Usefulness and limits of the SUV for tumor characterization and patient monitoring in FDG PET/CT

Irène Buvat

Imaging and Modelling
in Neurobiology and Cancerology lab
U8165 CNRS - Paris 7 & Paris 11 Universities

buvat@imnc.in2p3.fr

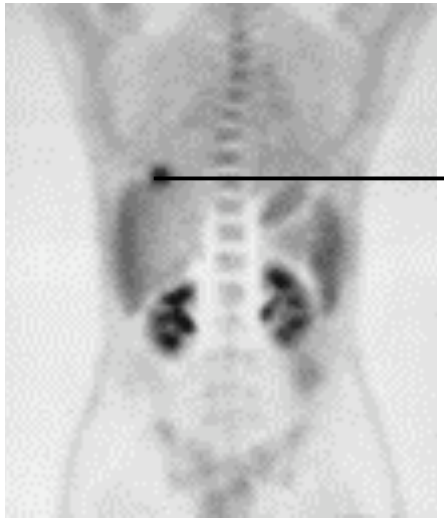
Introduction

- Need for quantitative indices to assist image interpretation in FDG PET
  - Diagnosis (+/--)
  - Patient monitoring (++/-)
  - Radiotherapy (+/-)

- Quantitative indices used in FDG PET:
  - SUV
  - Metabolically active volume
  - Total lesion glycolysis (TLG)

Can we trust them?
How to make the best use of quantitative indices?

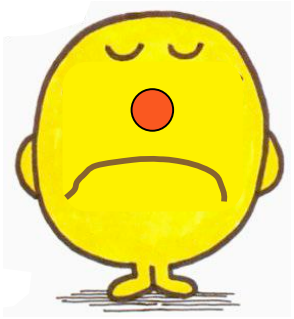
Definition of the SUV (Standardized Uptake Value)



$$SUV = \frac{\text{uptake (kBq/mL)}}{\text{injected dose (kBq) / “patient weight (g)”}}$$

If the tracer is uniformly distributed in the patient, SUV = 1 everywhere

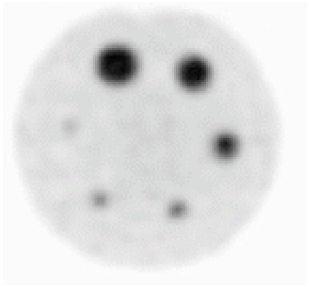
SUV ≠ 1, non uniform distribution

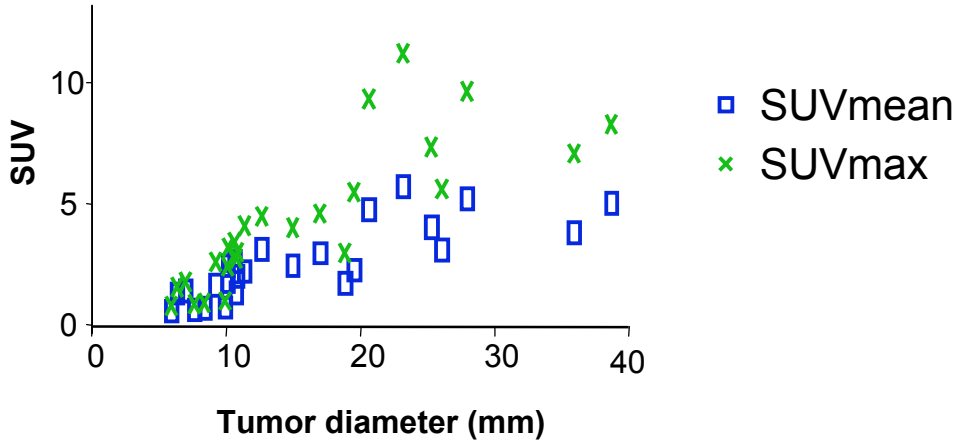


Clinical usefulness of the SUV

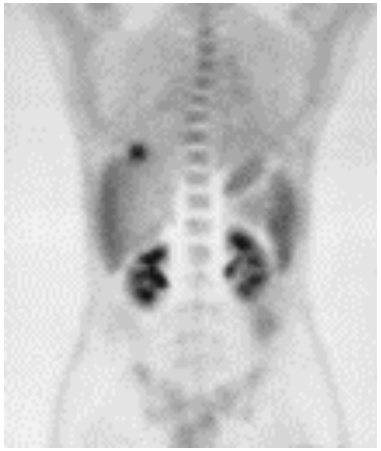
• Converting images into SUV makes them easily comparable : the expected value is 1 for any patient, whatever the injected activity and the body size

• Because of partial volume effect, SUV combines information regarding the metabolic activity AND the metabolically active volume, two quantities that are relevant for assessing tumor severity





Why should SUV be interpreted with cautious?



= estimate of the FDG distribution

not the same for all scanners and protocols

no standard way to measure it (SUVmean, SUVmax…)

$$SUV = \frac{\text{uptake (kBq/mL)}}{\text{injected dose (kBq) / patient weight (g)}}$$

estimate of the FDG made available to the tissue

Reliability of the uptake measurement

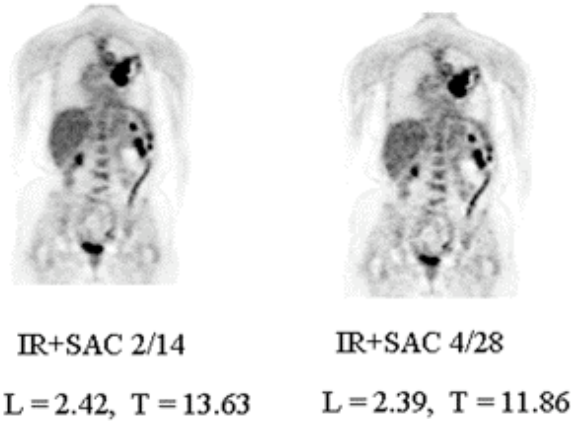
$$SUV = \frac{\text{uptake (kBq/mL)}}{\text{injected dose (kBq) / patient weight (g)}}$$

Affected by many factors:

- attenuation correction
- scatter correction
- respiratory motion
- partial volume effect
- tomographic reconstruction (noise, spatial resolution in the reconstructed images)
- measurement method

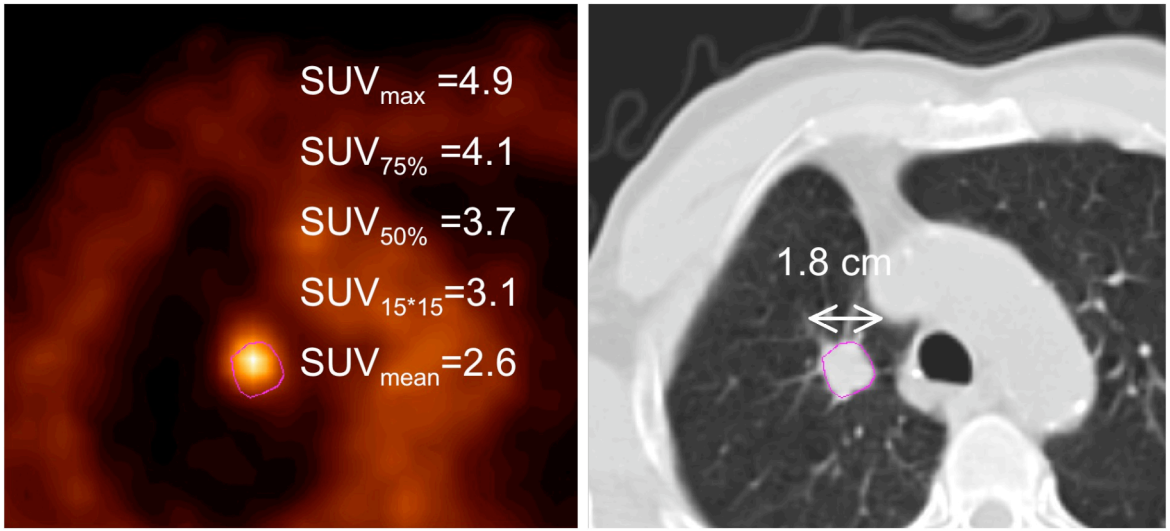
Examples of uptake variability

- Impact of image reconstruction



Schöder et al, J Nucl Med 2004:559-566

- Impact of the measurement method



Soret et al, J Nucl Med, 2007:932-945

Take-home message

Differences in uptake estimates greater than 100% (i.e., by a factor greater than 2) can be caused only by differences in the way data are acquired and processed. This suggests that comparison of SUV between PET centers using different scanning and processing protocols is almost impossible.

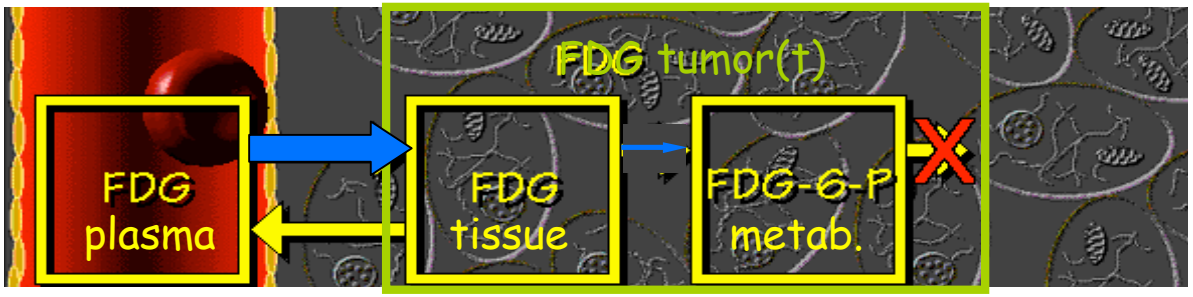
Feuardent et al, IEEE Trans Nucl Sci 2005: 1447-1452
Boellaard et al, J Nucl Med 2004:1519-1527
Jaskoviak et al, J Nucl Med 2005:424-428
etc…

- Potential solutions:
  - standardization of imaging and processing protocols
  - quality control of the estimated SUV using a standard procedure

Relevance of the normalization factor

$$SUV = \frac{\text{uptake (kBq/mL)}}{\text{injected dose (kBq) / patient weight (g)}}$$

- Crude estimate of the FDG made available to the tumor, ignoring:
  - the non steady state at acquisition time
  - the unmetabolized FDG
  - the plasma glucose level

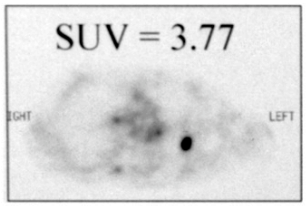


- No standard normalization (patient weight, lean body mass, body surface area, …)

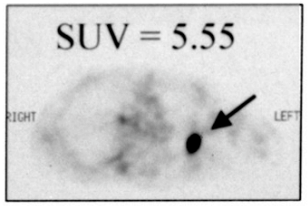
SUV estimates a normalized tumor uptake
but not the Glucose Metabolic Rate

Consequences of the crude normalization

- SUV still depends on time



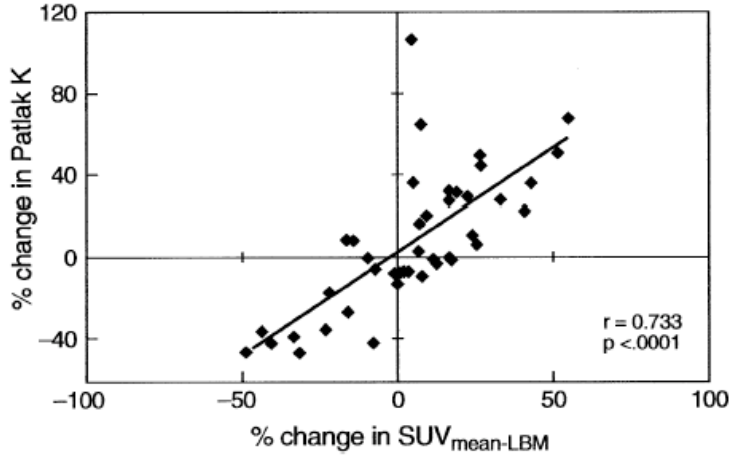
54 min post injection

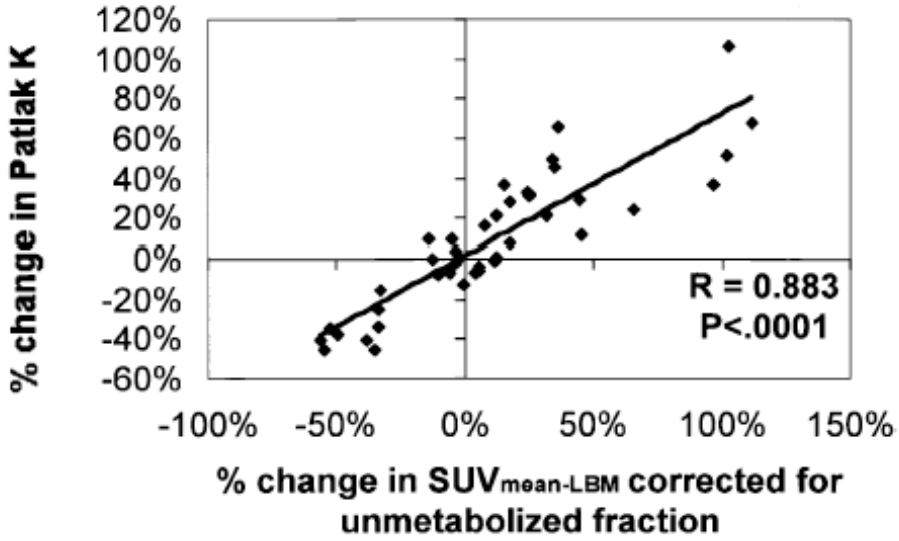


103 min post injection

Zhuang et al, J Nucl Med 2001: 1412-1417

- Mediocre correlation between SUV and GMR, partly due to unmetabolized FDG (between 6 and 67% of total FDG)





Freedman et al, Eur J Nucl Med 2003:46-53

Take-home message

SUV estimates a normalized tumor uptake, with no simple relationship with the Glucose Metabolic Rate.

How SUV is linked to GMR depends on patient specific parameters, that might change over time.

e.g., Freedman et al, Eur J Nucl Med 2003:46-53

- Potential solutions:
  - standardization of imaging and processing protocols (time between injection and acquisition, normalization factor)
  - in the future: dynamic imaging to better assess the GMR?

Practical results: huge variability of SUV estimates

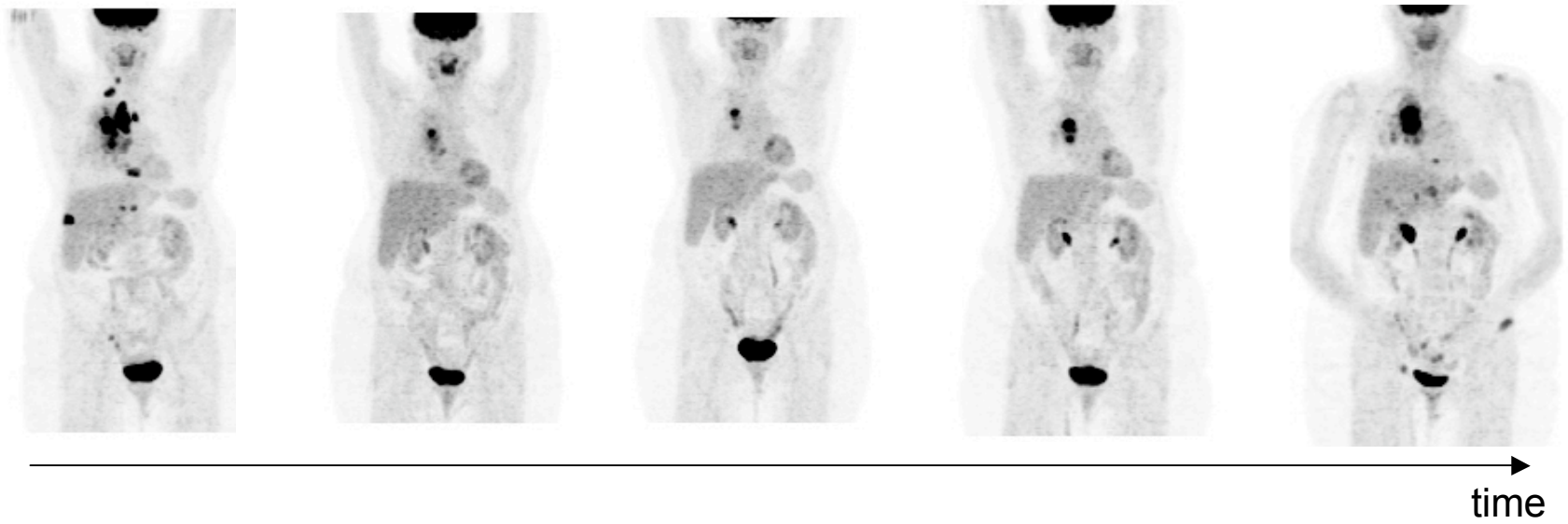
- Meta-analysis of the literature including 13 studies considering the prognosis value of the SUV in the primary tumor for NSCLC

Study	Type of SUV	SUV normalization	SUV threshold definition	SUV threshold
Ahuja (26)	SUV mean (SUR)	Weight	Best cut-off	10
Sugawara (23)	SUV max	Lean body mass	Median	8.7
Vansteenkiste (22)	SUV max	Weight	Best cut-off	7
Dhital (20)	SUV max	Weight	Best cut-off	15 or 20
Higashi (16)	SUV mean	Weight	Best cut-off	5
Jeong (18)	SUV max	Weight	Best cut-off	7
Downey (25)	SUV max	Weight	Median	9
Port (11)	Non specified SUV	-	Arbitrary	2.5
Sasaki (24)	SUV max	Weight	Best cut-off	5
Prevost (21)	SUV mean SUV max	Weight Lean body mass	Literature value	10
Eschmann (19)	SUV mean	Weight	Best cut-off	12
Borst (14)	SUV max	Weight	Median	15
Cerfolio (13)	SUV max	Weight	Median	10

Berghmans et al, J Thoracic Oncol 2008: 6-12

Context of patient follow-up

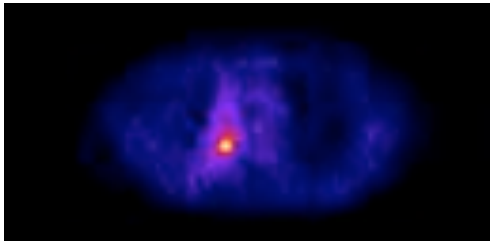
• Given that a department can operate always with the same scanner and protocol, can SUV be used for patient follow-up?



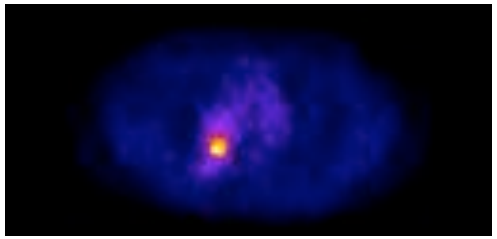
Are other parameters (tumor volume and TLG) more reliable?

Preliminary study

- 2 patients with NSCLC (grade IV) undergoing chemotherapy, with 5 and 6 PET/CT scans
- 17 tumor changes, rated as 9 partial responses, 2 stable tumors and 6 progressive tumors



Scan2



Scan3 (12 weeks later)

- For each tumor, calculation of:
  - 4 metabolically active volumes: $V_{40\%}$, V_{Nestle}*, V_{expert}, V_{optim}§
  - 6 SUV: mean SUV in the 4 volumes, SUV_{max}, SUV_{15mm}
  - 4 TLG: V x SUV_{mean} in the corresponding volume

For each index I ($V_{40\%}$, SUV_{mean}, etc) and each tumor:

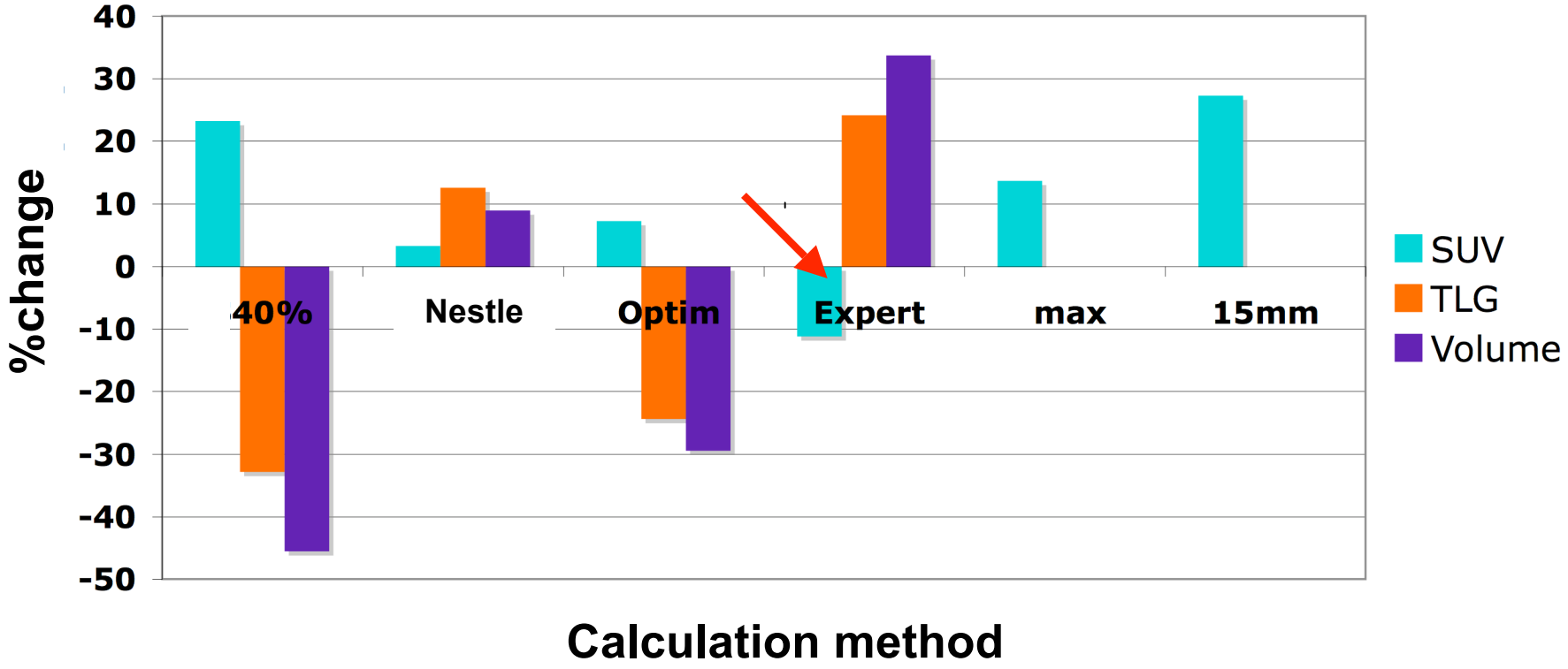
$$\%change = 100x(I_{scan(j+1)} - I_{scan(j)})/I_{scan(j)}$$

* Nestle et al. J Nucl Med 2005; 46:1342–1348

§ Tylski et al, J Nucl Med 2007; 48:43P

Example

- Patient 1, tumor 2, changes between scans 2 and 3

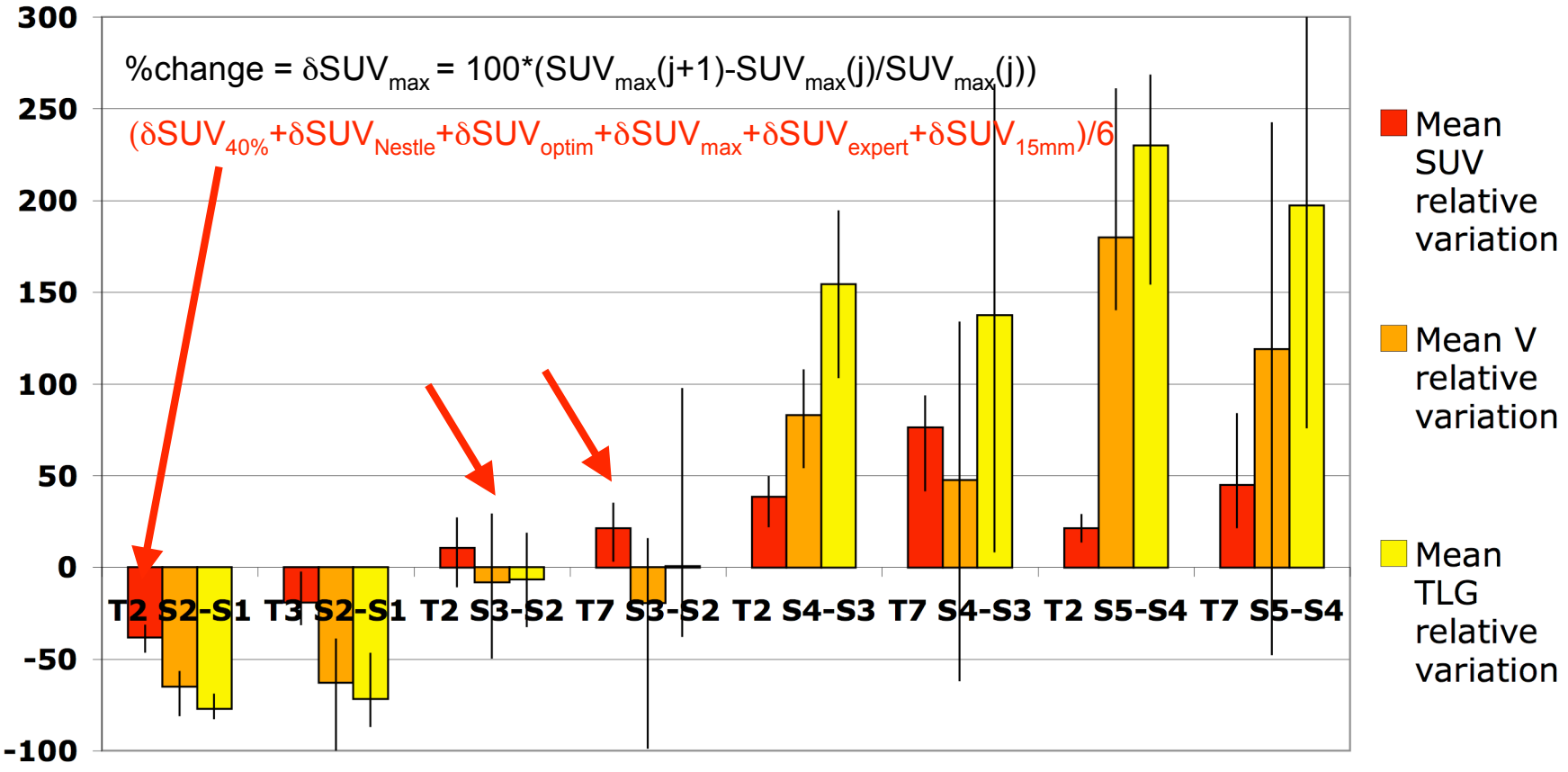


Results are not always consistent between measurement methods

Tylski et al, SNM 2008

Averaging over the different measurement methods

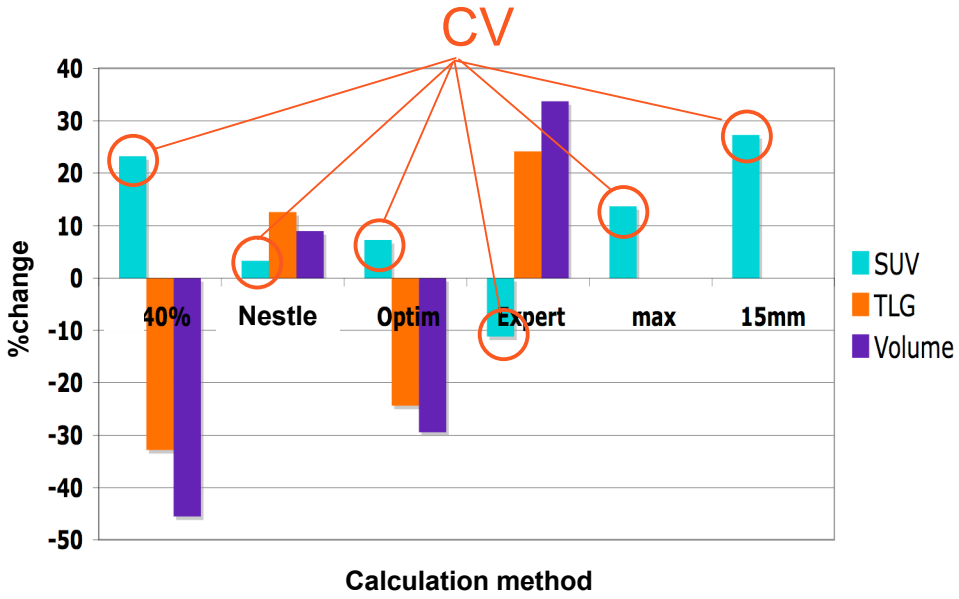
- Patient 1



SUV, volume and TLG do not always vary the same way

Tylski et al, SNM 2008

What is the most reproducible index?



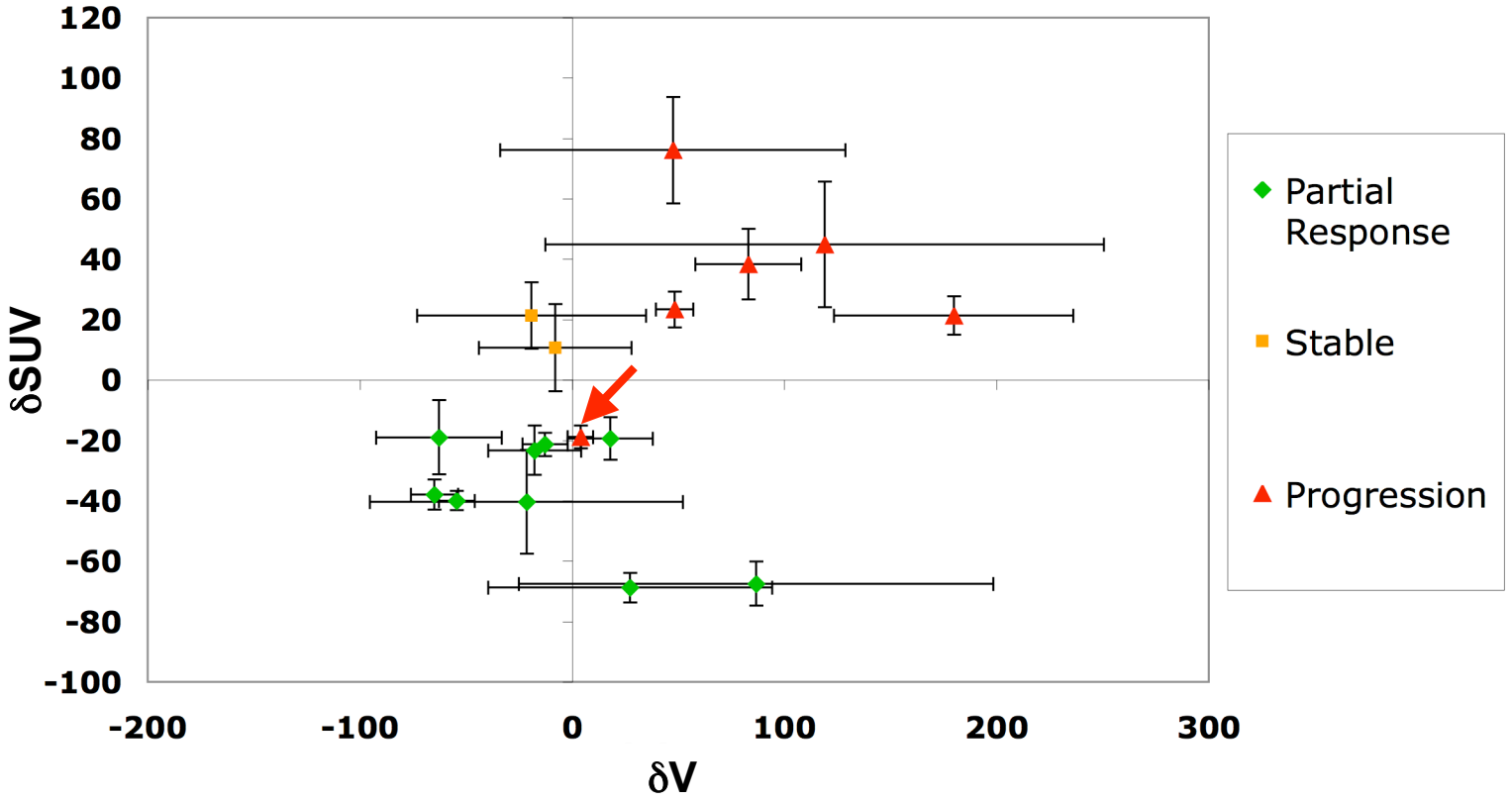
For each parameter (SUV, V or TLG), the coefficient of variation CV (σ/m) measures the variability of the parameter value as a function of the measurement method

- Averaged CV on the 17 couples of tumors :
  35% for δSUV
  80% for δTLG
  138% for δV

Changes in SUV appeared to be less dependent on the measurement method than changes in volume or TLG

Tylski et al, SNM 2008

Correlation with the clinical classification

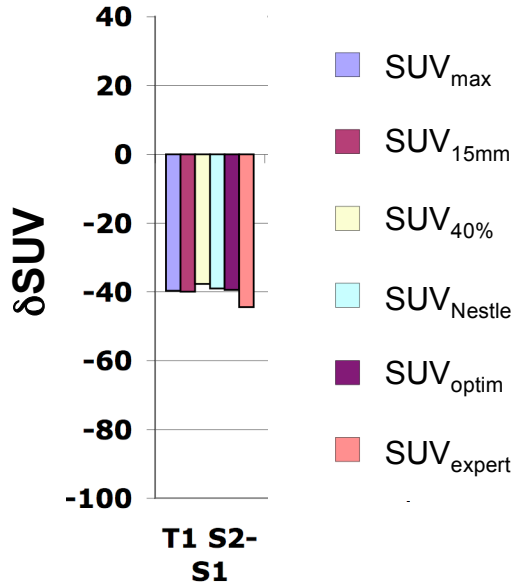


Joint analysis of δSUV and δV was needed to recover the clinical classification, while the analysis of δTLG alone did not distinguish between the 3 different types of evolution

Tylski et al, SNM 2008

How to take advantages of the different measurement methods?

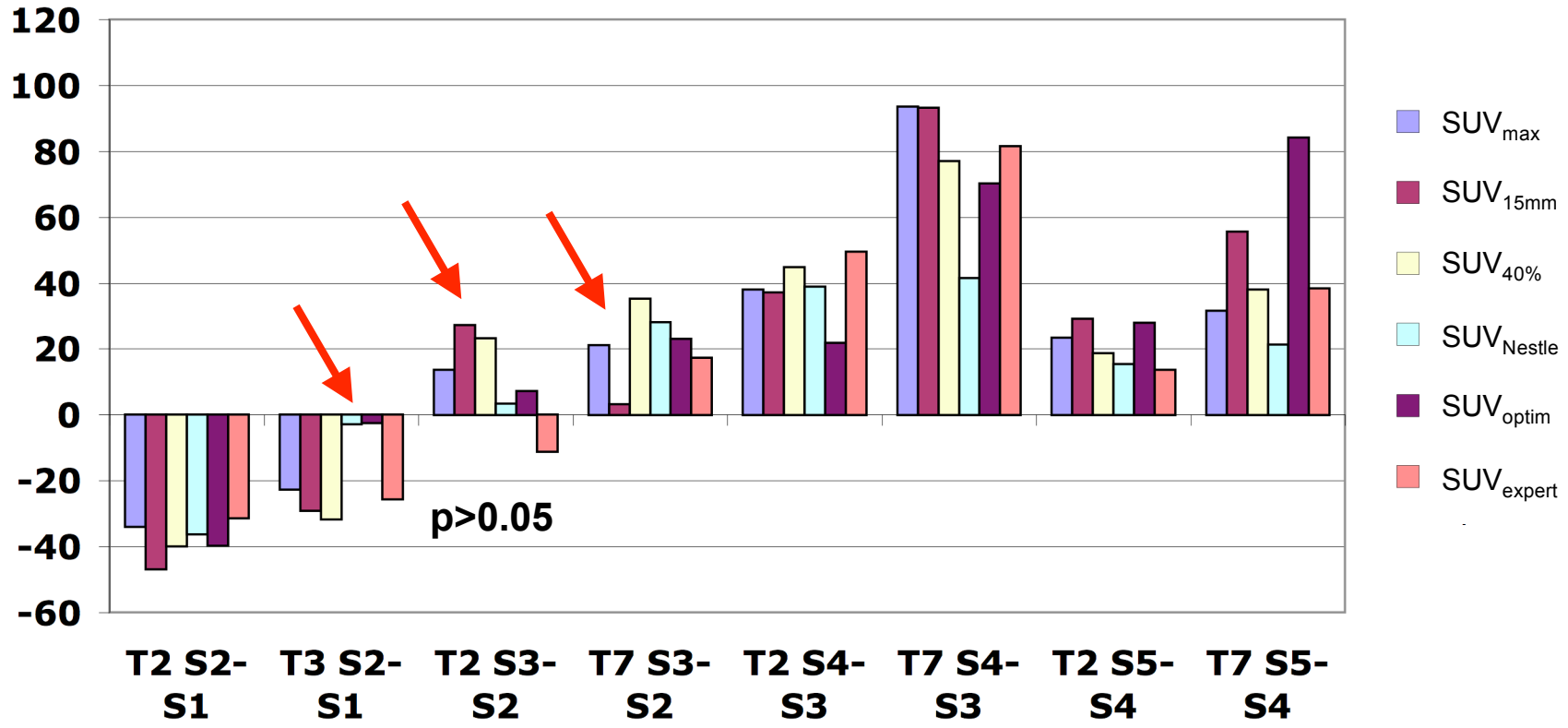
• If all SUV (SUV_{max}, SUV_{mean}, SUV_{expert}, …) vary in the same direction from one scan to the next, it is likely that the change is real



• Method: systematic calculation of the 6 δSUV and test of H0: δSUV=0 (non parametric test based on the bootstrap)

Results (example)

- Patient 1 : all changes were significant except one



Assuming that changes < 10% were not considered significant, considering a single measurement method would have led to a wrong conclusion in 7/102 cases

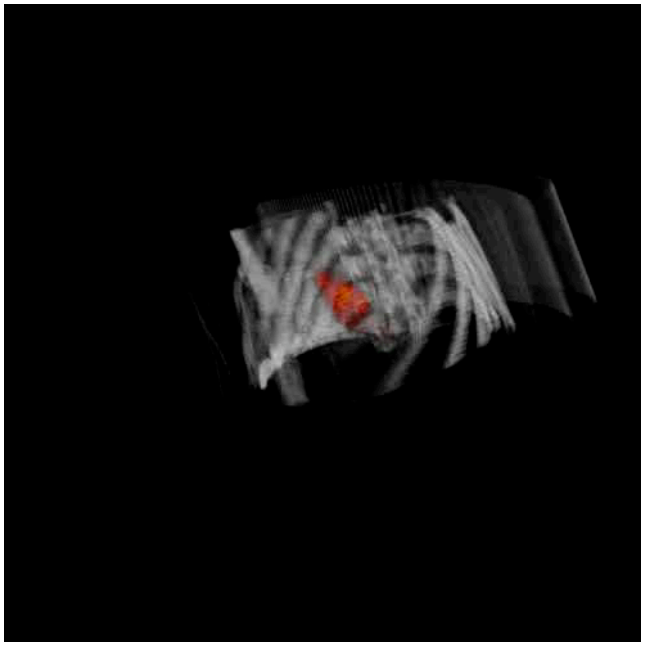
Tylski et al, SNM 2008

Take home message

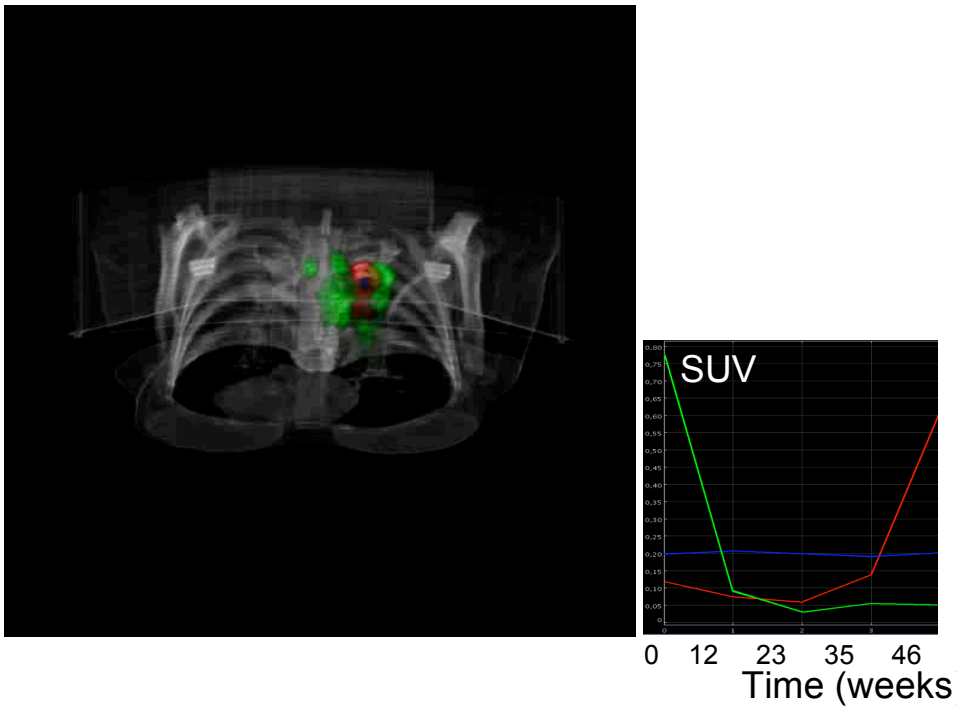
Having several measurement methods (SUV_{max}, SUV_{mean}, etc) for the same parameter can be taken advantage of to assign a level of significance to the observed change in SUV

Other approaches to go beyond the SUV for patient monitoring

- Parametric imaging: characterizing the tumor changes at the voxel level



between 2 scans: image subtraction
(similar to SISCOM or SPM)



over a series of scans:
using time series analysis

Necib et al, SNM 2008

General conclusion

- Standardization of measurement procedures and quality control are needed for consistent SUV measurements within and between departments

- Despite the intrinsic limitations of SUV, quantitative analysis can still be helpful and rigorous (e.g., statistical tests, aso)

- In any case, images have to be acquired in quasi-identical conditions to observe changes related to the tumors only

Acknowledgments

Perrine Tylski, Hatem Necib

Michelle Dusart, Camillo Garcia, Bruno Vanderlinden
from the Bordet Institute, Brussels

Slides available on
http://www.guillemet.org/irene/conference