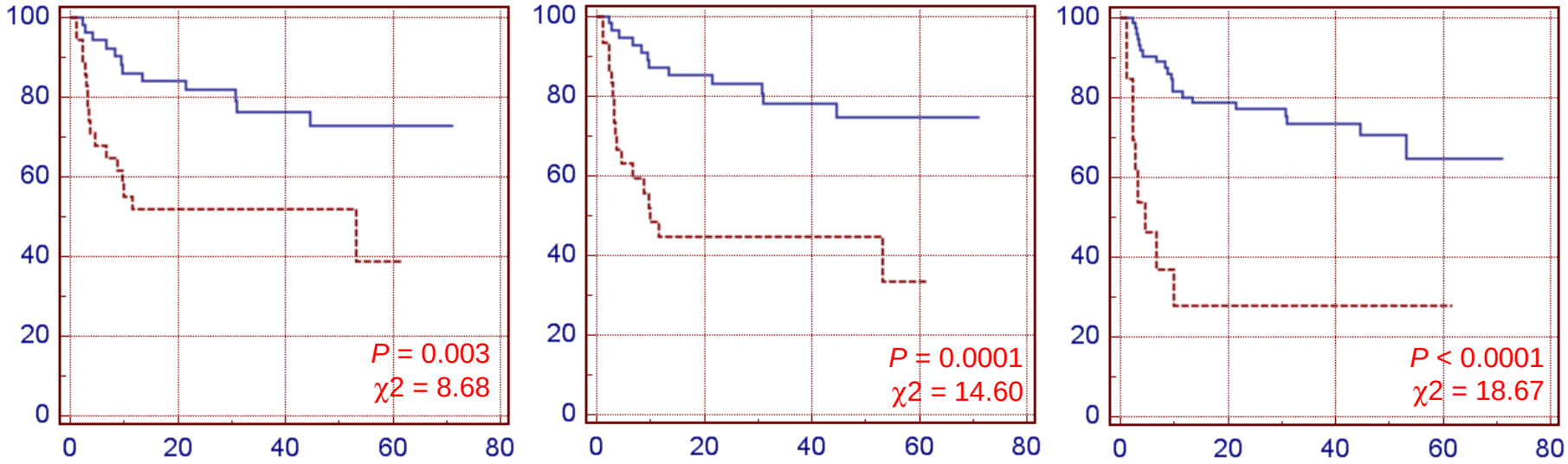
Conclusions

- 5PS: moderate to substantial agreement (κ 0.58-0.61)
- Ref. background must be high for interim PET/NHL
- Subjectivity $\rightarrow$ need for different opinions
  foci that were considered by Créteil/Dijon as “moderately increased above liver” (4) were considered “equal to liver” (3) by Cuneo
- Quantification may help the definition of scores 3-4
- ΔSUV is not observer-dependent for EFS prediction

125% liver

150% liver

200% liver



92 patients from Haioun, *Blood* 2005

[^{18}F]fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) in aggressive lymphoma: an early prognostic tool for predicting patient outcome

Corinne Haioun, Emmanuel Itti, Alain Rahmouni, Pauline Brice, Jean-Didier Rain, Karim Belhadj, Philippe Gaulard, Laurent Garderet, Eric Lepage, Felix Reyes, and Michel Meignan

Itti, Juweid, Haioun, et al. *SNM* 2010