Magnetic resonance imaging-guided partial volume correction of positron emission tomography data in PET/MRI

Authors:

Kjell Erlandsson, PhD*
Institute of Nuclear Medicine, University College London, London, UK
e-mail: k.erlandsson@ucl.ac.uk

John Dickson, PhD
Institute of Nuclear Medicine, University College London Hospital, London, UK
e-mail: john.dickson@ucl.ac.uk

Prof. Simon Arridge, PhD
Dept. of Computer Science, University College London, London, UK
e-mail: s.arridge@cs.ucl.ac.uk

David Atkinson, PhD
Centre for Medical Imaging, University College London, London, UK
e-mail: d.atkinson@ucl.ac.uk

Prof. Sebastien Ourselin, PhD
Centre for Medical Imaging Computing, University College London, London, UK
e-mail: s.ourselin@cs.ucl.ac.uk

Prof. Brian F. Hutton, PhD
Institute of Nuclear Medicine, University College London, London, UK, and
The Centre for Medical Radiation Physics at the University of Wollongong, Northfields Ave, Wollongong, Australia
e-mail: b.hutton@ucl.ac.uk

* Corresponding author
Address: UCLH, Euston Rd (T-5), London NW1 2BU, UK

Key words: Partial volume effects, Partial volume correction, PET/MRI, Quantification

Key points:

- PVC is important for accurate quantification in PET and for avoiding confounding factors related to changes in PVE, but is currently not routinely used in clinical practice.
- Co-registered anatomical information from MRI can be utilised for accurate PVC of PET data.
- Anatomically guided correction for PVE can be performed as a post-reconstruction procedure or during the image reconstruction process.
- The introduction of PET/MRI scanners could facilitate the use of PVC in routine clinical studies.
Disclosure:
The Authors have nothing to disclose.

Synopsis:
Partial volume effects (PVEs), are caused by the limited spatial resolution of the PET system, which results in a blurry appearance of the PET images and a reduction in the observed activity for small objects. There is increasing evidence that partial volume correction (PVC) is necessary in order to guarantee quantitative accuracy in PET. However, there is reluctance to apply PVC routinely in clinical practice, partly due to uncertainty regarding the method of choice.

In order to perform accurate PVC, it is necessary to introduce information from high-resolution anatomical images, such as CT or MRI. A number of correction methods have been developed in the past; either post-reconstruction or reconstruction-based methods. The latter can be more accurate, but the former are easier to implement. All the methods rely on accurate co-registration between the anatomical image and the PET image. PET/MRI offers clear advantages for PVC and can help alleviate the image registration issues.
**Introduction**

Partial volume effects (PVEs), are caused by the limited spatial resolution of the PET system, which results in a blurry appearance of the reconstructed PET images, and as a consequence a reduction in the observed maximum activity for small objects (1) (see Box 1). Partial volume correction (PVC) is essential for accurate quantification in PET, especially in small objects. PVC can also be important in situations when the PVE changes over time or between subjects (see Box 2).

PVEs can be regarded as two separate effects: spill-out of data from inside an image region and spill-in of data from outside into the region. The amount of spill-over between regions depends on the point-spread function (PSF) of the imaging system, which is often modelled as a 3D Gaussian function, characterised by its full-width at half-maximum (FWHM) – the larger the FWHM, the larger the spill-over. The blurring of the reconstructed PET distribution can be described by an integral transform, which, if the PSF is position-invariant, is called convolution. In principle, it is possible to reduce the severity of the PVEs by performing an inverse filtering or de-convolution operation on the reconstructed image (2, 3), or by using a resolution recovery technique during image reconstruction (4, 5). However, this type of methods cannot provide a full correction due to loss of high-frequency information from the measured PET data. In order to perform a more accurate correction for PVEs, it is necessary to introduce additional information, which can be obtained from high-resolution anatomical images, such as CT or MRI. A number of correction methods based on anatomical images have been developed in the past, although these are not in general being used routinely in clinical practice. Various reviews of PVC methods have been published previously (1, 6, 7).

In anatomically guided PVC methods, the anatomical image must, first of all, be co-registered with the PET image. Furthermore, most methods require that it be segmented into regions, which, for the purpose of the correction, can be assumed to contain a uniform distribution of activity. The result of the PVC is sensitive to errors in the co-registration and segmentation steps (8-10). The image co-registration can be performed with a range of available software tools (11-13), but it obviously
becomes easier if the images have been acquired on a multi-modality system, such as PET/CT or PET/MRI (14, 15). With a PET/CT scanner, the CT scan and the PET data acquisition are actually performed sequentially, so there is a possibility of patient motion between the two. On the other hand, a PET/MRI scanner technically allows for truly simultaneous acquisition of PET and MRI data. Although co-registration may still be needed, the adjustment needed should be minimal and any residual errors small. With regards to segmentation, MRI is in general preferable to CT, as it has better soft-tissue contrast, which is especially important in brain studies. For these reasons, the recent introduction of PET/MRI scanners could greatly facilitate the use of PVC in routine clinical studies.
[Box 1] The partial volume effect
Figure 1 shows results from a simulated NEMA-EIC phantom. The highest value in the largest spheres is correct in the noise-free case, but for the smaller spheres, the maximum value decreases as the diameter decreases. The true values in all the spheres can be restored by PVC.

Figure 2 shows the max and mean values as a function of diameter. With noisy data, the max value gives over-estimation for the larger spheres and under-estimation for the smaller spheres, while the mean value gives under-estimation for all spheres and is much less sensitive to noise.

Figure 1: Simulated phantom with 6 hot inserts. Left panel) Original phantom (a), reconstructed image without noise (b), and with noise, before (c), and after PVC with RBV (16) (d). Right panel) A circular profile through the centre of the spheres in the original phantom (blue line), and in the reconstructed image without noise (red line).

Figure 2: Maximum (blue lines) and mean values (red lines) in the spherical inserts as a function of sphere-size for noise-free data (nf, solid lines) and data with noise (wn, dashed lines).
[Box 2] Clinical PVEs

PVC can be important in situations when the PVEs change over time or between subjects, and could be a confounding factor in the interpretation of the PET images.

In neurology studies of patients with Alzheimer’s disease, atrophy can lead to thinning of the cortical gray matter layer as compared to normal subjects (Fig. 3a). In cardiology studies, the thickness of the myocardial wall changes during the cardiac cycle, being thicker at end systole as compared to end diastole (Fig. 3b). In oncology studies, the size of a tumour can change after therapy (Fig. 3c).

**Figure 3:** Schematic illustration of the effects of PVE in various clinical situations. a) Cortical thinning in a neurological study, leading to an apparent decrease in the tracer uptake. b) Cardiac contraction, leading to apparent increase in tracer concentration. c) A change in tumour size and tracer concentration, resulting in unchanged apparent mean concentration. In each panel, the top row shows the true tracer distribution, and the bottom row the PVE affected distributions. The left and right columns show anatomical changes occurring over time.
PVC methods

The correction for PVEs based on anatomical data can be considered to fall into two main approaches; post-reconstruction methods and reconstruction-based methods. A brief historical overview of these methods is presented as well as a review of the findings for various clinical applications. Of the two main approaches, the post-reconstruction methods appeared first. Initially the correction was applied to the mean value for one or more volumes-of-interest (VOIs), and later voxel-by-voxel methods were developed. The next step was to incorporate anatomical information directly into the image reconstruction algorithm. It has been shown that the reconstruction-based methods can be superior in terms of bias vs. noise trade-off (17). On the other hand, the post-reconstruction methods have the advantage of being easier to implement, and can be combined with the standard reconstruction software provided with the system. The historical evolution of the different methods is shown in Fig. 4.

Post-reconstruction methods

In 1979, Hoffman et al. (18) proposed a method to correct the mean value in an image region or VOI, assuming a uniform distribution within the VOI and no background activity. The corrected value was obtained by dividing the un-corrected mean value by a pre-calculated recovery coefficient (RC), which depended on the shape and size of the VOI and the system PSF. This method corrected for spill-out but not for spill-in. From here, multiple PVC methods evolved in different directions. RC-based methods have often been used for PVC in oncology (19).

Regional simultaneous estimation methods (A)

A series of method were developed, based on the principle of simultaneously estimating the mean values in two or more regions by first calculating the RCs for all regions and the cross-talk coefficients for all combinations of regions, and then solving a system of linear equations. This approach corrected for both spill-in and spill-out. In 1983, Henze et al. (20) developed a PVC method for cardiac studies, with simultaneous estimation of the activity concentrations in the myocardial
wall and the ventricular cavity of the left ventricle. Four recovery/cross-talk coefficients were determined, based on the geometry of the heart and the system PSF, and the corrected mean values were obtained by solving a system of 2 linear equations with 2 unknowns, assuming no background activity. In 1998, Rousset et al. (21) presented a similar method for neurology, but with many more regions, covering the entire brain. A matrix of recovery/cross-talk coefficients was generated (the geometric transfer matrix (GTM)), and the corrected mean values of all regions were obtained by multiplying the inverse of this matrix with a column-vector containing the un-corrected mean values. This method is known as the GTM method. A similar method was presented by Labbé et al. (22). Here the image was modelled as the sum of a number of components, each one corresponding to a single region convolved with the system PSF, multiplied by un-known coefficients. The authors used singular value decomposition to determine the values of the coefficients, which represent the corrected mean values for the different regions, but other mathematical techniques could be used. For example, an alternative algorithm was presented by Sattarivand et al. (23). In 2005, Du et al. (24) described a perturbation-based method for calculating the matrix coefficients for the GTM method (21). This technique allows one to take into account the non-linearity of iterative reconstruction methods. Although developed for SPECT, it is also applicable to PET data.

**Voxel-based additive and multiplicative methods (B)**

In previous methods, the outcome was a corrected regional mean value, but in 1988, Videen et al. (25) proposed a PVC method for voxel-by-voxel correction. A single VOI was used, and RC-values were determined for each voxel within the VOI. The application was neurology and a VOI corresponding to brain-tissue was obtained by segmentation of a CT image. It corrected for spill-out but not for spill-in, assuming no background activity. Meltzer et al. (26) implemented this method, using MRI instead of CT images. Müller-Gärtner et al. (27) developed this method further by introducing a spill-in correction term. The brain was segmented into gray matter (GM), white matter (WM) and cerebro-spinal fluid (CSF), and the correction was applied to GM voxels only, including correction for spill-in from WM, assuming the true mean value in WM could be directly estimated.
Erlandsson et al. (28) proposed a method, combining the GTM technique for estimation of the VOI mean values, with a Müller-Gärtner-type approach for performing a voxel-based correction of each VOI. With this method, called multi-target correction (MTC), the entire image could be corrected on a voxel-by-voxel basis, and no prior information was required.

**Purely multiplicative methods (C)**

In 1984, Kessler et al. (29) extended the RC-method in order to take into account spill-in from activity in the background. Correction factors were determined for given target-to-background ratios, which simultaneously corrected for spill-in and spill-out. In order to find the right RC value to use, it is necessary to know the lesion size and the target-to-background ratio. Methods to determine this information directly from PVE-affected PET images were proposed by Gallivanone et al. (30). A voxel-based method for neurology was presented by Yang et al. in 1996 (31). This is a purely multiplicative method, which is applied to the entire image. The brain was segmented into GM, WM and CSF, and an assumption had to be made regarding the relative activity concentrations in the 3 regions. In 2011, Thomas et al. (16) presented a method called region-based voxel-wise correction (RBV), similar to MTC (28), in which the GTM method (21) was combined with the voxel-based correction approach by Yang et al. (31). Erlandsson et al. (1) proposed an alternative algorithm, called “iterative Yang” (iY), in which the first step of RBV was replaced by an iterative procedure for estimation of the regional mean values, avoiding the GTM step.

**Wavelet-transform methods (D)**

In 2006, Boussion et al. (32) presented a novel approach for PVC, in which details from a high-resolution anatomical image were integrated into a low-resolution PET (or SPECT) image in the wavelet domain. Wavelet coefficients from the anatomical image, corresponding to high spatial frequencies, were transferred to the PET image, after scaling by the ratio of the wavelet coefficients from the two images at a lower frequency. Shidahara et al. (33) presented an improved version of this method, in which the MRI image was first segmented and parcellated using an anatomical atlas.
When the method was evaluated using simulated dynamic PET data, the results showed improved quantification (34), and when applied to real clinical data, the results were comparable to those of the GTM method (35).

**Deconvolution methods (E)**

In 2010, Segobin et al. (36) presented an iterative deconvolution algorithm, combined with regional averaging based on a partially segmented anatomical image. This approach could be useful where reliable segmentation was possible for only part of the image. Bousse et al. (37) proposed a post-reconstruction deconvolution technique, based on a maximum penalized likelihood algorithm, where the PET image is described by a model, assuming the existence of activity classes that behave like hidden Markov random fields, driven by a segmented MRI image. The method outperformed other post-reconstruction PVC methods when the segmented MRI was inconsistent with the PET image.

**Segmentation-free methods (F)**

In 2012, Wang & Fei (38) proposed a PVC method based on iterative deconvolution within a Bayesian framework, which included an edge-preserving smoothness constraint. The constraint was imposed using a neighbourhood structure, including voxels depending primarily on the difference in PET image values and secondly on the difference in MRI values. Yan et al. (39) also proposed a novel post-reconstruction method, which did not require segmentation of the MRI image. It was based on the assumption that a linear relationship exists between PET and MRI image intensities in a local neighbourhood around each voxel.

**Reconstruction-based methods**

Anatomical data can be incorporated into an iterative image reconstruction method as a priori information (or just “prior”) within a Bayesian framework. This means that the reconstruction algorithm is guided towards solutions, which have higher probability of being correct. This is known as a “maximum a posteriori” (MAP) algorithm. An alternative is to use penalised likelihood
algorithms, although, in practice, the two approaches are quite similar. Traditionally, a smoothing prior would be used in order to limit the noise-amplification, often seen in iterative algorithms. However, this also results in resolution degradation. The role of the anatomical prior is to restrict the resolution degradation and preserve anatomical features in the image. The correction for PVE is achieved by modelling the PSF during the reconstruction process. We can distinguish between different types of prior; some utilise the spatial distribution of the anatomical image, some the image intensity.

**Spatial prior methods (G)**

In 1991, Chen et al. (40) presented a Bayesian reconstruction algorithm, incorporating anatomical boundary information from CT or MRI images. They achieved significant improvement in image quality with simulated data. Fessler et al. (41) propose a penalised likelihood reconstruction method with boundary information from MRI images, which took into account possible errors in the boundary information due to errors in the co-registration or segmentation. This was done by blurring the corresponding weights in the penalty term. Gindi et al. (42) proposed a Bayesian reconstruction method, which utilised an edge-map from an anatomical image, taking into account the uncertainty in the location of the edges by a spatially varying modulation of the probability. Ouyang et al. (43) presented a Bayesian reconstruction algorithm based on the joint probability of structural and functional boundaries. Boundary information was extracted from anatomical images, but only those boundaries which had high joint probability with the corresponding PET data were used. Chiao et al. (44) presented a penalised likelihood reconstruction algorithm with anatomical boundary information for cardiac perfusion studies, with co-registration parameters included in the optimisation process.

In 1996, Ardekani et al. (45) proposed an anatomically guided reconstruction algorithm that did not require segmentation of the MRI image. It was based on the cross-entropy between the estimated PET image and a prior image model, obtained by edge preserving smoothing of the previous
estimate. The weights of the smoothing kernel were determined for each voxel, based on the difference in the MRI image values. Bowsher et al. (46) presented a Bayesian algorithm for reconstructing PET or SPECT images and, at the same time, segmenting them into a number of regions. Anatomical information was utilised by assigning higher prior probabilities to segmentations in which each segmented region stayed within a single anatomical region. Lipinski et al. (47) presented a Bayesian reconstruction method, which assumed that all voxels within an anatomical region have values belonging to a Gaussian distribution with a given mean value. The method resulted in images with low noise, but was sensitive to incomplete or erroneous anatomical information. Sastry and Carson (48) proposed a similar algorithm, in which a Gaussian distribution was assumed for different tissue types (GM, WM, CSF and “other”). Each voxel was modelled as being composed of the different tissue types in proportions determined by a segmented MRI image. Furthermore, a smoothness constraint was used for each separate tissue type.

In 2002, Comtat et al. (49) presented a modified version of the algorithm proposed by Fessler et al. (41), in which the anatomical labels were blurred in order to take into account possible registration-errors. Baete et al. (50) presented a MAP algorithm, for FDG brain studies. The MRI images were segmented into fuzzy classes corresponding to GM, WM, CSF and “other”. For the purpose of calculating spill-in to the GM region, the WM and CSF regions were assumed to be uniform. Bowsher et al. (51) proposed a segmentation-free approach for incorporating MRI information into the reconstruction process, similar to the one proposed by Ardekani et al. (45). An edge-preserving smoothing operation was used. The smoothing kernel was generated by choosing a subset of voxels within a neighbourhood, with the smallest difference in MRI values. In 2010, Vunckx et al. (52) described an improved version of this prior, which takes into account the fact that the voxel-neighbourhoods selected are not symmetric.
**Intensity prior methods (H)**

In 2000, Rangarajan et al. (53) introduced a novel Bayesian framework for incorporating anatomical information into emission tomographic reconstruction, which did not require that the anatomical and functional regions were exactly homologous. The algorithm used a prior based on the mutual information, calculated from the joint histogram for the two images. A segmented anatomical image was used. In 2005, Somayajula et al. (54) presented a segmentation-free algorithm, based on the mutual information between various image features in the PET and anatomical images. The algorithm was sensitive to local maxima, and it was necessary to start with a good initial estimate. Nuyts (55) compared two anatomically driven MAP reconstruction algorithms with priors based on mutual information and joint entropy of the image intensities, respectively. He found that the joint entropy prior was superior, although local maxima could lead to convergence problems. Tang and Rahmim (56) also developed a MAP reconstruction algorithm, with a joint entropy prior. Using simulated data, they obtained improved noise versus bias tradeoff even in a lesion in the PET image which had no anatomical correspondence.

In 2010, Tang et al. (57) presented a method, which incorporated anatomical information into an algorithm for direct reconstruction of parameteric images from dynamic PET data. The method was based on a graphical analysis approach (58), and the anatomical information was incorporated using a joint entropy prior within a MAP algorithm. Pedemonte et al. (59) presented a MAP reconstruction algorithm with a prior based on class conditional joint entropy, in order to take into account the underlying tissue composition. The MRI was assumed to be composed of GM, WM, CSF and “other” with Gaussian distributed image intensities. The method outperformed conventional methods based on joint entropy. Somayajula et al. (60) described an approach with priors based on mutual information and joint entropy. Scale-space theory provided a framework for the analysis of images at different levels of detail, and presented a solution to the non-spatial nature of these measures. The scale-space features were defined as the original image, the image blurred at different scales, and the Laplacians of the blurred images. In 2015, Tang and Rahmim (61) developed a MAP...
reconstruction algorithm with a joint entropy prior, in which local spatial information was incorporated using a wavelet decomposition. The algorithm performed better than a standard joint entropy MAP algorithm, based on image intensity only, in the case of noisy data.

**Diffusion prior methods (I)**

In 2009, Chan et al. (62) explored an alternative approach to incorporating an anatomical prior into PET image reconstruction, in particular for the situation where lesions are apparent in the emission images but not in the corresponding anatomical images. The proposed method was based on an anatomically adaptive anisotropic median-diffusion filtering prior. The proposed prior could yield improved lesion contrast and reduced bias without requiring knowledge of lesion outlines. Kazantsev et al. (63) proposed a penalised maximum-likelihood algorithm, incorporating an anatomically driven anisotropic diffusion filter, with edge-preserving de-noising characteristics, which has the ability to retain information that was absent in the anatomical image.

**A level set method (J)**

In 2011, Cheng-Liao and Qi (64) proposed an approach to MAP reconstruction of PET images with a level set prior guided by anatomical edges, which did not assume an exact match between PET and anatomical boundaries, but encourages similarity between the two.
Figure 4: The historical development of anatomically guided post-reconstruction PVC methods (left) and reconstruction-based methods incorporating anatomical data (right). (The letters A-J refer to corresponding sections in the text.)
Clinical applications of PVC

PET/MRI offers clear advantages for partial volume correction. The ability to acquire anatomical information sequentially or even simultaneously to the acquisition of PET data helps alleviate many of the image registration issues that can be problematic with these correction techniques. Furthermore, the exquisite level of soft tissue contrast offers advantages to CT when it comes to feature definition, and can also give a better understanding of the tissue mix in PET voxels. The recent introduction of PET/MR as an imaging modality means that there is a dearth of literature on clinical partial volume correction on these scanners. The techniques discussed in the following section describe the application of partial volume methodologies to clinical PET data. Most of the methods described can be applied to PET/MR data, and also give an appreciation of the challenges that will be faced when applying such techniques to PET/MR studies.

Cardiology

In nuclear cardiology, there is a clear need for partial volume correction, although it is not commonly applied in clinical practice. For the assessment of myocardial blood flow, uptake of tracer is measured in tissue which has a thickness of around 15 mm in its systolic phase, and typically becomes much thinner in the diastolic phase. The resulting changes in partial volume losses can be visualised when viewing ECG-gated PET data. Apparent changes in blood flow are seen throughout the cardiac cycle, when in reality, for most patients the blood flow should be relatively stable. The problem becomes more of an issue when we also consider the radionuclides typically used for myocardial perfusion imaging in PET. $^{13}$N-ammonia, $^{15}$O-water, and $^{82}$Rb-chloride are three agents commonly used in myocardial perfusion imaging, with each emitting more energetic positrons than the frequently used $^{18}$F. Indeed $^{82}$Rb emits very energetic positrons, which, because of the longer positron path, leads to poorer spatial resolution than that typically seen in PET. The correction for changes in apparent blood flow across the cardiac cycle, and the relative difference between $^{18}$F and $^{82}$Rb PET has been assessed in a paper by Johnson (65). To correct for the change in partial volume effects, the authors follow a two-step approach. In the first step the change in apparent uptake at
end-diastole compared to end-systole is assessed to calculate the relative difference in partial volume effects at the two phases. Then using phantom data with known thicknesses and partial volume losses, the correction for end-systole is determined, and from this the correction for end-diastole is derived. Using the ECG-gated data to determine the amount of time spent in the diastolic and systolic phases, it is then possible to obtain average whole heart cycle partial volume correction factors for ungated images.

The assessment of uptake in the myocardial tissue is not the only issue in myocardial perfusion imaging. Proper quantitative analysis of perfusion requires the use of kinetic modeling, which brings its own associated problems: the losses/spill-out of myocardial uptake into the surrounding areas including the ventricular cavities; and spill-in of activity into the myocardium from the ventricular cavities. While the latter is not too problematic at late time phases where blood pool activity is limited, at early and intermediate dynamic time frames there can be significant activity in the blood pool, and low but growing activity in the myocardium. This can be complicated further as a volume of interest in the ventricular cavity may be used as an image-derived input function for the kinetic modeling. One of the earliest PVC methods for myocardial imaging, suggested by Herrero (66), assumes a Gaussian point spread function representation of spatial resolution can be convolved with image data of known dimensions to estimate partial volume losses. The authors used a previously calculated value of the point spread function, an assumed thickness of myocardium, and $^{11}$C-monoxide PET as a way of measuring the size of the ventricular cavity to calculate the partial volume losses in a series of dog studies. In a study by Iida (67), two methods of partial volume correction were evaluated for the correction of myocardial perfusion data. The first measurement involved the inclusion of a tissue fraction in the kinetic model of $^{15}$O-water data, whereas the other method used a combination of transmission scan imaging and $^{11}$C-monoxide imaging to produce images of extravascular density. Of course, these techniques are applied to dynamic data, with no regard given to the differences across the cardiac cycle.
In recent years there has been a plethora of software available to calculate myocardial blood flow (68), with several different methods used to deal with partial volume effects. Some of these solutions take a simple approach to minimise PVEs by assuming fixed global values of partial volume loss, and by using image derived input function volumes placed in positions with minimal spill-in effects from surrounding tissue. Other researchers have used generalised factor analysis of the dynamic data itself in an attempt to intrinsically correct for PVEs (69). An alternative approach which has become popular in commercialised software (70, 71) is that proposed by Hutchins et al. 1990 (72). It assumes that if the blood to myocardium spill-over plus the tissue blood-fraction is equal to \( fb \), then the myocardial activity contribution can be represented by \( 1-fb \). This parameter can then be incorporated into the kinetic models to determine partial volume corrected kinetic parameters. This technique has since been modified further by using whole myocardium volumes of interest to reduce the spill-over effects into the image derived input function, which are in turn used to derive better regional values of myocardial blood flow (73).

**Oncology**

In oncology, the application of partial volume techniques in clinical use is more limited. In the most part, clinics report uptake in terms of maximum pixel value rather than the mean value in a region/tumour in an attempt to minimise the issues of partial volume losses. Of course, this is a primitive and rather limited solution because the region/tumour uptake is already likely to be affected by partial volume effects if it is small in size, and the reliance of a single maximum voxel value is also not robust. The introduction of SUV-peak, which looks at uptake in a group of pixels helps overcome the robustness to noise issues, but can still be affected by partial volume issues. There are also other issues to consider when introducing partial volume correction in oncology. Is correction required for a tumour volume, or for an image? Will the volume of interest be defined on the PET or corresponding anatomical image? Can the technique be easily applied and used in a busy PET clinic?
By the far the most common form of partial volume correction performed in oncology PET is the use of recovery coefficients. Assuming tumours are spherical in shape, and that their diameters can be determined, these methods use phantom measurements to predict the underestimation of uptake due to spill-out, and any spill-in of surrounding tissues. Adler et al. (74) was an early proponent of this method in the imaging of lymph nodes in breast imaging although spill-in effects were not considered. Avril et al. (75-77) expanded this method, again in breast cancer to look at different contrasts in phantoms and was thus indirectly looking at spill-in effects. Since these early attempts further refinements have been made. Vesselle et al. (78) working in lung cancer introduced a background region into the correction to deal with spill-in effects – something which is particularly important for small tumours. This approach was developed further by Hickeson et al. (79) who used very specific concentric background VOIs, and Hofheinz et al. (80) who developed a yet more complex method for background ROI determination. The Hofheinz technique has also been introduced into a clinical software platform to create a simple and automated approach to oncological partial volume correction with improved intra- and inter-operator error (81). Salavati et al. (82) applied this technique to FDG PET/CT studies of lung lesions with and without respiratory gating. They found that after PVC, SUV mean increased substantially, and that the PVE appeared to be the dominant source of quantitative error of lung malignancies. However, they found no clinically significant difference between respiratory-gated and non-gated data after PVC.

When applying such partial volume corrections, these and other authors have found improvements in diagnostic performance or better correlation with other physiological metrics. The success of the recovery coefficient technique is based on its simplicity, although it does have its limitations. The assumptions that the tumour is spherical, and that its uptake is homogeneous are questionable, particularly for larger tumours that may also have necrotic centres. Another issue is the determination of the tumour radius. The higher spatial resolution CT component from PET/CT, or indeed the MRI component from PET/MRI systems can potentially help, but unfortunately the metabolic volume does not always correspond to the anatomical volumes that these modalities
offer. Defining tumour volumes based on PET has been attempted (79, 83), but the robustness of this technique has yet to be proven particularly in smaller lower contrast lesions. A final issue with recovery coefficient techniques is that the correction is performed on a volume and not on the image as a whole. This is generally not a problem in oncological PET, where the tumour uptake is typically the main output metric of interest, but this approach is not helpful when the general distribution of uptake or heterogeneity of uptake within a tumour is required.

Other partial volume techniques in oncological PET are more complex, and because of this only appear in proof of concept studies with small patient numbers (32, 84, 85). An approach that has gained some interest is the use of deconvolution using Van Cittert (2) or Lucy-Richardson techniques (86) to iteratively recover resolution losses across an image. One problem with these methods is that noise can be amplified using these approaches, although de-noising using e.g. a wavelet methodology has been applied with good results (86, 87). Another issue with deconvolution techniques is that they are dependent on an accurate measurement of spatial resolution (point-spread function). Because spatial resolution changes across the PET field of view, this can be difficult to model. This issue affects other PVC methods as well, but it is more critical for deconvolution techniques, as no other information is utilised.

The wide choice of partial volume correction algorithms leads to difficulty in choosing the most appropriate algorithm for oncological imaging. In PET, although accuracy is important, because many patients are assessed longitudinally for treatment response and/or disease progression, test-retest reproducibility is equally important. Hoetjes et al. (83) evaluated three approaches of partial volume correction for accuracy and reproducibility in simulated and patient data. The study compared an iterative deconvolution approach, correction as part of iterative reconstruction, and a mask-based approach, and found reconstruction-based PVC performed best, although each of the other two approaches were adequate and easy to implement alternatives. Of course, it is assumed that PVC
provides improved quantitative performance. Although this may be the case, it does not necessarily lead to improved diagnostic outcomes (87, 88).

**Neurology**

In PET, the widest use of partial volume correction techniques is seen in neurological applications. The size and thickness of areas of interest in the brain together with the relative ease of image segmentation and image registration have led to a plethora of literature describing various methodologies and their clinical application. Since its initial conception and application in $^{11}$C-carfentanil PET by Meltzer et al. (26), a simple segmentation approach of dealing with atrophy related signal losses has been popular. Ibáñez et al. (89) used the method to show that hypometabolism as seen in FDG PET is a real phenomena, and not something caused by cortical atrophy related signal loss. Following on from this, Meltzer et al. (90) and Ibáñez et al. (91) using the same technique also found that apparent reductions in cortical glucose metabolism with increasing age were not real, but actually caused by normal age related increases in cortical atrophy. Refinements to the Meltzer technique have also been developed. Adapting the approach to include a spatially variant PSF, Labbé et al. (92) applied their technique in the analysis of 4 control and 8 Alzheimer's disease patients imaged using $^{18}$F-FDG PET. However, the most popular variant of the original Meltzer method, which replaced a single brain tissue compartment with individual grey and white matter compartments was that proposed by Müller-Gärtner et al. (27). Giovacchini et al. (93), using dynamic imaging of 8 young and 7 old subjects confirmed with the application of both Meltzer and Müller-Gärtner methods that blood flow was not affected by healthy ageing. Conversely, looking at a different biomarker, Bencherif et al. (94) imaging 14 healthy control subjects with $^{11}$C-carfentanil showed an age effect in uptake with partial volume correction, which disappeared when no correction was applied. In turn, the Müller-Gärtner method has also seen slight revisions (95) which have been applied in an FDG PET aging study (96). The relative advantages of two (Meltzer) or three (Müller-Gärtner) segmentations has also been explored (97). Although three segmented
compartments lead to slightly higher accuracy, they can be less robust than two compartments due to the mis-registration and segmentation errors that can prove problematic for these techniques.

There are also other PVC techniques that have been applied to clinical data. Goffin et al. (98) found benefits in applying anatomical based prior MAP reconstructions to reduce partial volume effects, when trying to localise focal cortical dysplasias in epileptic patients imaged with FDG. While in the area of Alzheimer’s disease, Thomas et al. (16) compared the Van-Cittert and RBV method together with the Müller-Gärtner method in a series of 70 amyloid PET scans, with the RBV method performing best out of the three methods. Another commonly used method of partial volume correction in neurological PET is the GTM method (21). In 2008, Rousset et al. (99) applied this method to 90 $^{11}$C-raclopride studies of healthy volunteers, showing increased and also less heterogeneous uptake (in terms of caudate vs. putamen) when compared to uncorrected studies.

The application of PVC techniques is becoming more widely reported in clinical dementia populations, where atrophy is clearly a common issue. Drzezga et al. (100) used commercialised Müller-Gärtner PVC software for correction of $^{11}$C-PIB and FDG studies of patients with Alzheimer’s disease and semantic dementia. The same group also used the technique in a longitudinal study of 15 patients, again with $^{11}$C-PIB and FDG in an attempt to follow the changes in Alzheimers disease over a period of two years (101). Further longitudinal amyloid PET studies by Su et al. (102) with the GTM method and Meltzer method, and Brendel et al. (103) using just the GTM method have also recommended the use of partial volume correction in these types of studies, with the GTM providing slightly better results than the Meltzer method.

Once more, one of the concerns with applying these techniques in longitudinal studies is the test-retest reproducibility of these methods with issues typically arising from errors in image registration and segmentation. In the study by Su et al. (102), segmentation errors led to test-retest variability of around 1.5%, which compared well to the 2% errors from the volumes of interest used to derive regional uptake. However it would appear that the use of different segmentation algorithms can
lead to different results in the same patients, suggesting that segmentation algorithms are not interchangeable when performing partial volume correction (9, 10). Nevertheless, the recent swell in the use of partial volume correction in dementia PET studies suggests that partial volume correction may become mainstream in these studies. See box 3 for a clinical example.

**Vascular**

Beyond the derivation of corrected image derived input functions in vessels (see article by Lepage in this issue), there have also been studies looking at correcting uptake in features within the vessel itself. Izquierdo-Garcia et al. (104) used a GTM methodology to correct uptake in plaques in 7 patients imaged with FDG and found corrected data to be both accurate and reproducible in the carotid artery wall. Reeps et al. (105) looked at performing partial volume correction in abdominal aortic aneurysms in 23 patients undergoing FDG PET imaging. Using a novel approach which created a geometric model of the area of interest, and a factor that accounted for the spatial resolution of the PET system, the study found that partial volume correction was not mandatory for such imaging, but may provide a role in helping to stratify patients. Burg et al. (106) compared various methods for quantifying tracer uptake in the aortic vascular wall in 18FDG PET studies using simulations based on an anthropomorphic digital phantom. They found that PVC using the GTM method gave the most reproducible results, although some bias remained due to the low dimensions of the vascular wall. Blomberg et al. (107) evaluated the impact of a novel PVC method, based on contrast-enhanced CT data, on quantification of arterial wall 18F-FDG uptake. The PVC method used was similar in principle to the Hofheinz method (80). Their study showed that PVC significantly influences quantification of arterial wall 18F-FDG uptake. On the other hand, they also found a good correlation between the values before and after PVC at 180 min post-injection, and concluded that the un-corrected values could be used as a measure of arterial 18F-FDG uptake (107).
Clinical example – neurology

Figure 5 shows images from a clinical study on a patient with semantic dementia and left temporal lobe atrophy. PET studies were performed with two different tracers: $^{18}$F-FDG for glucose metabolism, and the amyloid tracer $^{18}$F-AV45. Images are shown without PVC and with PVC using the iterative Yang method. (1) The two tracers have different distributions. The AV45 scan was amyloid negative, with mainly non-specific uptake in WM, while the FDG study shows mainly uptake in GM.

**Figure 5**: Data from a clinical study with $^{18}$F-FDG and $^{18}$F-AV45 on the same patient, including uncorrected FDG study (1st row), uncorrected AV45 study (2nd row), segmented/parcellated MRI scan (3rd row), corrected FDG study (4th row) and corrected AV45 study (5th row), with sagittal (left column), transaxial (middle column) and coronal (right column) sections. Each image volume was normalized independently.
Discussion

There is increasing evidence that PVC is necessary in order to guarantee quantitative accuracy in emission tomography but there is reluctance to apply PVC routinely in clinical practice where there continues to be reliance on visual interpretation for many routine studies. There continues to be uncertainty by many clinicians regarding the method of choice and the absence of widely accepted and validated commercial software does little to encourage more widespread use. There also continues to be concern regarding the various potential sources of error (e.g. Frouin et al. (8)). Most PVC methods rely on having accurately registered anatomical data (preferably MRI) but the lack of ready availability of these data and historical problems with cross-vendor compatibility have been barriers to routine use; the availability of PET/MRI systems certainly removes (or at least reduces) these barriers. As mentioned earlier registration is not entirely guaranteed with PET/MRI but the degree of mis-registration is certainly reduced and non-rigid registration is unlikely to be needed. Of course the underlying assumption in many PVC methods is that the anatomical edge information matches the tissue boundaries underlying the functional or molecular image, which is not necessarily the case (e.g. there may be cellular invasion of tissues not visible in the anatomical image). Note also that non-rigid registration between high resolution MRI and the limited resolution PET can attempt to align identified edges in a way that invalidates the PVC procedure. There are similar concerns regarding segmentation approaches that are often necessary, especially where structures are very small (e.g. blood vessels). A small error in segmentation may represent a sizeable percentage error in estimated volume, and hence a significant source of error. Approaches that avoid the need for segmentation may be preferred, for example techniques that utilise local variation in MR contrast directly (39, 45, 51).

In some cases it can be argued that PVC makes little difference to interpretation or indeed can have a negative influence due to enhanced noise. However, in situations where volume is changing in serial studies, misinterpretation is very common. Also if activity distribution changes in time, as is
clearly the case in kinetic studies, there can be significant errors introduced by PVEs. Take the example of a dynamic cardiac study where activity is initially circulating in the blood with significant spillover to the myocardium but later may be concentrated in the myocardium itself. Variation in the spillover will clearly affect the resulting time activity curves for both the ventricle (and image-derived input function) and myocardium, affecting any derived kinetic parameters. A recent paper by Su et al. (102) highlights the importance of performing PVC in kinetic amyloid studies which are being used to further elucidate uptake patterns in patients with dementia. Other groups have also shown the importance of PVC for kinetic analysis (8, 108, 109). PVC is also important for accurate estimation of image derived input functions (110). Further information on this topic can be found in the article by Lepage in this issue.

An aspect of PV correction less commonly addressed is the influence of the ‘tissue fraction’ effect, in general this accounts for the uptake of tracer in selected targets contained within a voxel (e.g. uptake in the cellular matrix rather than blood or air in the lung). In fact heterogeneous uptake of tracer is invariably the case within the volume of tissue contained in a typical image voxel. Kinetic analysis often does take into account a vascular component but again problems arise if the non-target tissue fraction varies spatially or temporally. Research on methods to tackle this problem, specifically applicable to lung studies, has been published (Lambrou et al. (111), Holman et al. (112)). This work relies on identifying air fraction from CT density and/or blood fraction based on kinetic analysis as a basis for deriving tissue fraction corrected values which may provide better indices of the underlying mechanism of uptake in disease or treatment response during therapy.

There is a distinct shortage of published work on the direct comparison of different PVC methods. Validation is not a trivial exercise as it requires ground truth knowledge of tracer distribution which is not possible in vivo. Comparison of discrimination power is commonly used to justify choice of individual methods although in practice the effect can be small (e.g. Thomas et al. (16)). The alternative approach is based on simulation. The challenge is to implement simulation methods that
are clinically realistic, that use analysis techniques that are independent of the methods used for PVC and that provide meaningful metrics for comparative purposes. Ideally test datasets should be made openly available so that PVC can be tested by individual developers, in much the same way as the clinical datasets made available for evaluation through ADNI (113). Development of such a library has been the aim of work undertaken in conjunction with the EU COST project (TD1007) with suggested approaches and preliminary results published by Hutton et al. (114) and Thomas et al. (115).

**Acknowledgements**

We would like to thank Dr Jonathan Schott, Dementia Research Unit, UCL for providing the clinical brain data included. The amyloid and FDG PET images shown were from a study supported by AVID Radiopharmaceuticals (a wholly owned subsidiary of Eli Lilly) and the The National Brain Appeal - Frontotemporal Dementia Research Fund. This work was supported by by the EPSRC (grant number EP/K005278/1), the EC under the FP7 INSERT project (Grant number 305311), and the National Institute for Health Research University College London Hospitals Biomedical Research Centre.
References


101. Forster S, Yousefi BH, Wester HJ, Klupp E, Rominger A, Forstl H et al. Quantitative longitudinal interrelationships between brain metabolism and amyloid deposition during a


