Third international workshop on interim-PET in lymphoma

Under the auspices of GELA, FIL, SFMN, EANM, EHA, SIE

Menton (France), Palais de l'Europe,
September 26-27th, 2011

Organization Committee
M. Meignan (France), A.Gallamini (Italy), C.Haioun (France)

Scientific Committee
S.Barrington (UK), B.Cheson (USA), U.Duhrsen (Germany), A.Gallamini (Italy), C.Haioun (France), E.Itti (France), M. Juweid (USA), L.Kostakoglu (USA), A. Lister (UK), M. Meignan (France), A.Polliack (Israel), Th.Vander Borght (Belgium).

Report on IVS in HL

Alberto Biggi

Nucl. Med.Dept.
AO Santa Croce e Carle
CUNEO

Patient selection

400 patients enrolled

336 patients with PET/CT scans uploaded & quality controlled

261 patients with PET/CT scans approved & sent to review

Reason for PET scan exclusion

- Absence of CT images 22
- Absence of baseline PET 25
- Absence of interim PET 1
- CT slices missing 3
- PET slices missing 10
- Poor quality scans 6
- Miscellaneous 8

- **REVIEWERS**
- Sally Barrington – London – UK
- Alberto Biggi- Cuneo – I
- Michele Gregianin – Padova - I
- Martin Hutchings- Copenhagen – DK
- Lale Kostakoglu – New York – USA
- Michel Meignan – Paris – F

KEOSYS Medical Imaging

Review results acquired and statistical analysed

Roadmap for IVS

- Clinical data retrieval should be complete by the end of September 2009
- Images retrieval should end by February 2010
- Preliminary results of the review could be obtained before February 2010 provided we have enough images on the WEB site
- Preliminary data should be presented to the Menton meeting 8 – 10 April 2010

First meeting of PET reviewers
Cuneo July 3th, 2009

General rules for PET interpretation

- Visual assessment should be the goal in a retrospective multicenter study
- Results should be reported using the five-point scale
- Semi quantitative analysis could be used to aid visual interpretation for discordant cases

PET reporting
- the Deauville criteria

Score 1: no uptake
Score 2: uptake ≤ mediastinum

Score 3: uptake > mediastinum but ≤ liver

Score 4: moderately ↑uptake > liver
Score 5: markedly ↑uptake > liver AND/or new sites of disease

Semi quantitative analysis
(for discordant cases only)

The Max SUV in the region of residual uptake measured in a ROI placed on the axial slice with highest intensity should be compared with the max SUV in a large ROI representative of uptake in the whole structure to estimate maximum uptake in mediastinal blood pool and the liver.

Rules for interpretation

***Score 5 → new lesions*.**

- New lesion at a different site from disease → score 1.
- New lesion at a different site from disease with clear evidence of PD at other sites → score 5
- New lesion at a different site probably NOT lymphoma but request for clinical information

Diffuse uptake in spleen or marrow on the interim scan is most likely due to chemotherapy and should be scored as no disease especially if growth factors have been used (even if focal uptake is present at baseline)

Focal uptake in marrow can be scored as no disease if there is reduced uptake at sites where there was disease on baseline (due to marrow ablation) and increased uptake at sites with no disease at baseline (due to chemotherapy effect). This means that uptake on the interim scan may be like a “mirror” of the uptake on the baseline scan

Symmetrical tonsillar uptake (on baseline or interim scan) is most likely to represent a normal variant or inflammatory changes. Asymmetric uptake on the interim scan should only be regarded as disease if there was clear evidence of tonsillar involvement at baseline but beware! as tonsillar involvement is very much less likely in HL than NHL.

After Menton Meeting (july 7th , 2010)

Rules for agreement

1. Agreement levels will be based on an analysis which is binary i.e. 1,2 vs 3,4,5 liver threshold) and 1,2,3, vs 4,5 (mediastinal threshold) for negative and positive respectively
2. cases where < 4 reviewers agree whether the scan is “negative” or “positive” must be reviewed to determine if consensus can be reached; agreement levels will be reported for independent reading only NOT for consensus reading. Issues where there are problems with reaching consensus should be identified by this process and it would be useful identify these types of cases for the paper, which could almost consitute a “manual “ for PET reporting in interim lymphoma.
3. It may be necessary review those cases scored 5 in initial 50 cases (score 5 was incorrectly published in article in Leukemia and lymhpoma as markedly increased uptake compared with liver AND new lesions rather than markedly increased uptake compared with liver AND/OR new lesions which is correct).

Second meeting of PET reviewers
Cuneo November 27th, 2010

Final meeting in London
february 2011

“To review scans where agreement is not reached by at least 4 reviewers”

The best cut-off of the score to be used for interim PET evaluation in the ABVD-BEACOPP protocol

Treat with BEACOPP only patients with a positive PET

The “best” is the score with the highest TP/FP ratio

Which is the score of the patients with 6 different reviewer and 6 different score?

Pat.	Rev. 1	Rev. 2	Rev. 3	Rev. 4	Rev. 5	Rev. 6	Final Score
1	2	4	2	3	1	3	**UD**
2	2	3	2	3	1	3	**UD**
3	2	1	**3**	**3**	**3**	**3**	**3 in 4 rev**
4	2	**3**	**3**	**3**	**3**	**3**	**3 in 5 rev**
5	**5**	**5**	**5**	**5**	**5**	**4**	**5 in 5 rev**
6	**5**	**3**	**4**	**4**	**3**	**5**	**UD**
7	**1**	**1**	**2**	**3**	**1**	**1**	**1 in 4 rev**
8	**4**	**4**	**4**	**4**	**4**	**4**	**4 in 6 rev**
9	**2**	**2**	**3**	**1**	**1**	**1**	**UD**
10	**3**	**3**	**3**	**3**	**3**	**3**	**3 in 6 rev**
…….							

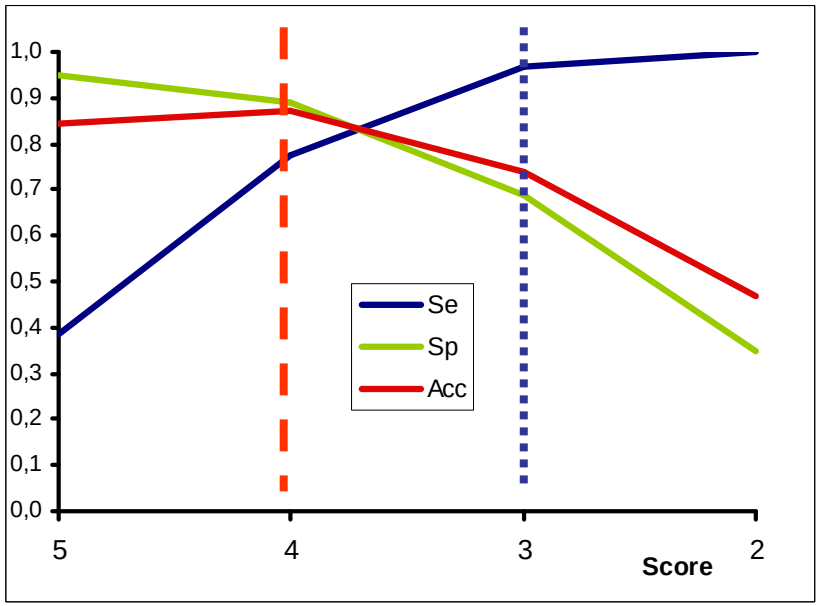
The final score for each patient is the score defined by the majority of the reviewers (4)

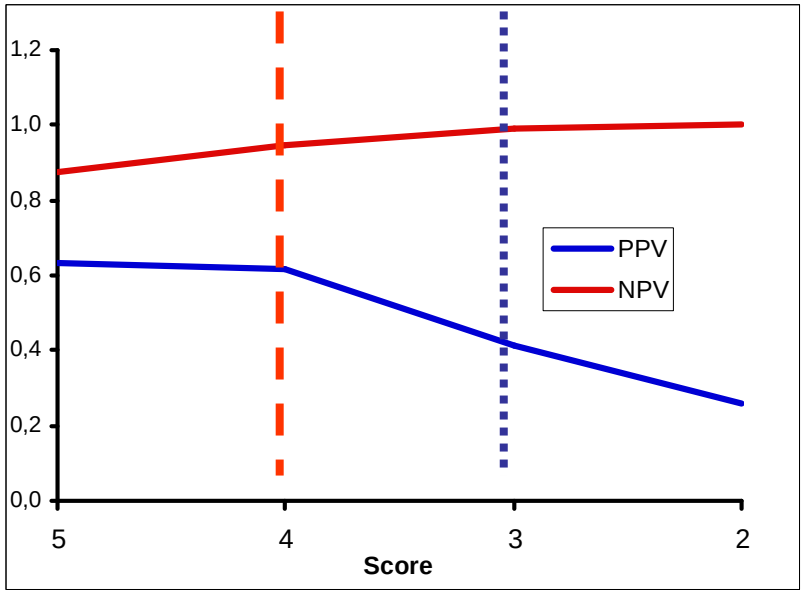
Score agreement between reviewers

Score 1	7 **+ 6 ***	17 **+ 2 ***	21	**53 pts**
Score 2	6	15	21	**42 pts**
Score 3	3	15 **+ 3 ***	12	**33 pts**
Score 4	11 **+ 3 ***	4	4	**22 pts**
Score 5	10	5 **+ 1***	3	**19 pts**
	46 pt	**62 pt**	**61 pt**	**169 pts**
	6 reviewers	5 reviewers	4 reviewers	

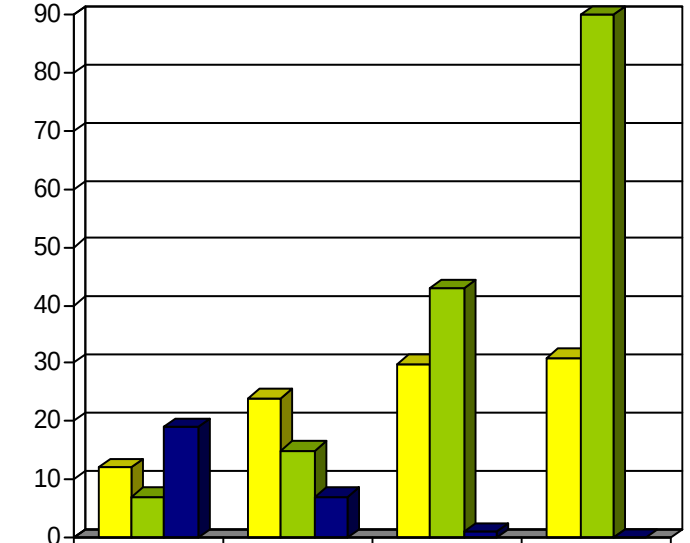
169/261 = 65%

* after consensus in London



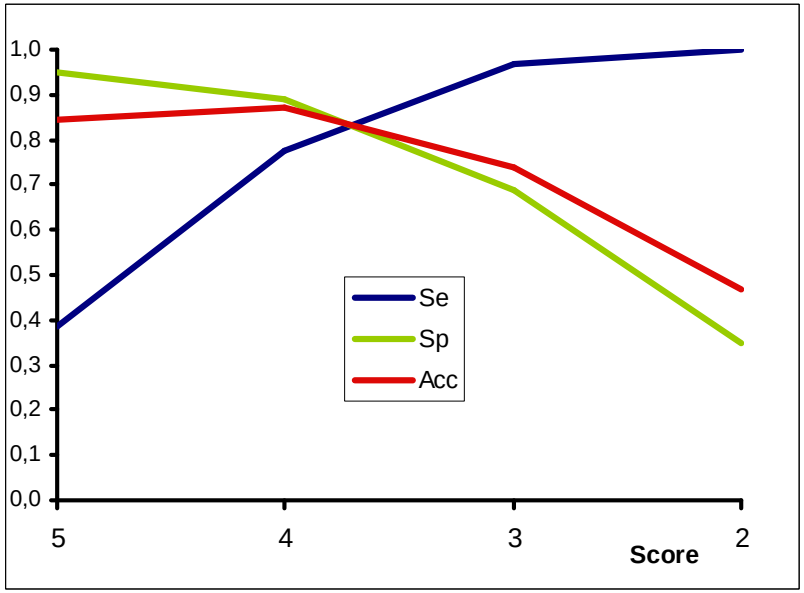


169 pts

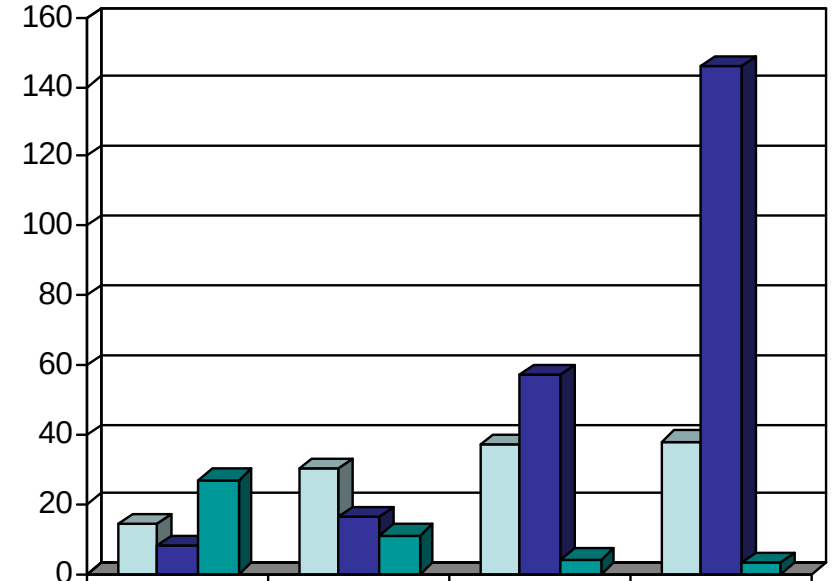


	Score 5	Score 4	Score 3	Score 2
TP	12	24	30	31
FP	7	15	43	90
FN	19	7	1	0

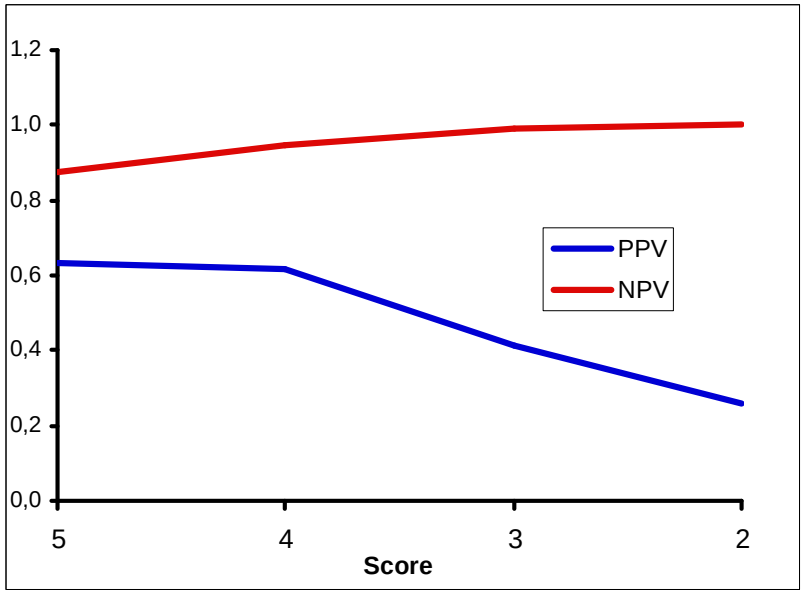
Sc. 5: TP/FP → 12/ 7 → 1.7
Sc. 4: TP/FP → 24/15 → 1.6
Sc. 3: TP/FP → 30/43 → 0.7



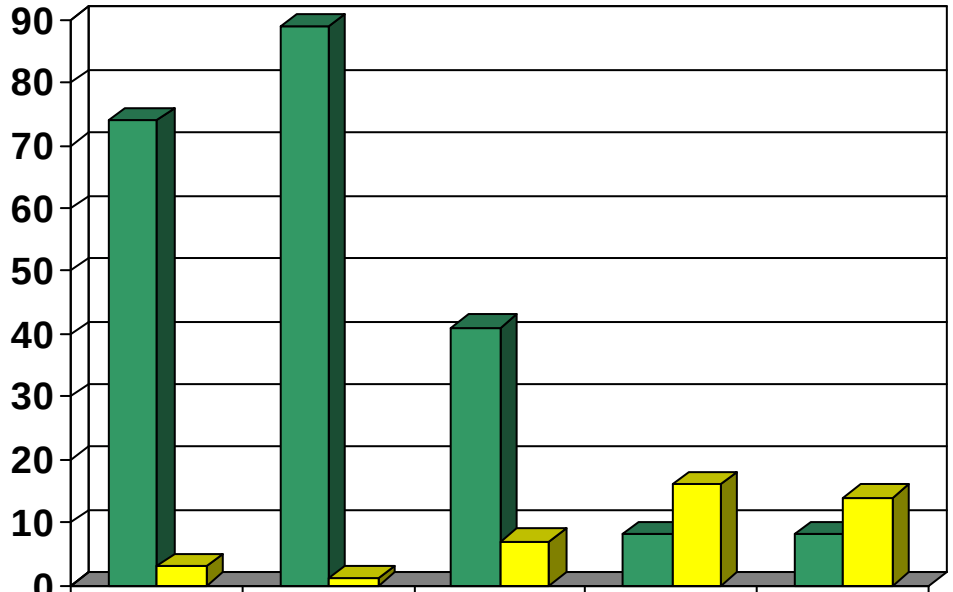
261 pt



	Score 5	Score 4	Score 3	Score 2
TP	14	30	37	38
FP	8	16	57	146
FN	27	11	4	3



Score and outcome



	Score 1	Score 2	Score 3	Score 4	Score 5
Responders	74	89	41	8	8
Non responders	3	1	7	16	14

Summary of methodological aspects

Agreement between reviewers for PET+ (score 4,5) vs PET- (score 1,2,3)

213/261	82%	full agreement	6 vs 0
29 /261	11%	minor discordance	5 vs 1

-to be discussed in the future?

11 /261	4%	major discordance	4 vs 2
8 /261	3%	true discordance	3 vs 3

Analysed in joint session

True discordance interpretation of :
Marrow (2)
Gut (1)
Brown fat/vascular (2)
Parotid adenoma (1)
Missed disease (2)

Agreement between reviewers

Binary concordance:

-ve vs. +ve

1,2,3 vs. 4,5

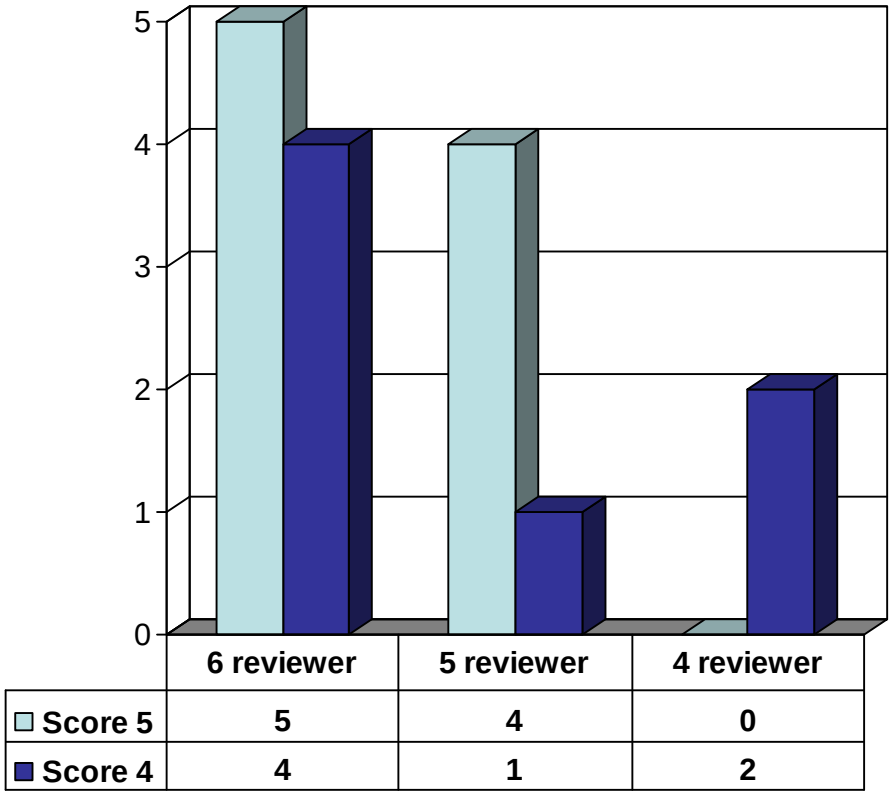
Cohen's K	
< 0.2	poor
0.21–0.40	fair
0.41–0.60	moderate
0.61–0.80	good
>0.81	very good

Cohen's Kappa:

	Mean						
	0.75	1	0.73	0.77	0.78	0.75	0.73
	0.73	0.73	1	0.75	0.75	0.70	0.71
	0.78	0.77	0.75	1	0.83	0.77	0.77
	0.81	0.78	0.75	0.83	1	0.84	0.84
	0.77	0.75	0.70	0.77	0.84	1	0.78
	0.77	0.73	0.71	0.77	0.84	0.78	1

Alpha	D-obs	D-exp	N	R	Reviewers
0.758	118.400	489.908	261	6	Hutchings, Meignan, Barrington, Kostakoglu, Biggi, Gregianin

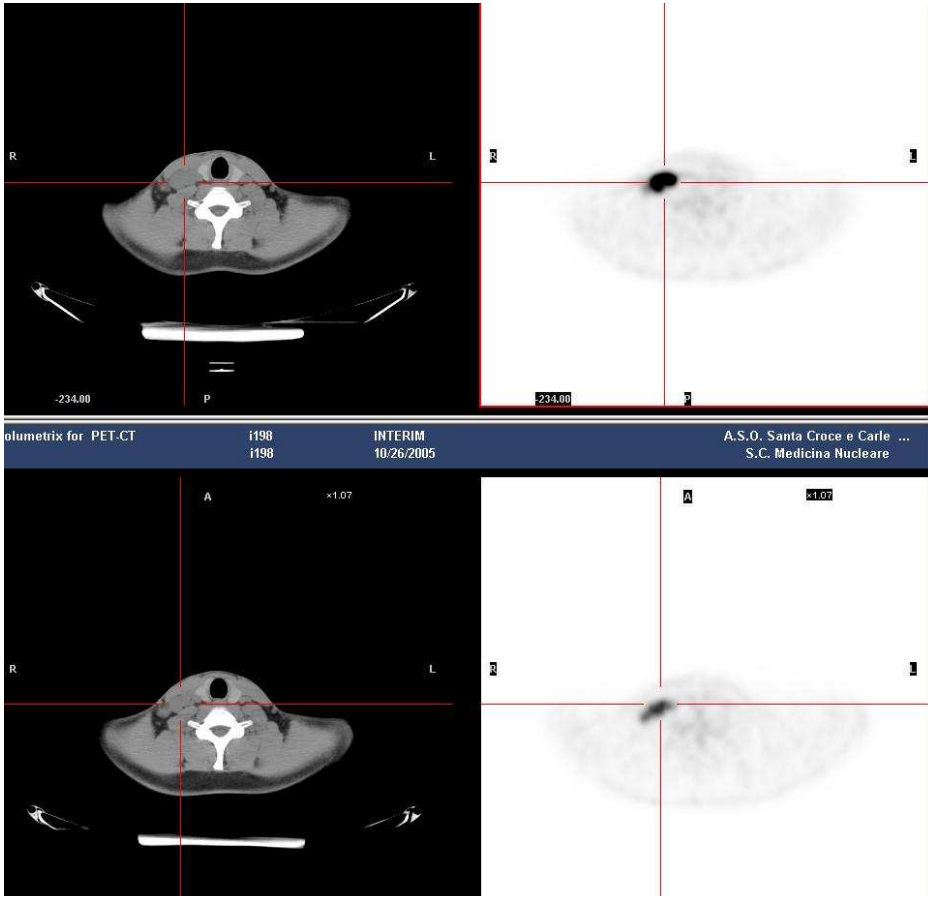
False positive results



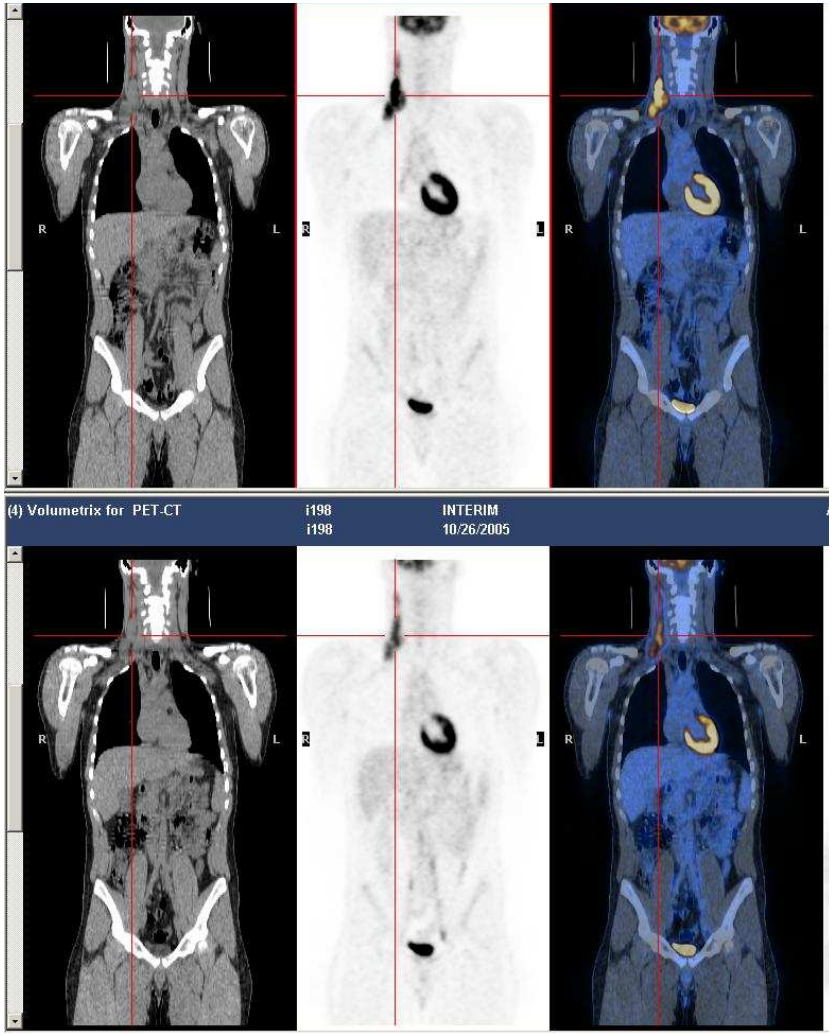
	6 reviewer	5 reviewer	4 reviewer
Score 5	5	4	0
Score 4	4	1	2

16 pts

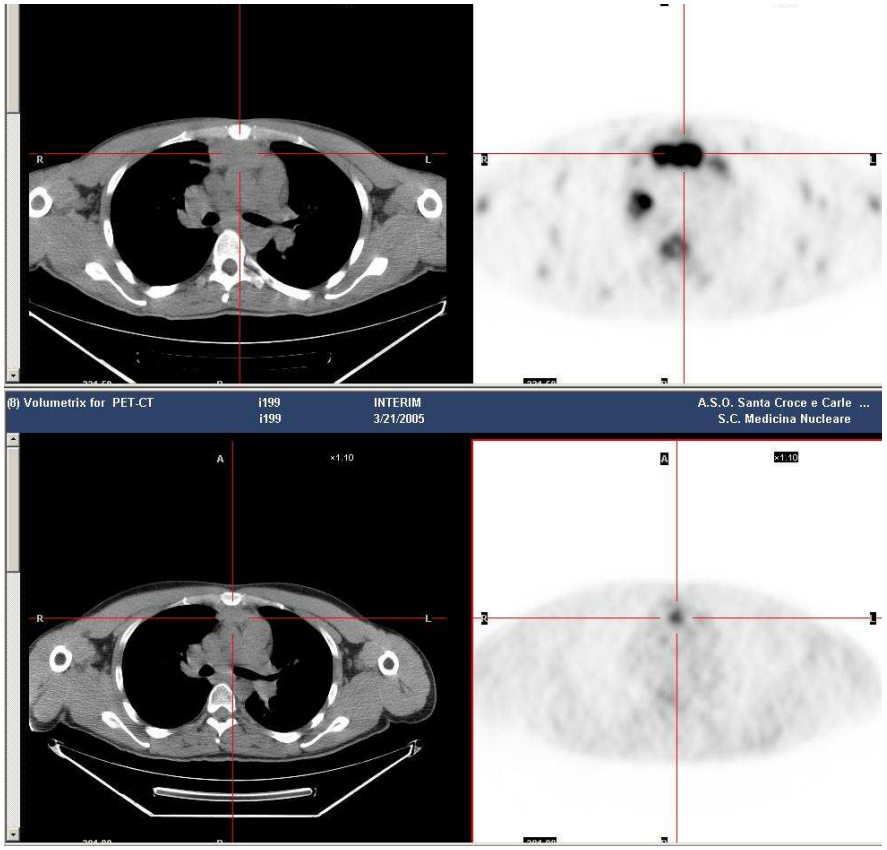
- 10 mediastinum
- 2 laterocerv.
- 1 right pulmonary hilum
- 1 axilla
- 1 lung
- 1 bone



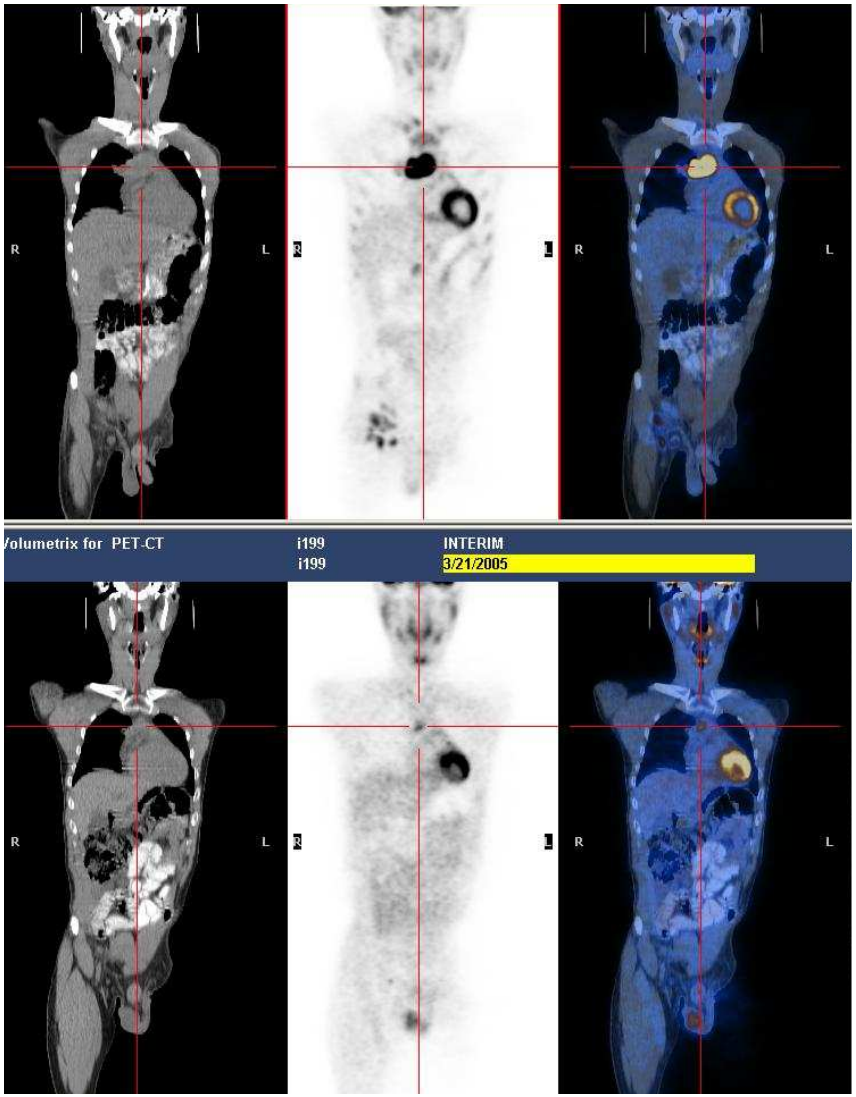
Case 198
♂ 16 y
HD sclerondular
Stage IIIA
CR - FFS 60 months



Case 198
Score 5 for 6/6 reviewers
SUVMax lesion 12.1 → 9.6
SUVMax liver 2.0 → 2.6

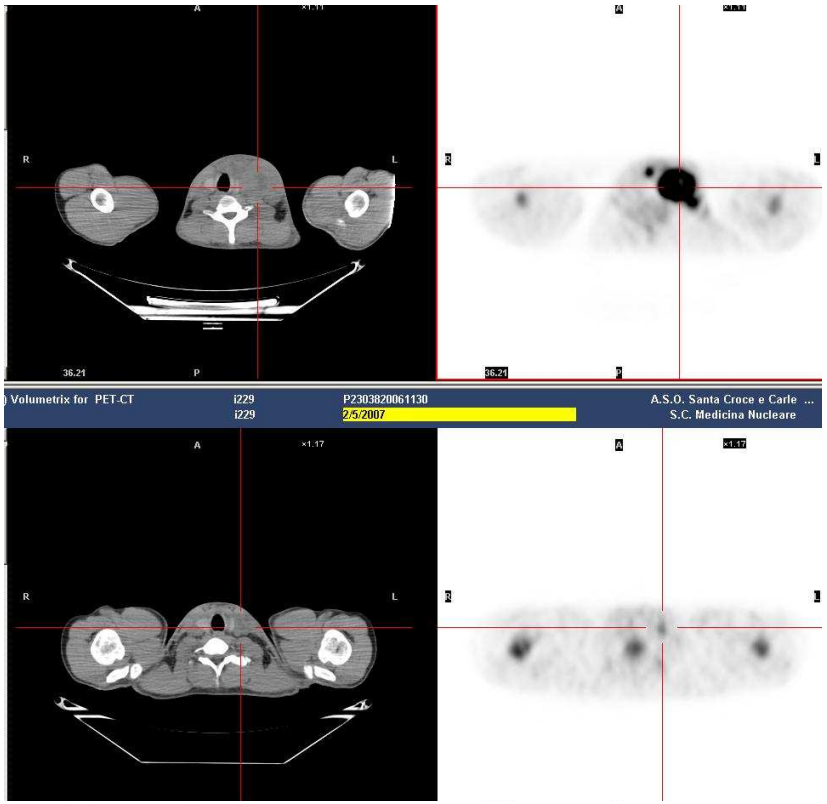


Case 199
♂ - 23 y
HD scleronodular
Stage IVB
CR - FFS 72 months

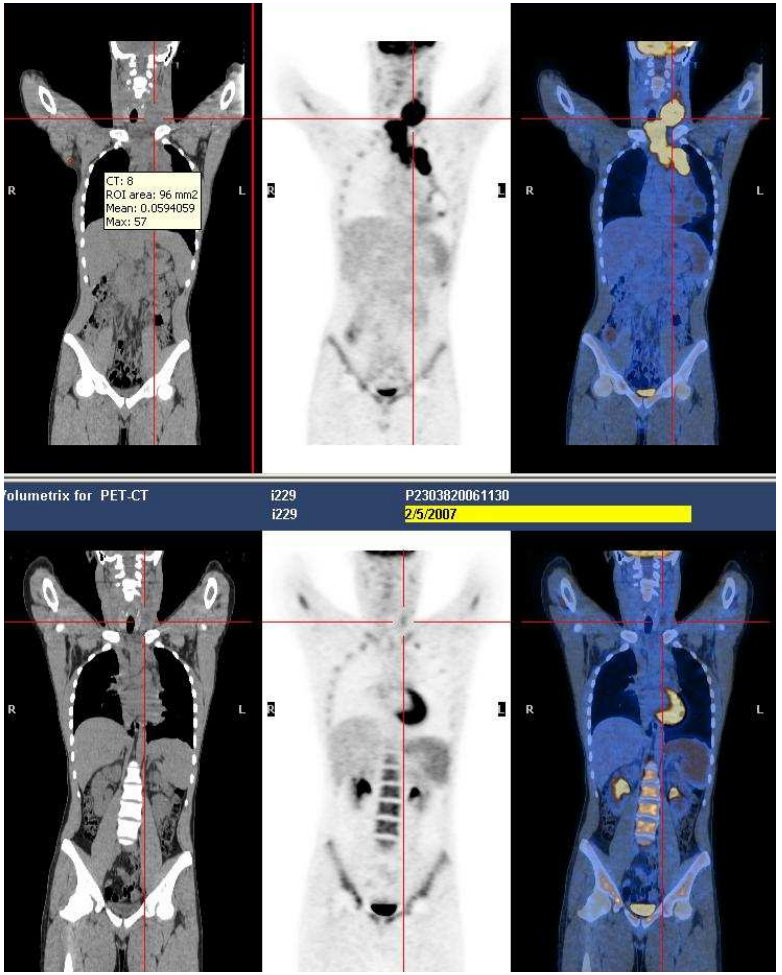


Case 199
Score 5 for 6/6 reviewers
SUVMax lesion 17.6 → 5.2
SUVMax liver 2.0 → 3.2

Consensus reevaluation

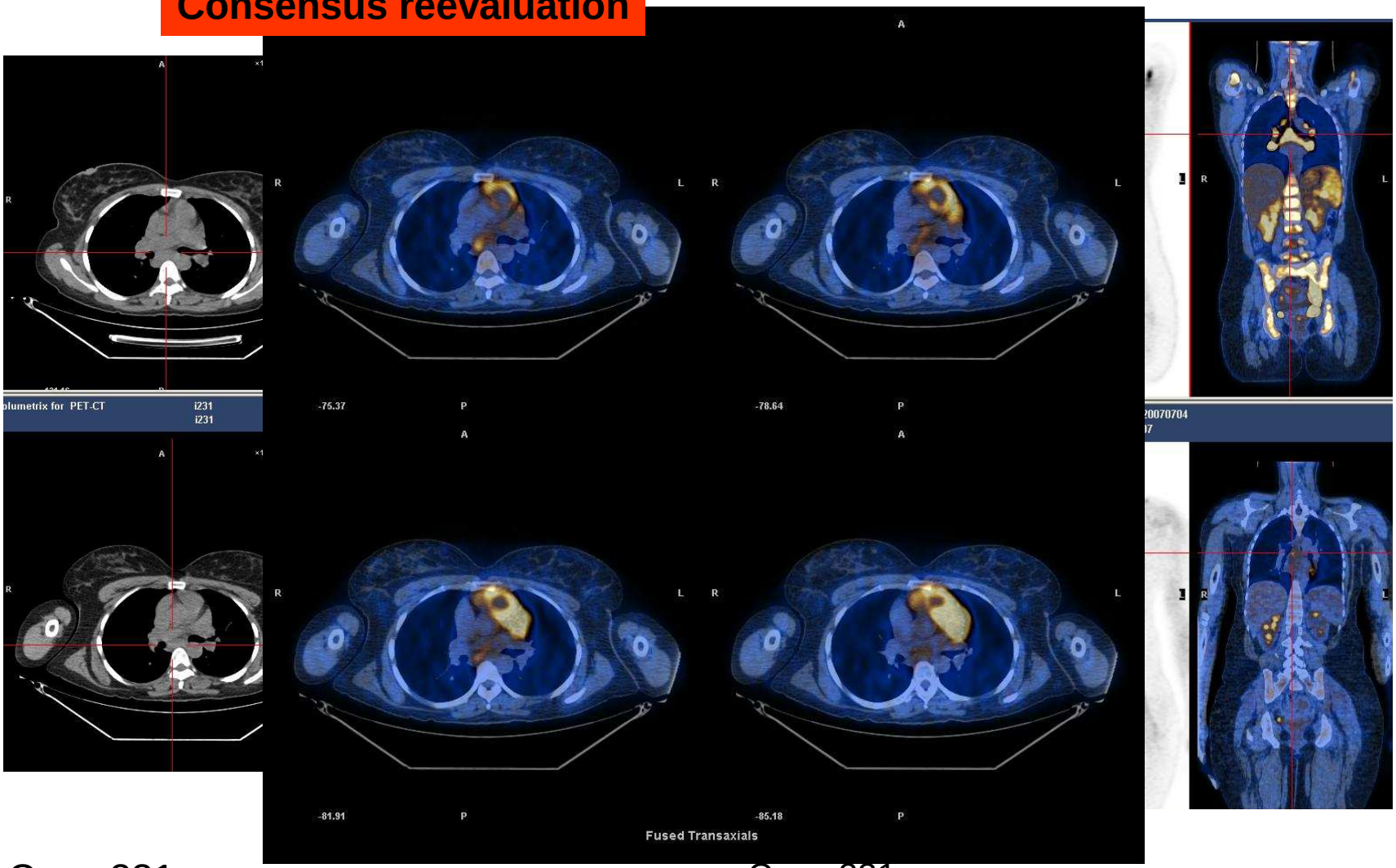


Case 229
♂ - 20 y
HD scleronodular
Stage IIB
CR - FFS 35 months



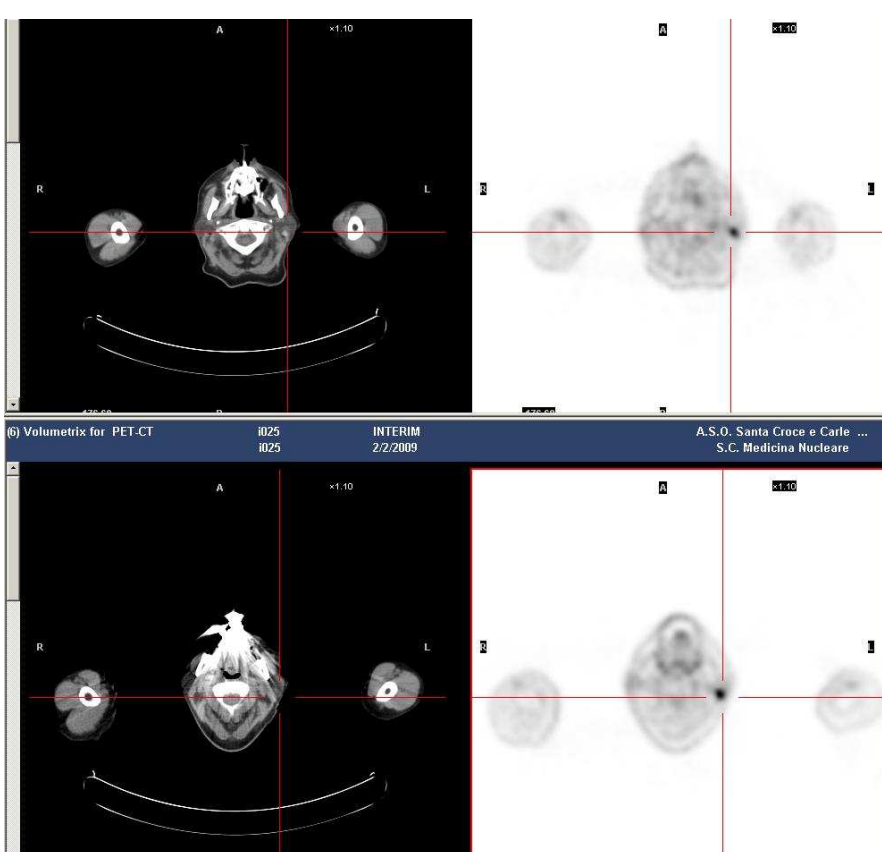
Case 229
Before consensus: 2 – 1 – 4 – 4 – 1 - 4
After consensus: 4 – 4 – 4 – 4 – 4 - 4
SUVMax lesion 15.1 → 2.5
SUVMax liver 2.3 → 1.8

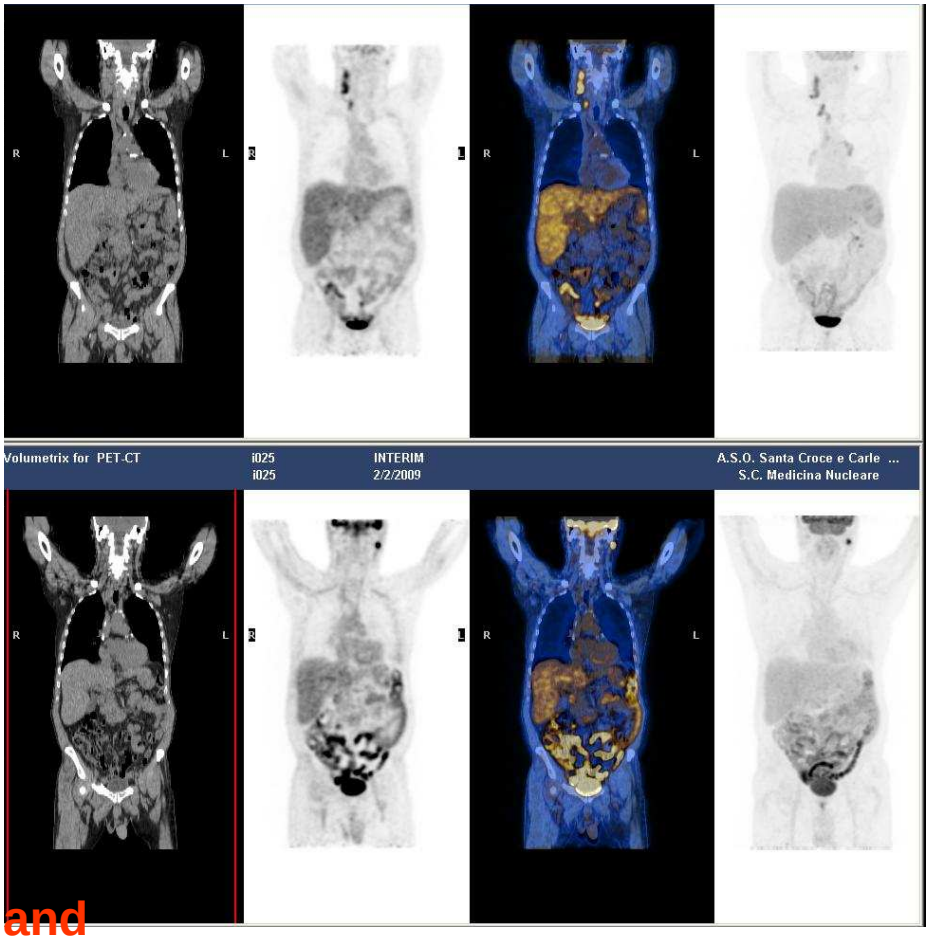
Consensus reevaluation



Case 231
♀ - 35 y
HD mixed cellularity
Stage IVB
CR - FFS 43 months

Case 231
Before consensus: 1 – 2 – 4 – 5 – 4 - 1
After consensus: 4 – 4 – 4 – 5 – 4 - 1
SUVMax lesion 12.0 → 4.7
SUVMax liver 4.7 → 2.6



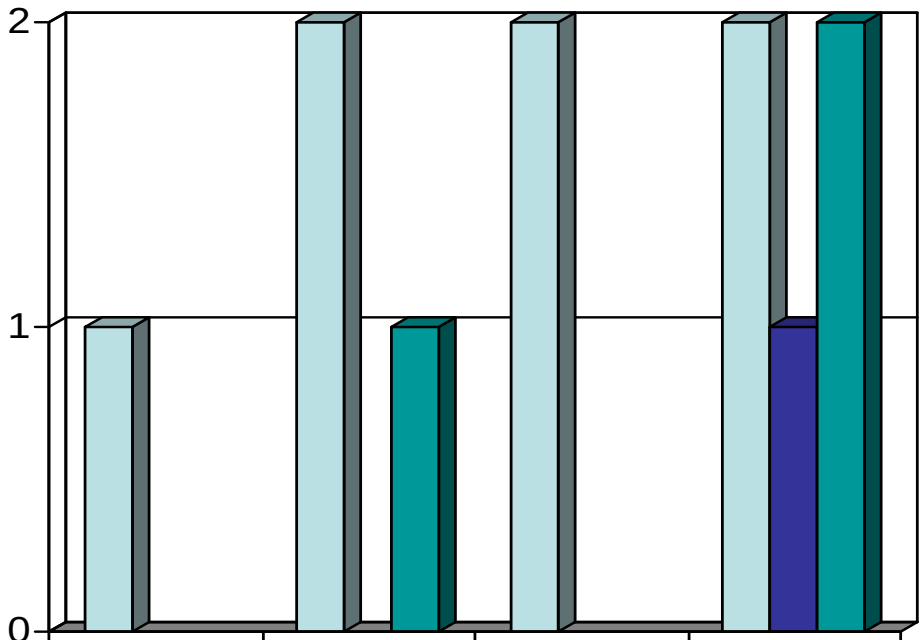


Warthin tumor of the left parothid gland

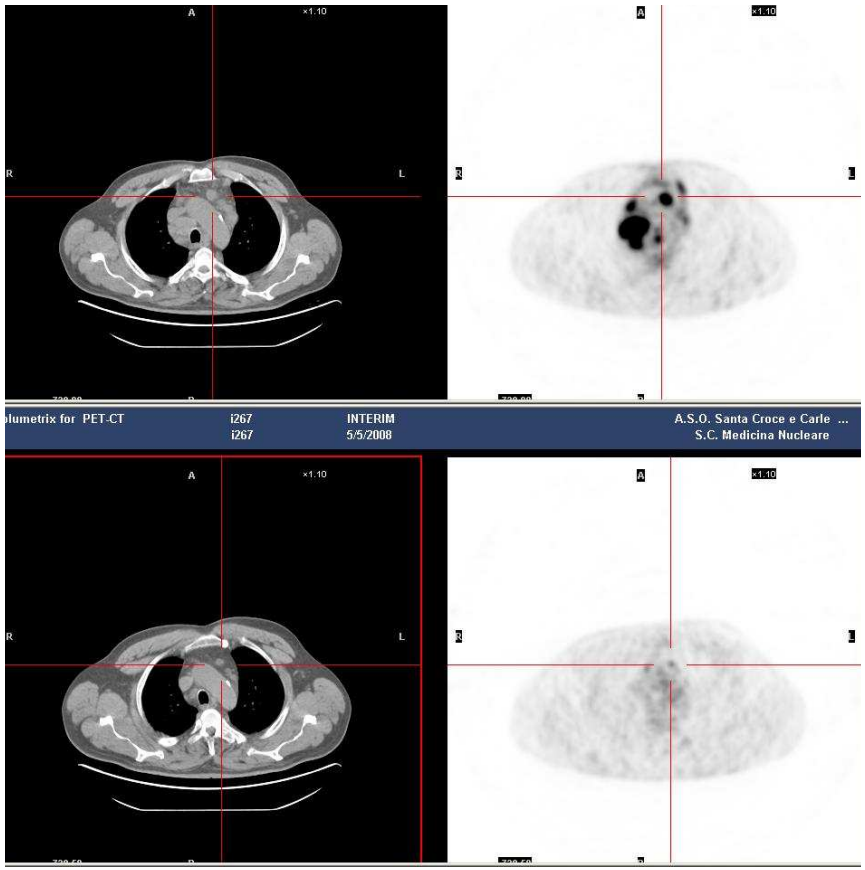
Case 25
♂ - 63 y
HD scleronodular
Stage III B
CR - FFS 26 months

Case 25
Before consensus: 5 – 4 – 2 – 5 – 2 - 1
After consensus: 2 – 1 – 1 – 1 – 2 - 1
SUVMax lesion 12.0 → 4.7
SUVMax liver 4.5 → 2.8

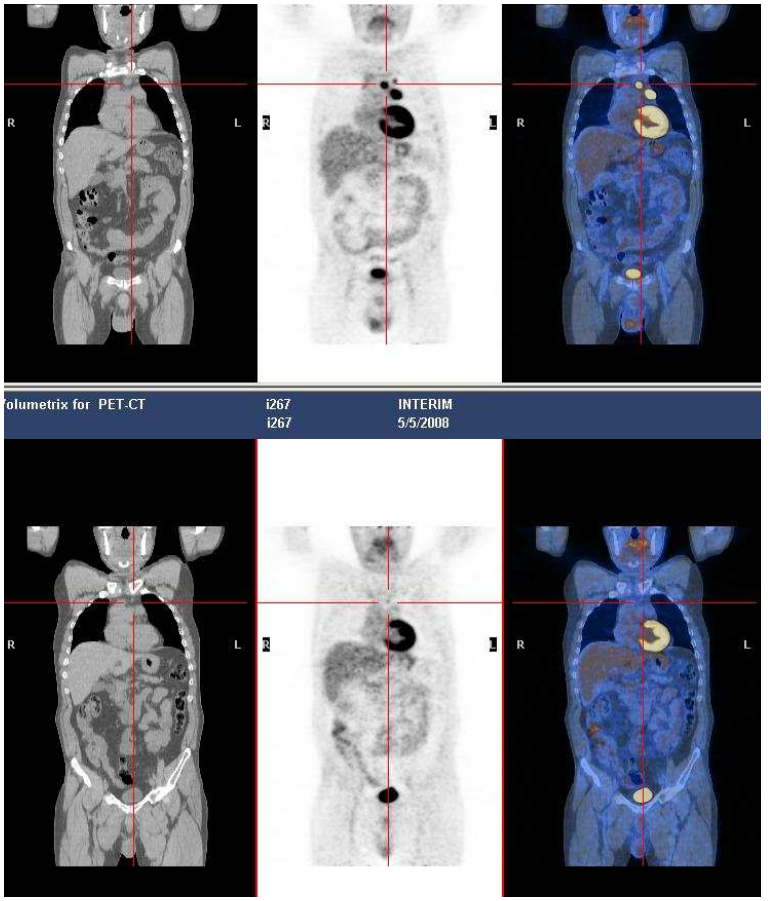
False negative results



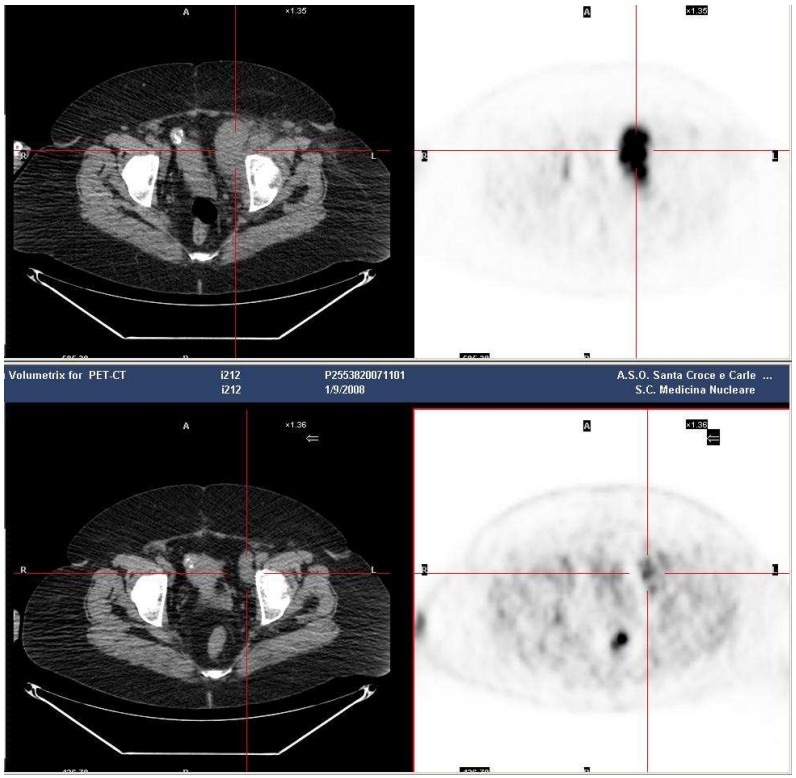
	6 reviewer	5 reviewer	4 reviewer	3 reviewer
Score 3	1	2	2	2
Score 2				1
Score 1		1		2



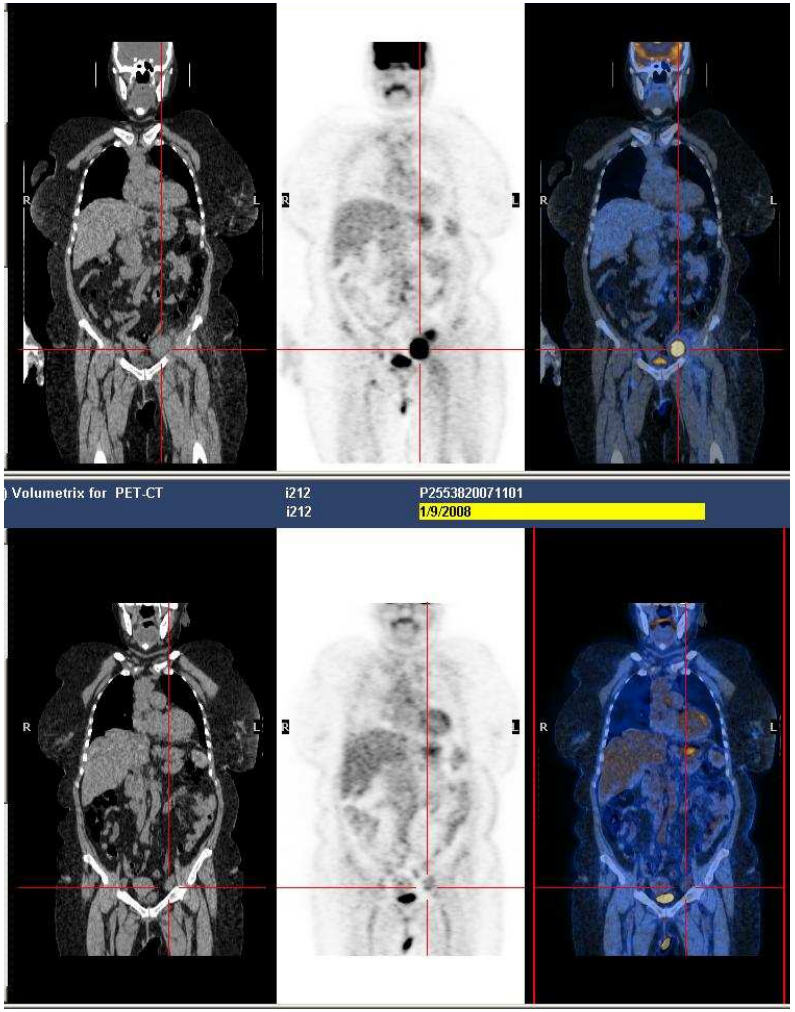
Case 267
♂ - 56 y
HD scleronodular
Stage III B
PD - FFS 10 months
Alive after II° line therapy (1.2011)



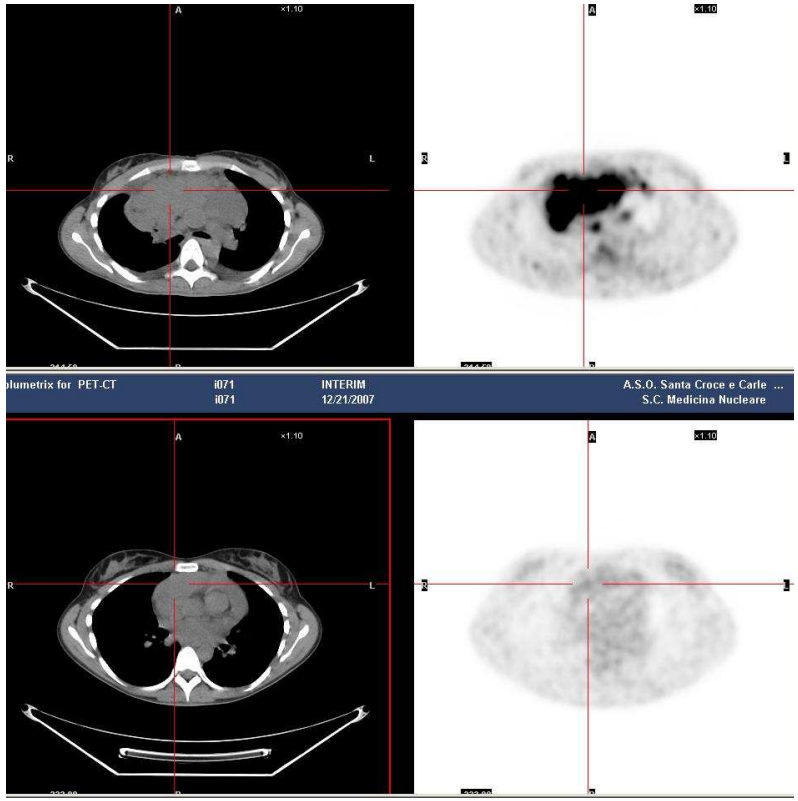
Case 267
Score 3 for 4 reviewer; score 2 for 2
SUVMax lesion 11.6 → 2.5
SUVMax liver 3.7 → 4.1



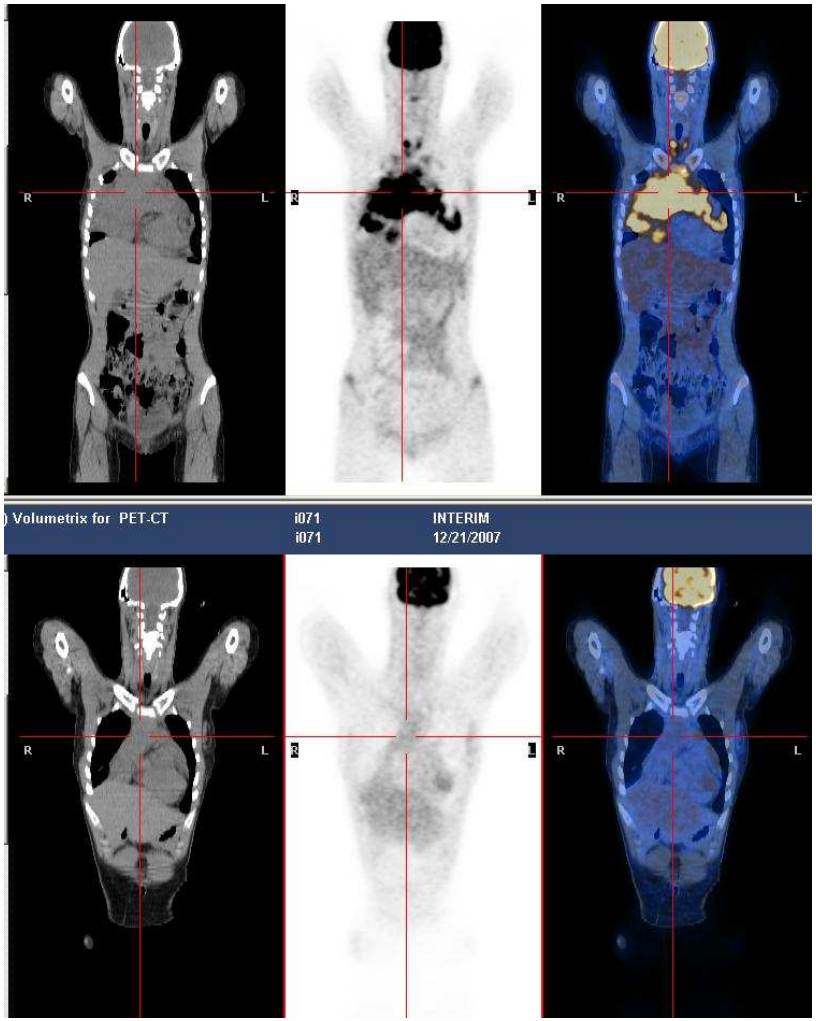
Case 212
♂ - 69 y
HD scleronodular
Stage IIIA
PD - FFS 8 months
CR 14 months after II° line therapy



Case 212
Score 3 for 5 reviewer; score 1 for 1
SUVMax lesion 23.2 → 3.6
SUVMax liver 3.8 → 3.8

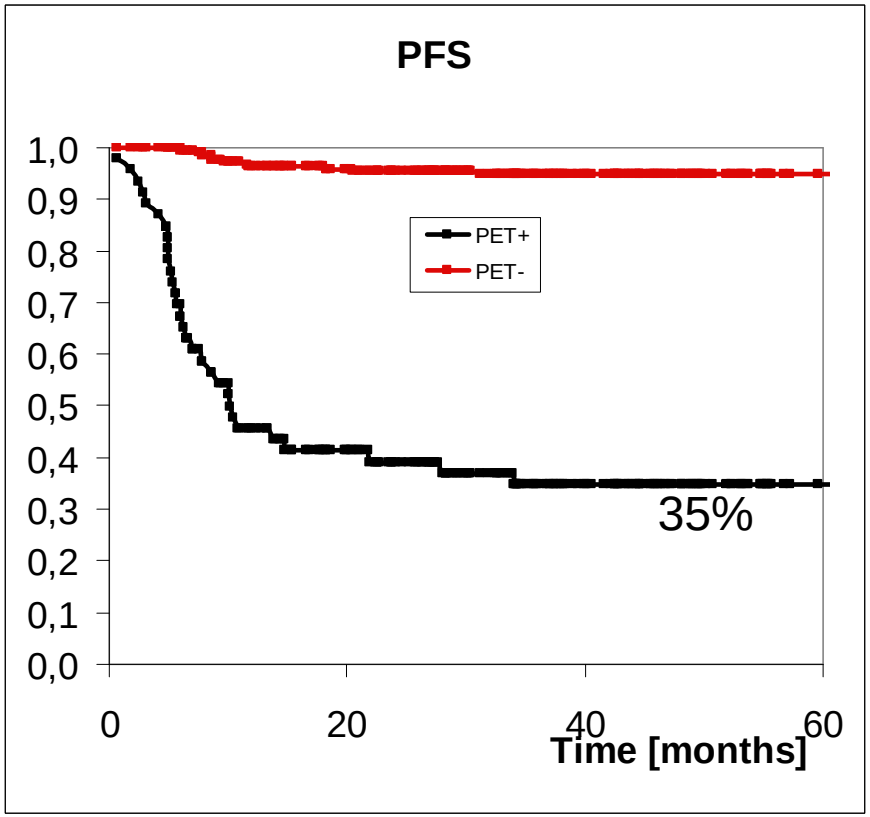


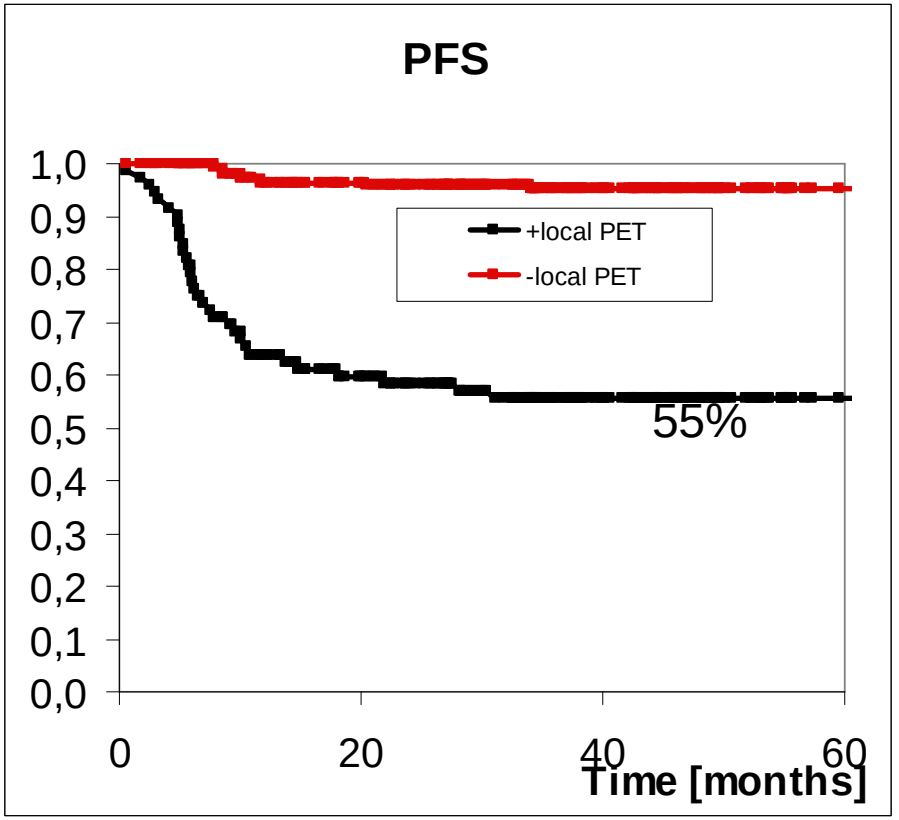
Case 71
♀ - 24 y
HD scleronodular
Stage IIA adverse prognostic factor
PD - DFS 12 months
CR after II° line therapy (2.2011)



Case 71
Score 3 for 6 reviewer
SUVMax lesion 13.6 → 3.0
SUVMax liver 2.8 → 2.6

Review panel vs local center






Thank’s for your attention