A Monte Carlo Study of the Dependence of Early Frame Sampling on Uncertainty and Bias in Pharmacokinetic Parameters from Dynamic PET

Ida Håggström¹, Jan Axelsson¹, Charles Ross Schmidtlein², Mikael Karlsson¹, Anders Garpebring¹, Lennart Johansson¹, Jens Sörensen³, and Anne Larsson¹

¹Department of Radiation Sciences, Umeå University, Umeå, Sweden; ²Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, New York; and ³Nuclear Medicine and PET, Department of Radiology, Oncology and Radiation Sciences, Uppsala University, Uppsala, Sweden

Compartmental modeling of dynamic PET data enables quantification of tracer kinetics in vivo, through the calculated model parameters. In this study, we aimed to investigate the effect of early frame sampling and reconstruction method on pharmacokinetic parameters obtained from a 2-tissue model, in terms of bias and uncertainty (SD). Methods: The GATE Monte Carlo software was used to simulate 2 × 15 dynamic 3-deoxy-3-[18F]-fluorothymidine (18F-FLT) brain PET studies, typical in terms of noise level and kinetic parameters. The data were reconstructed by both 3-dimensional (3D) filtered backprojection with reprojection (3DRP) and 3D ordered-subset expectation maximization (OSEM) into 6 dynamic image sets with different early frame durations of 1, 2, 4, 6, 10, and 15 s. Bias and SD were evaluated for fitted parameter estimates, calculated from regions of interest. Results: The 2-tissue-model parameter estimates \( K_1, K_2 \), and fraction of arterial blood in tissue depended on early frame sampling, and a sampling of 6–15 s generally minimized bias and SD. The shortest sampling of 1 s yielded a 25% and 42% larger bias than the other schemes, for 3DRP and OSEM, respectively, and a parameter uncertainty that was 10%–70% higher. The schemes from 4 to 15 s were generally not significantly different in regards to bias and SD. Typically, the reconstruction method 3DRP yielded less frame-sampling dependence and less uncertain results, compared with OSEM, but was on average more biased. Conclusion: Of the 6 sampling schemes investigated in this study, an early frame duration of 6–15 s generally kept both bias and uncertainty to a minimum, for both 3DRP and OSEM reconstructions. Very-short frames of 1 s should be avoided because they typically resulted in the largest parameter bias and uncertainty. Furthermore, 3DRP may be preferred over OSEM for short frames with poor statistics.

Key Words: dynamic PET; Monte Carlo; GATE; compartment modeling; frame sampling

DOI: 10.2967/jnmt.114.141754
3'-deoxy-3',18F-fluorothymidine (18F-FLT), have proven successful alternatives to 18F-FDG for brain tumor studies (11–14). This work focused on the tracer 18F-FLT, which can be used to image tumor proliferation (11–14).

In dynamic PET, the tracer distribution over time is observed. The collected projection data are sorted into frames from which the kinetics of the tracer in a tissue or organ of interest can be quantified. Commonly, the dynamic data are analyzed using a tracer-specific compartment model, yielding a set of model-determined parameters representing the tracer kinetics. For dynamic PET imaging, list-mode data can be collected and binned to arbitrary frame durations. When reconstructing the data into a dynamic image set, the user is faced with the choice of using longer frames with better counting statistics but poor temporal resolution or shorter frames with poor counting statistics but better temporal resolution.

Short frames may be of interest during the first minutes of the dynamic PET study to better capture the fast variations in the tracer uptake and clearance, which are usually largest for early frames, and for a better definition of the early blood peak (6). This is especially true for image-derived TACs of the blood, because they are sharper than the TACs of the tissue and thus require higher sampling rates (7). Clinical PET frames are often sampled at intervals ranging from 10 to 300 s, although frames shorter than 5 s exist but are rare. Although short frames are often desired in dynamic PET imaging, they are seldom used because of the poor statistics associated with short frames and poor quality and bias of image reconstructions of such frames (6).

The aim of this study was to investigate the effects of different sampling schemes (frame durations) for early frames have on pharmacokinetic parameter estimates in regards to bias and SD for typical dynamic 18F-FLT brain studies with typical dose administrations (noise level). The aim was also to investigate the effect of the reconstruction method used for the dynamic PET data by comparing the results for analytic 3-dimensional (3D) filtered backprojection with reprojection (3DRP) and 3D ordered-subset expectation maximization (OSEM). In order to obtain data, the MC software GATE was used to simulate 2 × 15 separate dynamic PET studies with a digital head phantom with added tumors, representing 18F-FLT with realistic kinetic parameter values.

### MATERIALS AND METHODS

#### Compartment Model

The 2-tissue-compartment model seen in Figure 1 is commonly used to describe 18F-FLT (15). The TAC describing the activity concentration in arterial blood plasma is known as the input function \( C_p \), and the 2 TACs describing the tissue concentrations, \( C_{p+NS} \) and \( C_S \), represent free-plus-nonspecific (nondisplaceable) and specifically bound tracer in tissue, respectively. The model parameters that govern the tracer exchange are the uptake rate from blood to nondisplaceable tracer in tissue \( K_1 \) (mL cm\(^{-3}\) min\(^{-1}\)), the clearance rate from nondisplaceable tracer in tissue \( k_2 \) (min\(^{-1}\)), and the rates between the nondisplaceable and specifically bound tracer tissue compartments \( k_3 \) (min\(^{-1}\)) and \( k_4 \) (min\(^{-1}\)), respectively.

### FIGURE 1.

2-tissue-compartment model consisting of compartments of arterial blood plasma (\( C_p \)), free-plus-nonspecific (nondisplaceable) and specifically bound tracer in tissue (\( C_{p+NS} \) and \( C_S \)), 4 rate constants (\( K_1, k_2, k_3, k_4 \)), and tissue blood fraction (\( V_a \)). \( C_{PET} \) is the apparent concentration in a PET image VOI or voxel.

The activity measured by the PET camera is the sum of \( C_{p+NS} \) and \( C_S \) plus an added blood component to account for signal contamination from blood vessels within the voxel or a volume of interest (VOI) (8,16,17).

#### Monte Carlo Simulation

GATE (18) was used to perform MC simulations of 3D dynamic PET 18F-FLT brain scans. The camera simulated was a previously validated Discovery LS PET scanner (GE Healthcare) (19), consisting of 18 detection rings with 672 bismuth germanate crystals of approximately 4 × 8 × 30 mm each. The transaxial field of view was 550 mm, and the axial field of view 152 mm.

The phantom used in the simulations was the voxelized digital BrainWeb head phantom (20). The phantom was simulated and described in a previous study (21) and is briefly described here. The phantom consisted of 7 main materials, with 7 uniform spheric tumors (diameters, 3, 6, 9, 12, 18, 24, and 30 mm) distributed in each hemisphere (14 inserted tumors). Finally, a 25-mm spheric blood region was placed centrally in the brain. The constructed phantom had an isotropic voxel size of 1 × 1 × 1 mm and is depicted in Figure 2. The size, shape, and location of the blood region was designed to be practical for extraction of an image-derived input function of small variance and little influence of partial-volume effects, rather than to be realistic. Furthermore, the inclusion of a blood region within the phantom eliminated the need for additional simulations (e.g., the heart) for input-function derivation.

Realistic pharmacokinetics were assigned to all normal phantom tissues, and fitted TACs from 2 clinical dynamic brain 18F-FLT scans performed at Umeå University Hospital were used for those regions. The same input function \( C_p \) was assigned to the blood region in both setups, denoted FLT\(_1\) and FLT\(_2\). \( C_p \) was generated using Matlab (version 8.1.0; The MathWorks Inc.) and had a typical input function appearance (22), with a realistic peak amplitude of around 50 kBq/mL\(^3\) (measured at Umeå University Hospital). For each of the 2 setups, Matlab was used to
generate the corresponding tissue TAC (TTAC) CPET, according to the 2-tissue model with the realistic \(^{18}\text{F}-\text{FLT}\) parameter values for gliomas, seen in Table 1 \((12,13)\). For both FLT1 and FLT2, all 14 tumor regions were assigned the same TTAC CPET. The simulated blood and tumor TACs are seen in Figure 3. The source particles simulated were \(^{18}\text{F}\) positrons with an electron range production cut of 2 mm, a \(\delta\)-ray production cut of 10 keV, and an x-ray production cut of 10 keV \((19,23)\). Physical decay of the sources was applied with a half-life of 6,586.2 s. The activities in the tissue and blood regions of the phantom were read from the generated TACs and updated every second of the simulation. For practical reasons, the detector dead time was not included in the simulation. The total GATE acquisition time was set to 60 min, and simulated data were stored in list-mode. Thirty MC simulations were performed: 15 of the setup with FLT1 and 15 with FLT2, to improve the statistics for the kinetic parameter analyses. Although simulated, the random coincidences were not included in the study. The effect of random coincidences was considered small because the random fraction for the 2 setups was merely 2%.

As previously described by Håggström et al. \((21)\), data from 1 normalization and 1 calibration simulation were also used for the image reconstruction normalization and quantitative calibration. All simulations were performed using the computer cluster Akka (Intel Xeon quad-core L5420 central processing units) at the HPC2N collaboration, Umeå University. Each of the 30 dynamic simulations required about 3,300 central-processing unit hours.

### Image Reconstruction

The list-mode true plus scattered coincidences were binned into 3D sinograms \((19)\) and reconstructed using 2 methods: analytic 3DRP \((24)\) and OSEM \((25)\) iterative reconstruction. The reconstructions were performed using the software STIR \((26)\). A Colsher filter (cutoff frequency, 0.5 pixel\(^{-1}\)) was applied for 3DRP, and the OSEM settings were chosen so that the tumor VOI values had reached convergence at 60 subiterations and 12 subsets (5 iterations). Both methods had normalization, decay, attenuation, and scatter corrections applied.

For the attenuation correction, a linear attenuation coefficient data map (\(\mu\)-map) relevant for 511-keV photons for the respective phantom materials was generated from the BrainWeb phantom. The normalization simulation data were binned into a sinogram to create the normalization sinogram \((27)\). A scatter sinogram estimate was created from the single scatter simulation algorithm implemented in STIR \((28,29)\) and used as additive sinograms in the OSEM loop. Attenuation and normalization data were also included in the loop, whereas all 3 corrections were used as precorrections for 3DRP.

In accordance with clinical settings of the Discovery LS (GE Healthcare), all reconstructed images were postfiltered with a gaussian transverse filter of 6.0 mm in full width at half maximum and a 3-point smoothing filter \([1 2 1]/4\) in the axial direction. Reconstructed image sizes were 165 × 165 × 35 voxels, with a voxel size of \(2 \times 2 \times 4.25\) mm.

Finally, a scale factor to calibrate all images from counts to Bq/mL was created from a 3DRP reconstruction of the true coincidences from the calibration simulation.

The dynamic PET data were reconstructed into 6 dynamic image sets, with the first 2 min (covering the early blood peak) sampled at \(120 \times 1, 60 \times 2, 30 \times 4, 20 \times 6, 12 \times 10, \) or \(8 \times 15\) s. The PET data between 2 and 60 min were all sampled at \(2 \times 30, 2 \times 60, 2 \times 150, \) and \(10 \times 300\) s.

### Model Fitting and Parameter Analysis

The input function was image-derived from a spheric VOI covering the complete blood region (25-mm diameter, 488 voxels, 8.3 mL), and 2 tumor TTACs from each simulation were derived from complete VOIs of the largest left and right tumors, respectively (30-mm diameter, 843 voxels, 14.3 mL). In this study, only the 2 largest 30-mm tumors were used to minimize partial-volume effects (the additional 12 tumor spheres are intended for another study). The data used for analysis thus comprised 15 input functions and 2 TTACs per input function, making a total of 30 tumor TTACs for each of FLT1 and FLT2, for which the TAC values were calculated as the mean value of the VOI at each frame.

Kinetic parameter estimates were obtained by nonlinear least-squares (NLS) fitting of each of the TTACs with the input function considered a noiseless model input. The NLS fitting is commonly weighted, and each TTAC value should be weighted according to its inverse variance. Because the true variance is typically unknown, it is usually approximated by considering radioactive decay, frame duration, and often also frame count. However, weighting according to noisy counts can degrade the parameter estimation severely \((30,31)\). Furthermore, in order not to force a fit only to the last few TTAC values with long frame durations and in essence ignore the short early frames, a uniform weight was used for all frames in this study. The Matlab function lsqnonlin was used for the fitting, and the true values were used as start values to avoid any effects from the choice of initial parameter guesses. The mid time of each frame was used as the time point, and the influx rate parameter \(K_i\) was calculated according to Equation 1 for each VOI.

The 30 sets of kinetic parameters were finally averaged into 1 single set of estimated \([K_1, k_2, k_3, k_4, V_a, K_i]\) for each of the 6 sampling schemes, for both FLT1 and FLT2. The relative bias of all 6 pharmacokinetic parameter estimates \(P\) was calculated as:

### Table 1

<table>
<thead>
<tr>
<th>Parameter set</th>
<th>(K_1)</th>
<th>(k_2)</th>
<th>(k_3)</th>
<th>(k_4)</th>
<th>(V_a)</th>
<th>(K_i^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLT1</td>
<td>0.071</td>
<td>0.091</td>
<td>0.047</td>
<td>0.018</td>
<td>0.066</td>
<td>0.024</td>
</tr>
<tr>
<td>FLT2</td>
<td>0.111</td>
<td>0.131</td>
<td>0.017</td>
<td>0.012</td>
<td>0.122</td>
<td>0.013</td>
</tr>
</tbody>
</table>

*Values calculated by Equation 1.*
RESULTS

In this study, we investigated the effect of early frame duration on bias and SD in pharmacokinetic parameter estimates obtained from the 2-tissue-compartment model for typical $^{18}$F-FLT brain studies.

Two sets of kinetic parameter value sets were chosen, representing the tracer $^{18}$F-FLT, but because the simulations themselves were parameter-specific, not tracer-specific, the results are likely valid for any tracer suited for the 2-tissue-compartment model with similar pharmacokinetic parameter

<table>
<thead>
<tr>
<th>Parameter set</th>
<th>1 s</th>
<th>2 s</th>
<th>4 s</th>
<th>6 s</th>
<th>10 s</th>
<th>15 s</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLT$_1$</td>
<td>0-44</td>
<td>0-87</td>
<td>0-173</td>
<td>3-258</td>
<td>125-429</td>
<td>282-642</td>
</tr>
<tr>
<td>FLT$_2$</td>
<td>0-57</td>
<td>0-113</td>
<td>0-226</td>
<td>4-338</td>
<td>162-564</td>
<td>368-843</td>
</tr>
</tbody>
</table>

DISCUSSION

On average, the total registered and kept coincidences from the 15 simulations was $119 \times 10^6$ and $137 \times 10^6$, for FLT$_1$ and FLT$_2$, respectively. The total kcounts per frame for the early frames (first 120 s) of the different sampling schemes are seen in Table 2.

Examples of the reconstructed images are seen in Figure 4, and representative TTACs from 1 of the 30 FLT$_1$ tumor VOIs, with corresponding NLS fit, are seen in Figure 5. As can be seen in both figures, the noise level of the images and subsequent TACs decrease as the frame duration increases. The calculated bias and SD of the parameter estimates from the NLS fits of the image-derived input

$$\frac{(P - P_{true})}{P_{true}} \times 100\%,$$

where $P_{true}$ is the true parameter value. The accompanying relative SE in the bias was

$$\frac{s_P}{\sqrt{30}} \times \frac{1}{P_{true}} \times 100\%,$$

where $s_P$ is the SD of parameter estimate $P$. The relative SD in the parameter estimates was calculated according to

$$s_P \times \frac{1}{P_{true}} \times 100\%.$$
values and noise level (dose administration). Two image reconstruction methods were also studied, 3DRP and OSEM.

Of the statistically significant results, the bias was generally smallest for early frame durations of 6–15 s. A frame-sampling dependence was generally found for \( K_1, k_2, V_a \), and more so for OSEM than 3DRP. The closer statistical analysis revealed that it was the 1-s, and occasionally also the 2-s, scheme that stood out from the rest. With 3DRP, it was only for the faster kinetics described by FLT2 that a frame-sampling dependence was found for \( K_1 \) and \( k_2 \). The slower kinetics in FLT1 did not show a significant sampling dependence. These results are not surprising because faster kinetics should intuitively be more prone to undersampling than slower kinetics for which the TACs do not vary as quickly from frame to frame. Parameter \( K_1 \) mainly governs the amplitude of the TTAC, whereas \( k_2, k_3 \), and \( k_4 \) together contribute to the shape of the TTAC. This makes predictions regarding individual parameter responses to changes in the TTAC shape difficult. Apart from the highest bias, the SDs were also highest for the shortest frames of 1 s, which is expected because the signal-to-noise ratio of the frames is roughly proportional to the square root of the number of counts.

The bias for OSEM appeared to have a frame duration dependence, which was more pronounced than for 3DRP, for parameters \( K_1 \) and \( k_2 \). It is known that iterative reconstruction (OSEM) of low-count images may result in (positively) biased images whereas analytic reconstruction methods (3DRP) do not (32,33). The bias introduced due to low count reconstruction may explain the larger bias for short frames and hence more prominent frame duration dependence for OSEM, compared with 3DRP. The parameter \( K_1 \), which mainly controls the peak amplitude of the TTAC, would thus be most affected by a biased ROI value. Compared with OSEM, 3DRP images have a high background noise and streak artifacts (Fig. 4). However, the lowest-count OSEM images (1- and 2-s frames) appear to have artifacts in the form of high uptake spots. The effect of the low count bias for short framed OSEM images and how it transfers to the parameter biases and SDs is hard to include properly in studies in which the TACs are simulated directly with added noise profiles. This effect is more realistically depicted using full MC simulations with complete image reconstructions.

Furthermore, according to typical practice, the same setting (here 60 subiterations, 12 subsets) was used for all frame-sampling schemes and for all frames of each scheme. However, because of the difference in statistics between the different frames, they should ideally (despite practicality issues) have the OSEM reconstruction settings optimized individually. Furthermore, OSEM images have noisier hot regions than 3DRP, and vice versa for cold regions, as known from literature (9,34). Cerebellum VOI (cold) SDs and tumor VOI (hot) SDs were on average 71% higher and 28% lower, respectively, when 3DRP was compared with OSEM. This fact, together with the low count bias, may explain the 15% larger parameter uncertainties obtained for OSEM than 3DRP images. However, the noise level in OSEM images is heavily dependent on the number of iterations, and other reconstruction settings could yield slightly different results.

It is apparent that \( V_a \) is dependent on the frame duration. Even the theoretic TACs (Fig. 7) resulted in large biases for the longer sampling intervals. Because \( V_a \) is determined by the early blood peak seen in the TTAC, a long frame duration will effectively lower the peak due to smearing and
cause an underestimated $V_a$. An input function with a wider peak would most likely decrease the frame-sampling dependence and yield less bias in $V_a$. On the other hand, a more narrow blood peak would likely cause $V_a$ to be more underestimated.

Both $k_3$ and $K_i$ were found independent of early frame sampling. Even though Figure 6 shows a larger bias for the shortest 1-s frame sampling, the results were not significant. These 2 parameters have been shown to be of potentially larger clinical value (13,14), so a small dependence on frame-sampling scheme is desired. Because $K_i$ is a macro-parameter calculated from 3 other parameters, it is likely to be more stable than single parameters (6). The image-derived $k_3$ was not different from the theoretic noiseless case. Thus, the effects of, for example, the camera system blurring and reconstruction process did not add additional bias to the estimate, making $k_3$ reliable and strengthening its role as a clinically relevant parameter.

The bias and SD of parameter $k_4$ were large. A longer scan time (>60 min) would likely improve the $k_4$ estimation (15), but as this parameter is rarely considered clinically important because of large uncertainties, we chose not to take this into further consideration. Jovkar et al. (8) found that keeping $k_4$ fixed in the fitting procedure resulted in more stable estimates of all other parameters.

According to our findings, noise and bias in the image-derived TACs can affect the parameter uncertainty and bias to a large extent. Because the input function is derived from only 1 ROI and assumed true for the fitting procedure, any bias in it will affect all parameter estimates. If the chosen blood ROI is not representative, the calculated bias and SD in the parameter estimates may be largely affected. To minimize uncertainty and bias, the use of a population-based input function might be helpful (22,35). However, the population-based input function is subject to interpatient variability and may be biased in itself.

The parameter bias was found larger for 3DRP than OSEM and the uncertainty smaller on average. As shown by Thiele et al. (31) and Yaqub et al. (30), the choice of weight factors for weighted NLS can affect the results to a relatively large extent.

FIGURE 6. Relative error (bias) (A) and uncertainty (SD) (B) in all parameter estimates for 6 different early frame-sampling schemes. Error bars in bias represent SE. $y$-axes are scaled differently for better visibility.
where \(3DRP\) showed minimum bias for reconstructions benefited from uniform weighting, considerably better estimates than \(OSEM\). The bias for \(OSEM\) in this study, the focus was to evaluate a typical \(18F\)-FLT study and because a larger ROI or higher activity yields better statistics. In general, the parameter uncertainties were generally stable for frames of 2–15 s but higher for the short 1-s sampling.

The estimation of \(k_3\) was most reliable (bias < 5% in general), and the parameter \(V_a\) was overall most dependent on frame duration.

Despite the visually more appealing \(OSEM\) images, the analytic image reconstruction method \(3DRP\) showed less dependence on early frame sampling, compared with the iterative technique \(OSEM\). Moreover, \(OSEM\) reconstructions of the shortest early frames (1 and 2 s) had artifacts in the form of high uptake spots, not seen in the \(3DRP\) images.

DISCLOSURE

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734. This work was supported in part by the Swedish Cancer Society and the Cancer Research Foundation at Umeå University. No other potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

We thank Prof. Jun Yu from the Department of Mathematics and Mathematical Statistics at Umeå University for his support on statistical calculations.

REFERENCES

A Monte Carlo Study of the Dependence of Early Frame Sampling on Uncertainty and Bias in Pharmacokinetic Parameters from Dynamic PET

Ida Häggström, Jan Axelsson, Charles Ross Schmidtlein, Mikael Karlsson, Anders Garpebring, Lennart Johansson, Jens Sørensen and Anne Larsson

Published online: January 22, 2015.
Doi: 10.2967/jnmt.114.141754

This article and updated information are available at: http://tech.snmjournals.org/content/43/1/53

Information about reproducing figures, tables, or other portions of this article can be found online at: http://tech.snmjournals.org/site/misc/permission.xhtml

Information about subscriptions to JNMT can be found at: http://tech.snmjournals.org/site/subscriptions/online.xhtml