Mr Chairman, Ladies and Gentlemen,

Accuracy of standard errors associated with region of interest measurements in quantitative PET

I. Buvat, C. Riddell, M. Soret, S. Hapdey

U494 INSERM, CHU Pitié-Salpêtrière, Paris, France
For quantitative analysis of PET data, it would be useful to associate a standard error to the measurements that are performed from the images. This would help to have an idea about the soundness of the estimated parameters. In kinetic modeling, it would enable a study of error propagation. Weighted fits accounting for the statistical uncertainties could be used with advantages instead of unweighted fits.

When the error associated to an ROI measurement is required, a common practice to estimate this error is by calculating the standard error of the value inside the ROI. Such a procedure assumes that the variance within the ROI is identical to the variance that would be obtained had the measurement been made several times on repeated acquisitions performed under exact identical conditions.

The purpose of this paper was to assess the accuracy of such a practice in PET and to suggest alternative methods.
First, let us precise what we are willing to measure and what we are actually measuring.

When calculating a mean value in an ROI, the uncertainty we would like to estimate is the variability that one would get if one could repeat the same measurement on several images acquired in exactly the same conditions.

If one estimates this variability by the variance associated with the mean within the ROI, one implicitly makes two assumptions:

1) first, that the signal within the ROI is uniform
2) second, that the statistical fluctuations for repeated measurements across time are identical to the statistical fluctuations of the signal within the ROI: this is the hypothesis of ergodicity.

Actually, both assumptions can be questioned:

1) Indeed, if the region of interest encompasses a small structure, it might include signal variation not due to noise but due to the fact that the noise free signal is actually not uniform in the ROI. This is what we will call the signal variance.

2) Second, even if the actual underlying signal in uniform, noise can be spatially correlated, which makes the ergodicity assumption wrong. This is what we call the noise spatial variance.
So, the variance within an ROI includes two components: the signal variance and the noise spatial variance.

We can identify these two components using repeated measurements performed under the same conditions.

Indeed, a robust estimate of the total ROI variance can be obtained by measuring the ROI variance for each of the replicated image and taking the mean. This ROI variance includes both the signal variance and the noise spatial variance.

To determine the noise spatial variance, one can calculate the mean of the N replicates and note that the noise spatial variance of this mean image is N times less than that measured on one of the replicated data set only, while the signal variance part is identical. So the noise spatial variance can be estimated by solving this system, which yields this expression.

Knowing the noise spatial variance, the signal variance can be deduced by subtracting the noise spatial variance estimate from the robust estimate of the ROI variance.

On the other hand, the ideal variance one would like to estimate is given by the variance of the mean measured on each of the replicates.

Note that the difference between the noise spatial variance and the ideal variance measures the departure from ergodicity.
To assess the importance of the different variance components, we considered a cardiac phantom for which 40 replicated acquisitions were acquired. The phantom had FDG activity within the left ventricle wall and the soft tissues but no lung activity.

Each dataset was reconstructed using 5 different schemes. The different variance components were assessed in 14 ROIS as shown here: 3 were located in the LV wall and 11 were located in soft tissues.
First, let’s look at the goodness of approximation according which the ROI variance is identical to the true variance.

Here, you can see the ratio of the ROI variance to the true variance for the 14 regions for FBP reconstructed with a Ramp filter and for OSEM with 24 iterations. Ideally, this ratio should be 100%.

For these two reconstruction algorithms, the goodness of the approximation depends a lot on the considered ROI. For FBP with a Ramp filter for instance, there are some regions in which the ideal variance is very well approximated while in others, there are underestimation by as much as 60%.

So the first point is that in a given image, the discrepancy between the spatial variance and the true variance varies a lot across the image and also depends on the reconstruction algorithm.
To better examine the impact of the reconstruction algorithm, here are the same graphs as before for all algorithms.

As previously observed, for all reconstruction algorithm, the discrepancy between ROI variance and true variance strongly depends on the ROI.

This slide also shows that the smoother the reconstructed images, the more the ROI variance underestimates the true variance. This is because the smoother the image, the more correlated the noise and correlated noise means violation of the ergodicity assumption.

Approximating the variance by the ROI variance can therefore be very misleading when using images with highly correlated noise.
Let's now look at the two components of the ROI variance for OSEM with 24 iterations. You can see here the ratio of signal variance to ROI variance for the different regions. It appears that the signal variance is only significant in the LV regions, that is in the regions that are strongly affected by partial volume effect. In all soft tissue regions, the signal variance is negligible.

The signal variance was only significant in the LV regions for all reconstruction algorithms. However, the ratio of signal to total ROI variance was affected by the reconstruction algorithm. Here, you can see the proportion of signal variance in the ROI variance in the LV for the different algorithms.

Again, the smoother the image, the greater the contribution of signal variance. This is because smooth image yields more partial volume effect, hence greater intrinsic signal variance within ROIs that encompass small structures such as the LV.
Now, one might wonder how we could avoid that.

We think that a proper estimate of the variance regardless of the reconstruction algorithm absolutely requires some repeated measurements. Of course, in practice, only a very small number of repeated acquisitions could be obtained. To best take advantages of only few measurements, we proposed last year at the SNM a non parametric bootstrap approach. We compared the variance values obtained by this approach to those obtained from ROI measurements.

Here are the ratios of the estimated variance over the true variance for different estimation procedures and for the different reconstruction algorithms. The ratios were averaged over all 14 ROIs.

As shown previously, the mean bias varied a lot for different reconstruction algorithms when using the ROI method to estimate the variance. When using a bootstrap approach from only 2 measurements, there was a significant underestimation of the true variance, but the bias was almost identical regardless of the reconstruction algorithm and of the ROI.

When using the bootstrap with 5 repeated measurements, the estimates was quite good and again, the bias did not vary much for the different reconstruction algorithms and different regions. The greater the number of replicates, the more accurate the variance estimate.
In conclusion, an important take home message is that you should not use ROI variance to estimate the variance associated with the mean measured in an ROI. The resulting biases strongly depend on the region and on the reconstruction algorithm.

Actually, even if you use this index to compare the noise level for different reconstruction algorithms, your conclusions might be quite wrong. For instance, from our results, using variance values as estimated by the ROI method, we would conclude that changing from 24 to 240 iterations in OSEM increase the variance by 7.5 while the true answer was 1.5.

We suggested an alternative approach based on the bootstrap which improves the variance estimate even with only few measurements. If such replicated data sets could be acquired systematically, this approach would certainly be very helpful to associate variance values with quantitative measurements in PET.

Thank you for your attention.