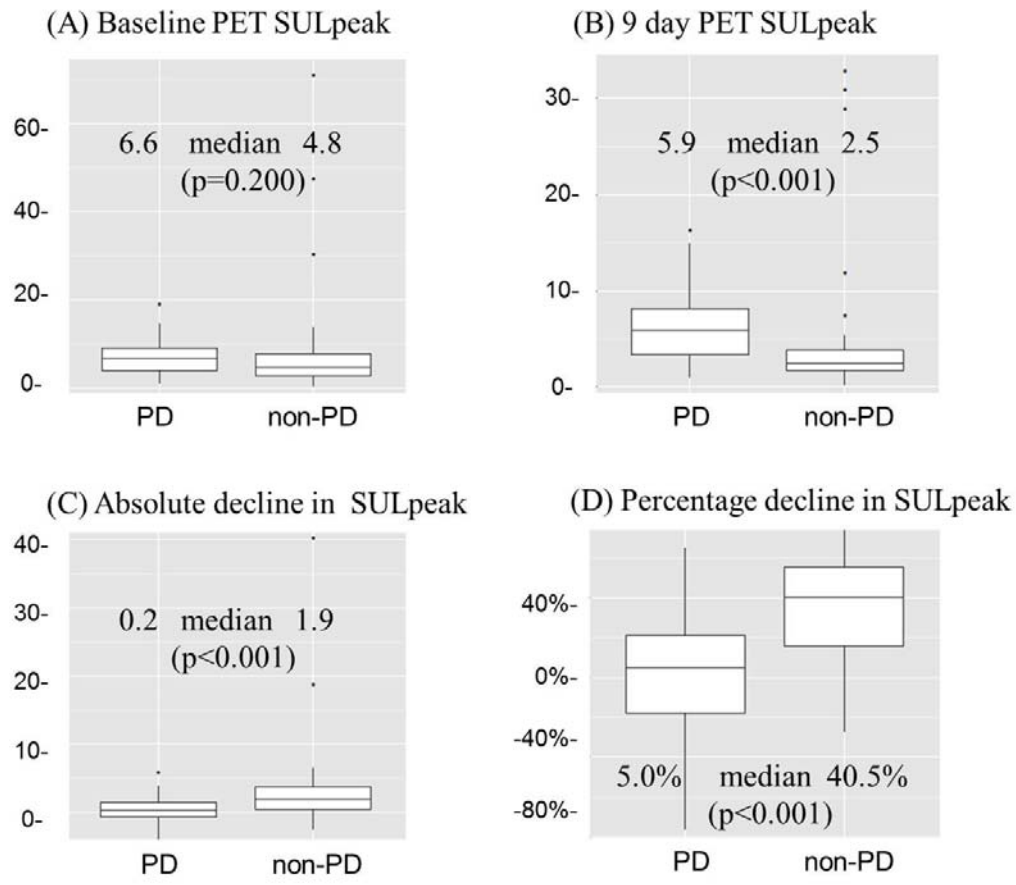
Supplemental Table 1. The details of the technical variations are shown. Of note is that despite technical variance being common, the PET method was predictive of outcome.

Technical deviation					Number of cases (%)	
Injected FDG dose					21 (22.5%)	
Time from FDG injection to image acquisition					55 (59.1%)	
Different scanner systems between baseline and follow-up studies					11 (11.8%)	
Liver FDG uptake variation between baseline and follow-up studies					23 (24.7%)	
Lesion at baseline below the PERCIST threshold for measurability					16 (17.2%)	
			mean	standard deviation	minimum	maximum
Baseline	FDG dose	mCi	6.82	2.86	1.63	17.60
		mCi/kg	0.10	0.05	0.04	0.29
	uptake time	Minutes	74.70	18.78	36	151
	weight	Kg	68.73	23.45	23.0	170.1
Follow-up	FDG dose	mCi	6.37	2.23	2.27	12.60
		mCi/kg	0.10	0.04	0.04	0.23
	uptake time	Minutes	75.37	20.78	33	167
	weight	kg	68.68	23.62	23.0	169.4

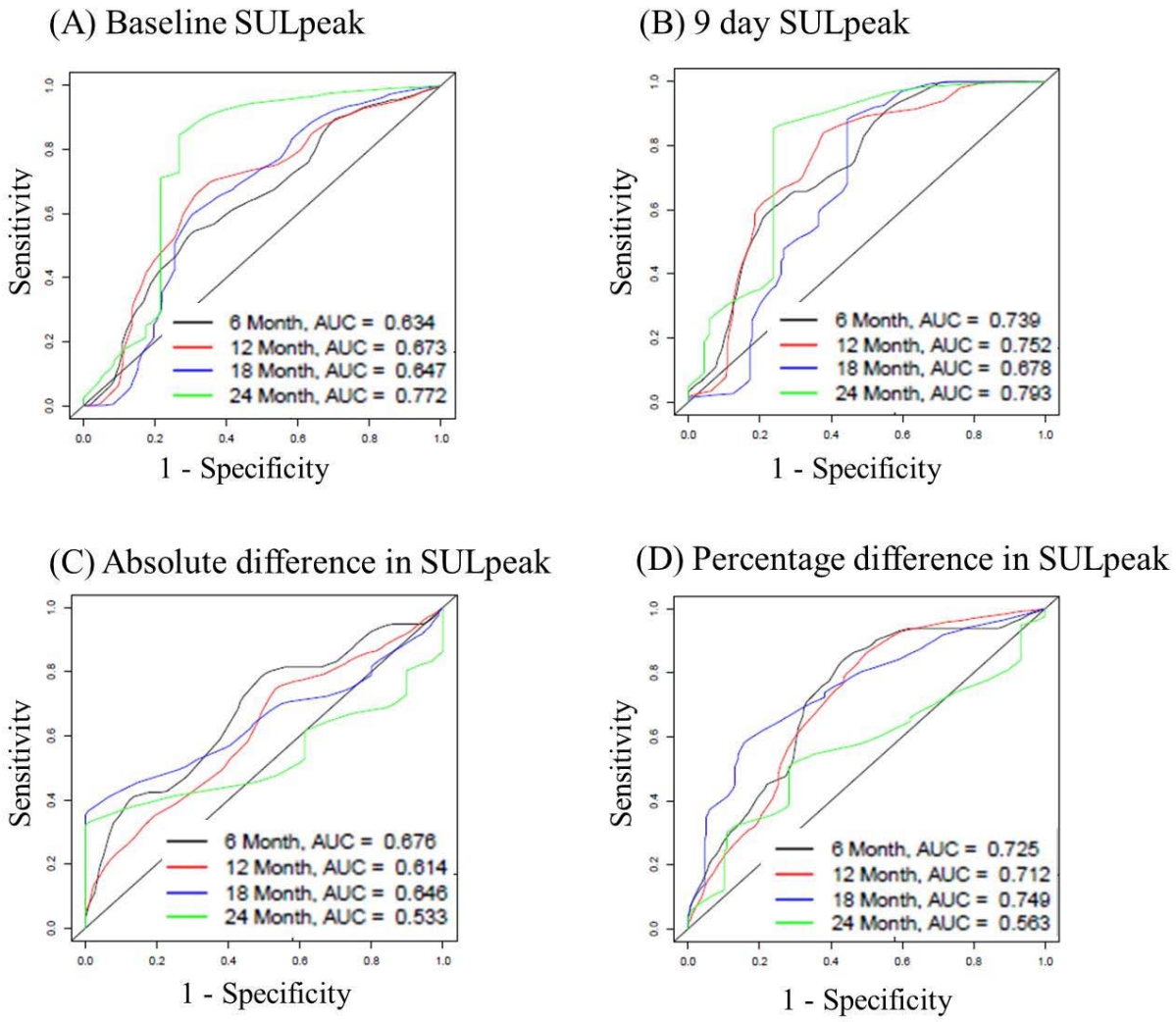
Supplemental Table 2.

Number of patients		Week 12 FDG PET response			
		CMR	PMR	SMD	PMD
Day 9 FDG PET response	CMR	1	1		
	PMR	1	9	4	7
	SMD		1	1	1
	PMD				

Supplemental Figure 1. Box plots of SULpeak measurements at baseline and 9 days after therapy, and the status of the clinical response at 6 weeks after therapy including the outlier values.

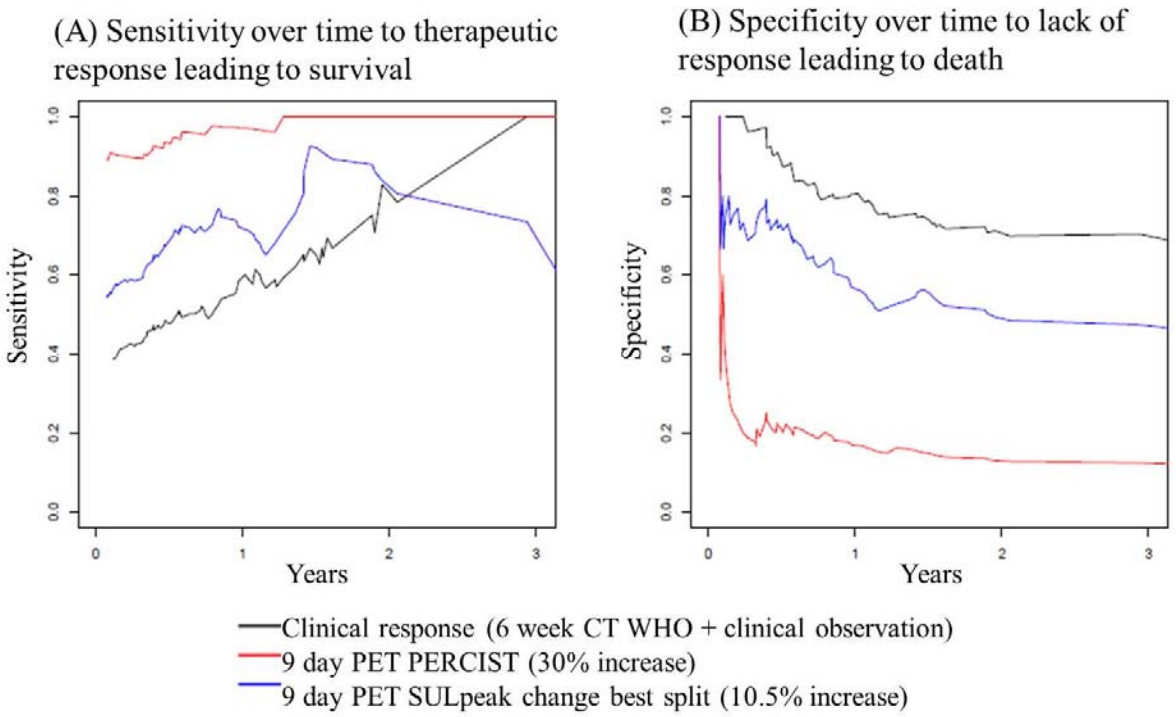


Supplemental Figure 2. ROC curves of different PET measurements for survival at 6, 12, 18 and 24 months.

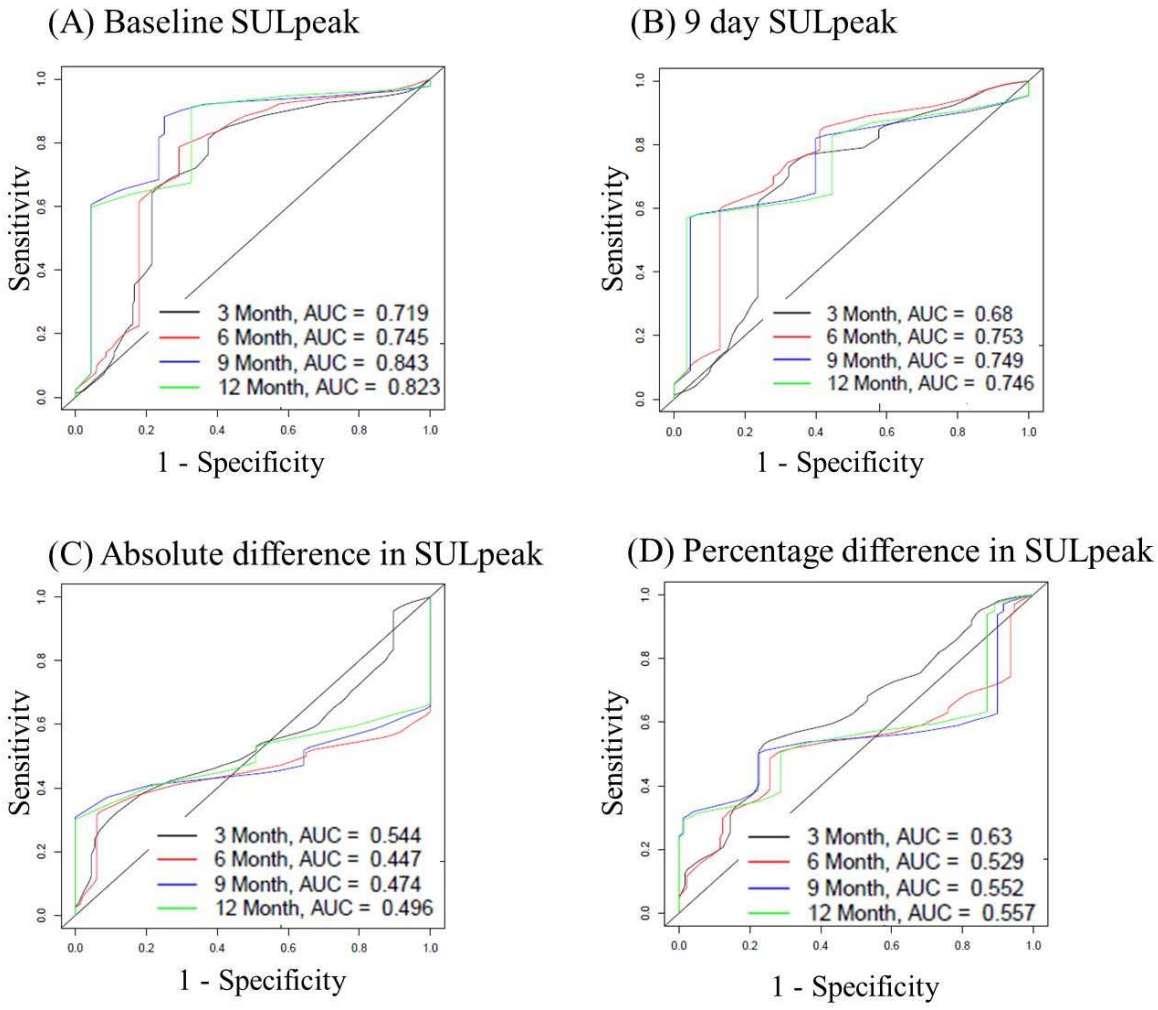


Supplemental Figure 3. PERCIST and clinical response. Sensitivity to a favorable therapeutic response leading to survival. Specificity to lack of a favorable therapeutic response, then leading to death. The figure shows the trade-offs of sensitivity and specificity of the day 9 PERCIST and week 6 clinical response to predict survival at different time points. The sensitivity and specificity of PERCIST over time were compared to those of clinical response. For a given time point, the reference standard is defined by the binary vital status of the patient at that time point. As shown on the left panel of figure 5, among the patients who live past 1 year, 95% could be correctly identified by PERCIST as non-progressive disease at day 9 after therapy; 70% could be correctly identified by greater than 10.5% decline of $\%\Delta SUL_{peak}$ at day 9; and 60% could be correctly identified by clinical observation and WHO criteria of non-progression at week 6.

The specificity of the three criteria to identify those who did not respond to the therapy and died within 1 year is 18% by PERCIST criteria of progressive metabolic disease at day 9; 58% by SULpeak with 10.5% split at day 9; and 80% by clinical response at week 6 (right panel of figure 5).It suggests that non-progressive disease based on PERCIST at day 9 is very sensitive in identifying who would respond to the treatment and live longer, and thus should be kept on the treatment. Using a SULpeak cutoff point of 10.5% or more in decline improves the specificity but lowers the sensitivity compared to the PERCIST definition based on 30% SULpeak change.


(A) Sensitivity over time to therapeutic response leading to survival
(B) Specificity over time to lack of response leading to death
Sensitivity
Specificity
Years
Years
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
Clinical response (6 week CT WHO + clinical observation)
9 day PET PERCIST (30% increase)
9 day PET SULpeak change best split (10.5% increase)

Supplemental Figure 4. ROC curves of the PET measurements for progression free status at 3, 6, 9 and 12 months.



Day 9 PET vs. progression free survival (PFS)

Progression free survival was based on clinical observation and week 6 CT response assessment by WHO criteria. The median PFS was 40 days (95% CI, 37 to 42 days) for the 115 patients. The Kaplan Meier estimation of PFS by dichotomized PET_{day9} response by PERCIST showed that patients with PMD had shorter PFS (median 31 days, 95% CI 27 to 37 days)

compared to patients with non-PMD (median 41 days, 95% CI 38 to 57 days; log rank p value=0.023).

The $PET_{baseline}$ and PET_{day9} SULpeak values were associated with progression at 3, 6, 9 and 12 months by ROC analysis (figure 6). This was more robust than the absolute and percentage change in SULpeaks. The $PET_{baseline}$ SULpeak had an AUC of 0.84 for progression at 9 months, PET_{day9} SULpeak had an AUC of 0.75 for progression at 6 months, and the $\%\Delta SUL_{peak}$ had an AUC of 0.63 for progression at 3 months. The PET_{day9} response of PMD showed increased hazard ratios for progression (hazard ratio 2.09, 95% CI, 1.10 to 3.97, p=0.024).