Reliable Blood Data for Full PET Quantification

PMOD APPLICATION NOTE

PMOD is a software
FOR RESEARCH USE ONLY (RUO)
and must not be used for diagnosis or treatment of patients.

1 Reliable Blood Data for Full PET Quantification

It has become widely recognized that qualitative imaging can be superseded by absolute quantification of tissue parameters for the understanding of biological systems. One of the gold standard methods in quantification is PET, which has sensitivity down to the picomolar range and a unique specificity due to the use of targeted molecular probes.

PET studies intended for full quantification require two sources of data, the reconstructed PET images measuring the tracer activity concentration in tissue over time, and the activity of the tracer and its metabolites in blood. While much effort is spent ensuring that PET data is accurate, the acquisition and processing of blood data is often treated as an inconvenience, despite the fact that it is crucial for the quality of the outcome.

The main information to be extracted from the blood data is the concentration of the authentic (unchanged) tracer in arterial plasma throughout the PET acquisition, as it represents the tracer available for uptake into tissue. For the quantification procedure it represents the input into the system and is therefore called the Arterial Input Function (AIF).

Whole blood consists of red blood cells and plasma with components such as proteins, white blood cells and original tracer as well recirculating tracer metabolites originating from tracer metabolism in e.g. the liver. Therefore whole blood has to be separated into these components in order to assess the authentic tracer concentration in arterial plasma alone.

It is a widely shared opinion that the AIF is the weakest part in full PET quantification. In this application note we therefore discuss the methods to reliably obtain the authentic tracer concentration in arterial plasma and avoid the many pitfalls.

1.1 Components of the Arterial Input Function

Three components of blood data need to be considered for PET kinetic modeling:

1. Whole blood (WB): Radioactivity in the vessels contributes to the signal measured by PET and should be accounted for in quantification. Note that the radioactivity originates from the tracer and its metabolites which are in almost all cases contained in the blood plasma.
2. Authentic tracer (synonyms: unchanged tracer, parent, AIF): The authentic tracer in arterial plasma is the most important element in full PET quantification, as it is exchanging with tissue and interacting with the target. Its concentration has to be determined from the plasma fraction of arterial whole blood.
3. Tracer metabolites (if any): If the tracer is metabolized (e.g. in the liver), recirculating radio-labeled metabolites contribute to the radioactivity measured in plasma. In this case, the relative signal contributions of authentic tracer and metabolites to the total plasma activity have to be assessed by a metabolite analysis.

The radioactivity concentrations of these blood components are time-dependent ("time-activity curves") and have to be assessed with enough temporal resolution.

Radio-labeled metabolites require careful consideration in both target tissue and blood. If these metabolites enter the target tissue they present a major confound and make quantification very difficult. Using brain tracers as an example, most metabolites are more hydrophilic than the authentic tracer, meaning they do not cross the blood-brain barrier. Hence these metabolites increase the measured activity in the blood, but are not present for target binding within brain cells (astrocytes, neurons, microglia).

In practice the following elements are used for converting the radioactivity measured in a whole blood sample into an AIF:

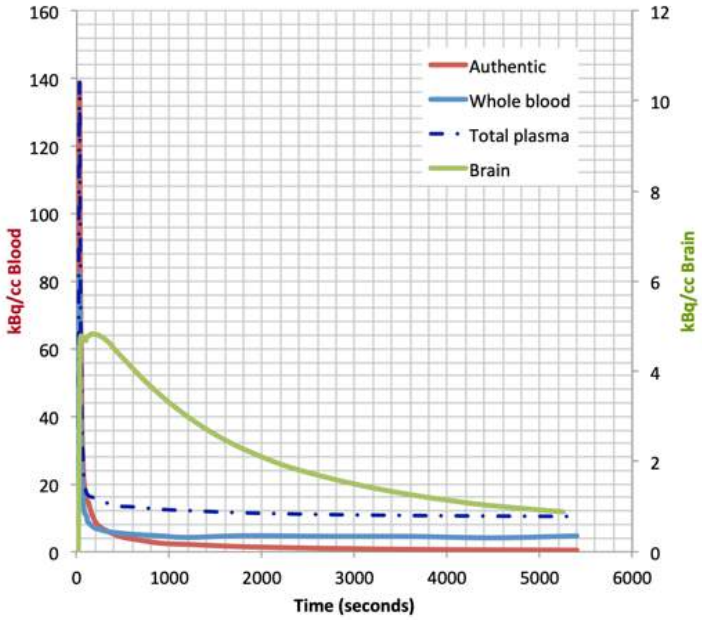
- Whole blood activity curve: The radioactivity concentration in whole blood needs to be assessed by taking physical blood samples throughout the PET acquisition, or alternatively measuring it in big enough blood vessels in the image. Note that the latter (image-derived whole blood curve) can be compromised by partial-volume effects or motion.

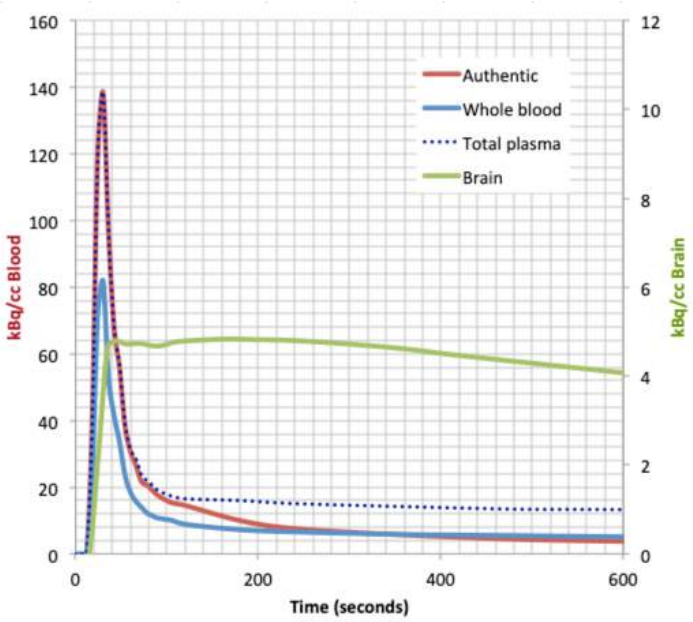
- Plasma/WB ratio: Radioactivity concentration in (total) plasma is usually higher than in blood as a whole. The relation changes over time and can be expressed as ratio of plasma activity concentration divided by whole blood activity concentration. Once this Plasma/WB ratio is known the total plasma activity can be calculated by multiplying the WB activity curve with the Plasma/WB ratio curve. While the Plasma/WB ratio may vary across individuals, for some populations a representative ratio can be assumed.
- Parent fraction: Radioactivity in plasma originates not only from authentic tracer (parent), but also from any radio-labeled tracer metabolites. Again, the relative contributions are variable over time and their relation can be characterized by a concentration ratio, the "parent fraction" defined as activity of parent divided by total plasma activity. It starts at a value of 1 when the tracer is injected and declines as the metabolites build up.

The arterial input function can easily be obtained from these three elements as follows: (1) The measured whole blood activity curve is multiplied by the plasma/WB ratio to obtain the total plasma activity curve. (2) The new total plasma activity curve is multiplied by the parent fraction to obtain the authentic tracer activity curve, the AIF.

Example

The plots from simulated data shown below illustrate the resulting activity concentration curves for whole blood, total plasma (after multiplication with a fixed value plasma/WB ratio), and authentic (metabolite-corrected) tracer (after multiplication with a parent fraction over time determining that less and less of the total plasma activity is authentic tracer). Other examples of plasma/WB ratio and parent fraction will be shown in the relevant sections later.





The process of collecting the data described above, recommendations, and solutions when data is missing will be discussed below.

1.2 Manual Blood Sampling

In the past, the gold standard method to measure the AIF was discrete, manually withdrawn, blood samples. In human subjects, this was typically done via a cannula in the radial artery, and in small animals via a catheter in the femoral artery.

Thus, plasma could be separated from every sample and the concentration of authentic tracer measured (discussed in detail below 12). However, the practicalities of withdrawing blood into syringes or sample tubes for this method means that the realistic maximum temporal resolution is 5-10 seconds. Following IV bolus tracer injection, the true peak concentration could be missed. Additionally, plasma separation and analysis requires a substantial volume of blood. This limits both the temporal resolution achievable, and the number of samples that can be taken. i.e. withdrawing 5 mL of blood from a human or large animal subject takes some seconds, and will likely involve the use of flow-control valves. For a 5 minute PET study of tissue perfusion using O-15-water, an AIF measurement at 5 second temporal resolution, with 5 mL per sample, totals 300 mL of blood. Such a volume may be of clinical significance in human subjects. This volume is intended as a guideline figure.

Pitfalls

As a critical assumption of quantitative PET is a steady physiological state, the loss of this blood volume can be detrimental to the experiment. Sampling with the highest possible temporal resolution when the peak concentration is expected, and a lower resolution thereafter can reduce the blood volume required, but in sensitive subjects or small animals this can remain a problem.

The timing recorded for each sample is prone to errors, and actually represents the duration required to actually collect sufficient blood. It is not generally possible to take duplicate (or triplicate) samples to reduce measurement error, and thus the loss of a single sample leads to gaps in the data. Pipetting errors when transferring blood to well counter tubes are a risk, and the limited half-life of isotopes such as C-11 and O-15 necessitate rapid processing.

1.3 Continuous Blood Sampling

1.3.1 Overview

The AIF should describe the tracer concentration over time as accurately as possible. This implies not only minimization of measurement errors, but also sufficient temporal resolution (sampling rate) to capture rapid changes in concentration.

Bolus Injection of Tracer

Most PET radiotracers are injected as a rapid intravenous "bolus". Although the exact duration used to deliver the bolus varies, this implies a very rapid increase in the blood activity concentration, which will then reduce over the time of measurement due to binding (both specific and non-specific) and metabolism.

Note: A true bolus would be instantaneous. The opposite of a bolus would be a continuous infusion of the tracer. In practice, tracer delivery over 0.5-2 minutes is still commonly known as bolus injection. Many centers use a syringe pump to standardize the rate of tracer delivery. Delivery of the tracer over 2-5 minutes may be referred to as a "slow bolus", and some protocols even combine bolus injection and infusion in order to reach an equilibrium state as quickly as possible. The results of kinetic modeling are independent of tracer delivery as long as the AIF is correct.

Continuous Sampling

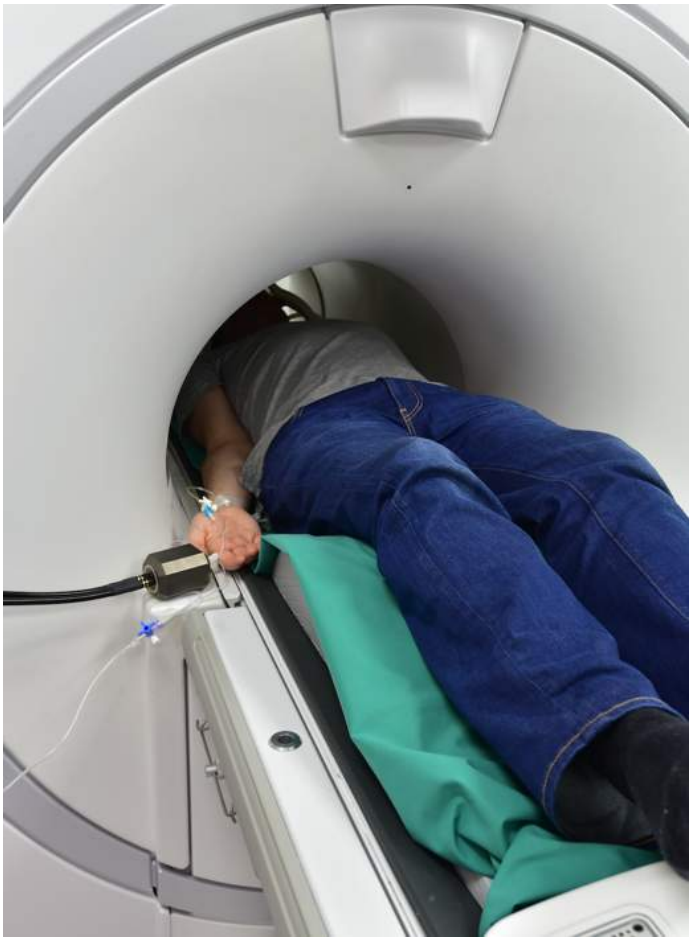
Measurement of the activity in blood flowing continuously through a catheter using small-scale coincidence detectors can yield the whole blood activity concentration with a temporal resolution of 1 second. The flow rate can be used to control the volume of blood withdrawn, or the arterial blood may even be returned to a vein in select species to truly minimize blood loss. This method can be considered the new gold standard for the measurement of peak tracer concentration following bolus injection.

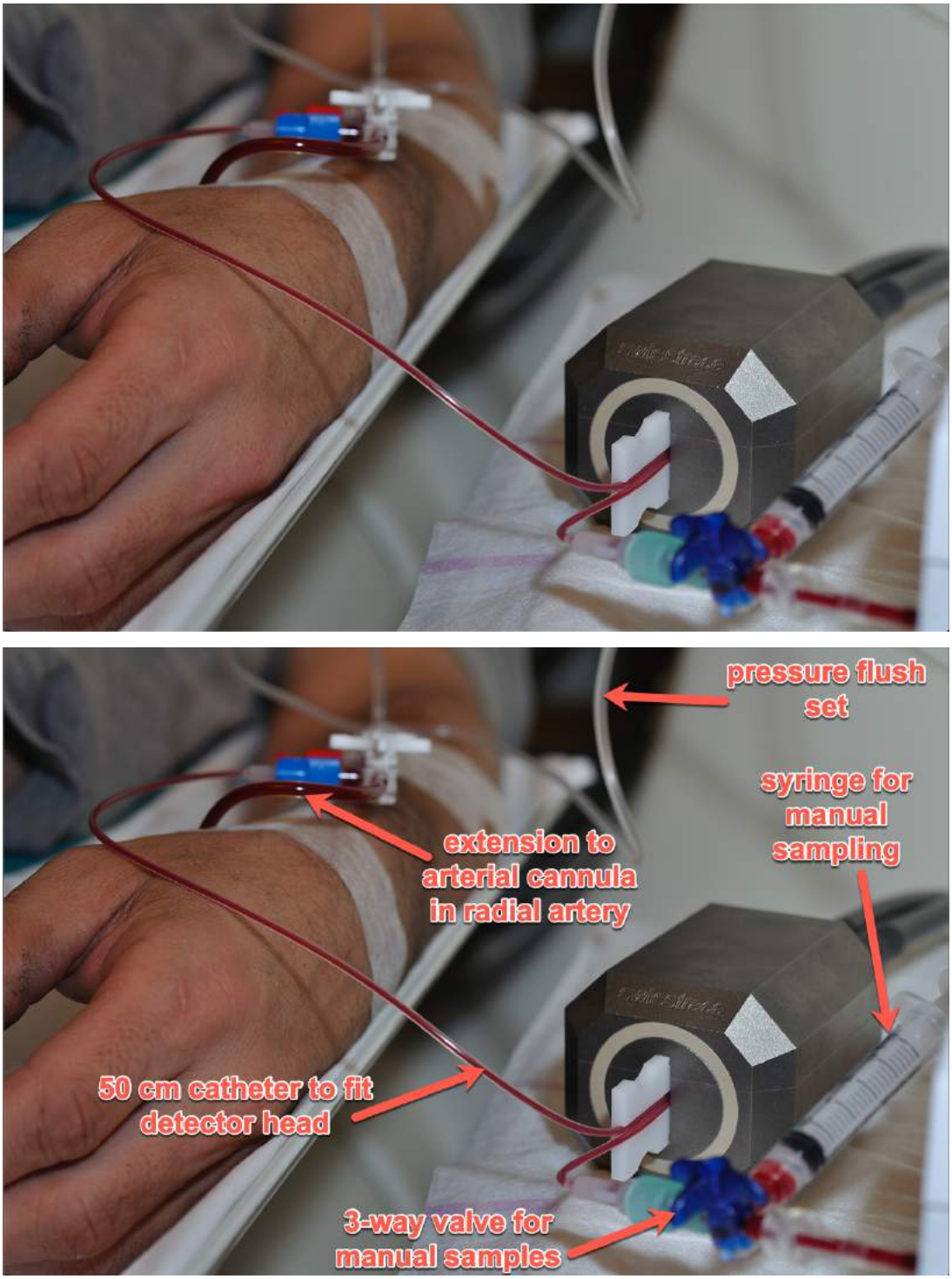
However, some samples must still be withdrawn in order to measure the plasma/WB ratio and parent fraction, a time-delay between tracer delivery to the tissue-of-interest and the detector may be introduced or extended, flow through a catheter may introduce "dispersion" (discussed later) and rapid changes in the plasma/WB ratio in the first minutes after tracer injection may be difficult to detect without disturbing the flow through the detector at this critical stage. Overall, the increased ability to accurately capture the peak activity concentration is considered superior to these drawbacks, and corrections for several of them exist.

The twilite three from swisstrace (www.swisstrace.ch) is an example of a small-scale coincidence detector and used in the illustrations below. It has been applied in several species from mouse to human, for both PET/CT and PET/MR research applications.

1.3.2 Continuous Blood Sampling in Humans

In human research subjects, arterial blood is withdrawn from the radial artery and pumped at a controlled flow rate to waste collection. The catheter passes through the detector head, where counts are collected in 1-second time bins. The addition of a 3-way valve to the system allows manual samples to be withdrawn and used to measure the plasma/WB ratio and parent fraction.



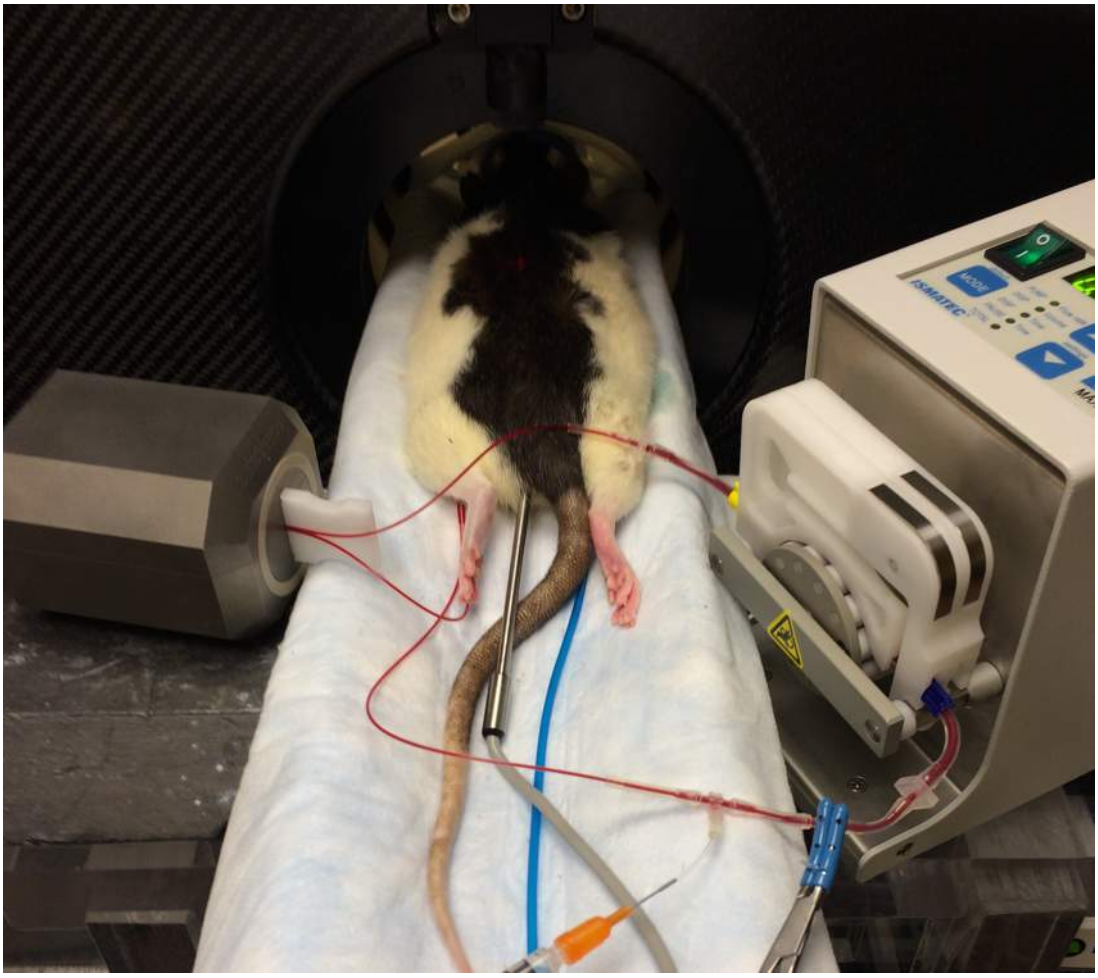


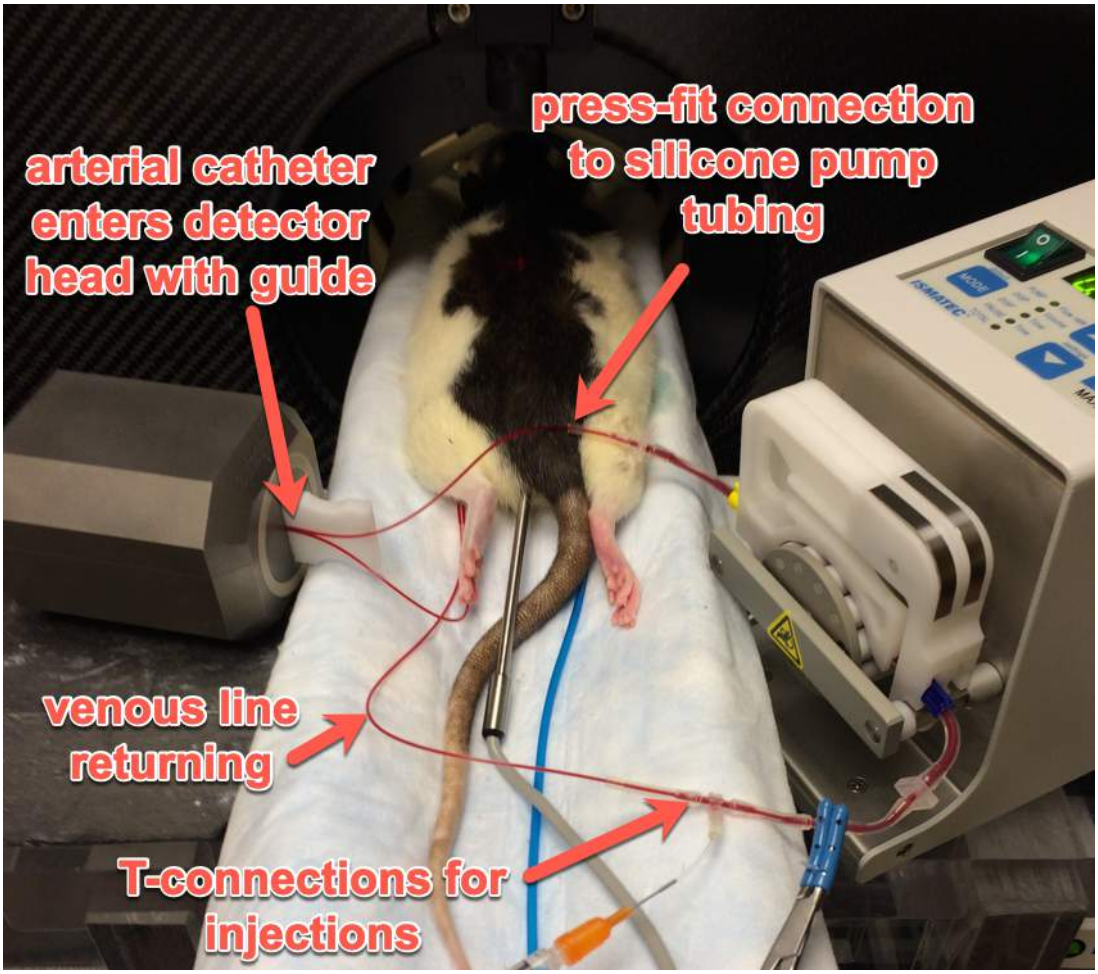
1.3.3 Continuous Blood Sampling in Small Animals

In small animals (rats, mice) blood loss is a crucial constraint. In terminal experiments this can be avoided by using an arteriovenous shunt between femoral artery and vein. Hereby, catheters are surgically inserted into the artery and vein and ensure reliable blood flow and return.

The arterial catheter is passed through the coincidence detector head for the activity measurement, then connected to silicone tubing in a peristaltic pump. The pump controls the flow rate and connects the arterial catheter to the venous one, creating a shunt. T-connections in the venous line allow the tracer to be injected directly, without using the tail vein, and for saline or investigational compounds to be injected.

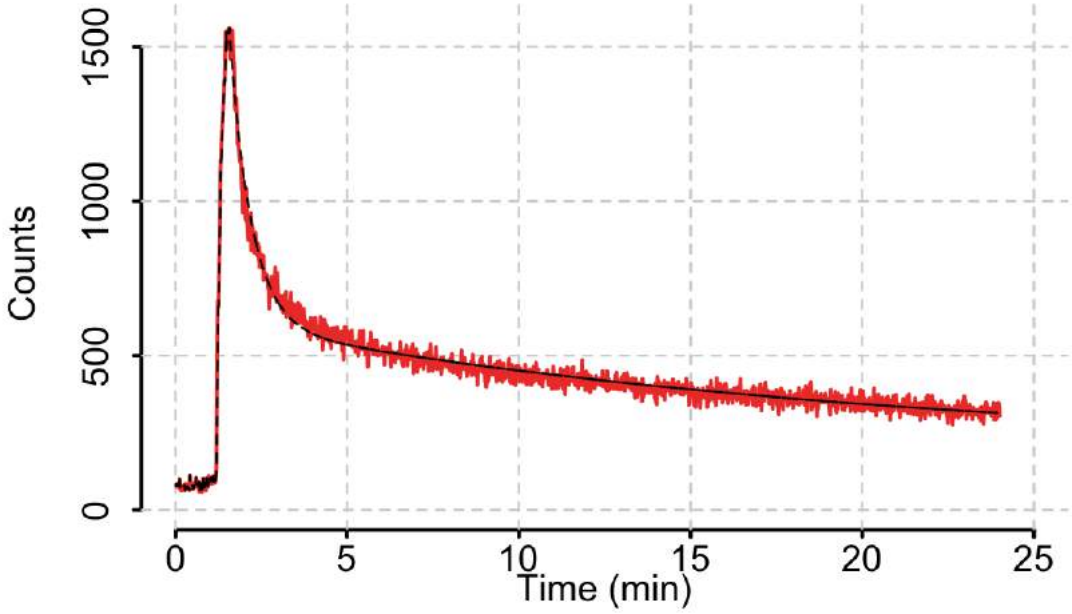
Manual blood samples may still be taken from the arterial side as required. This can be achieved by pausing the pump, disconnecting the arterial catheter, and allowing arterial blood to flow freely into a collection tube. With practice, this allows the user to collect 400 uL blood in approximately 30 seconds. Alternatively, a small cut in the silicone tubing of the peristaltic pump can be used to collect small drops of blood without pausing the pump, as described by Weber et al., 2002. A T-connection may also be added, but this can increase the risk of blood clotting in the system. Only limited manual sampling during PET is recommended in order to not alter the animal's physiology. Methods to obtain full plasma/WB ratio and parent fraction data despite this limitation are discussed later.





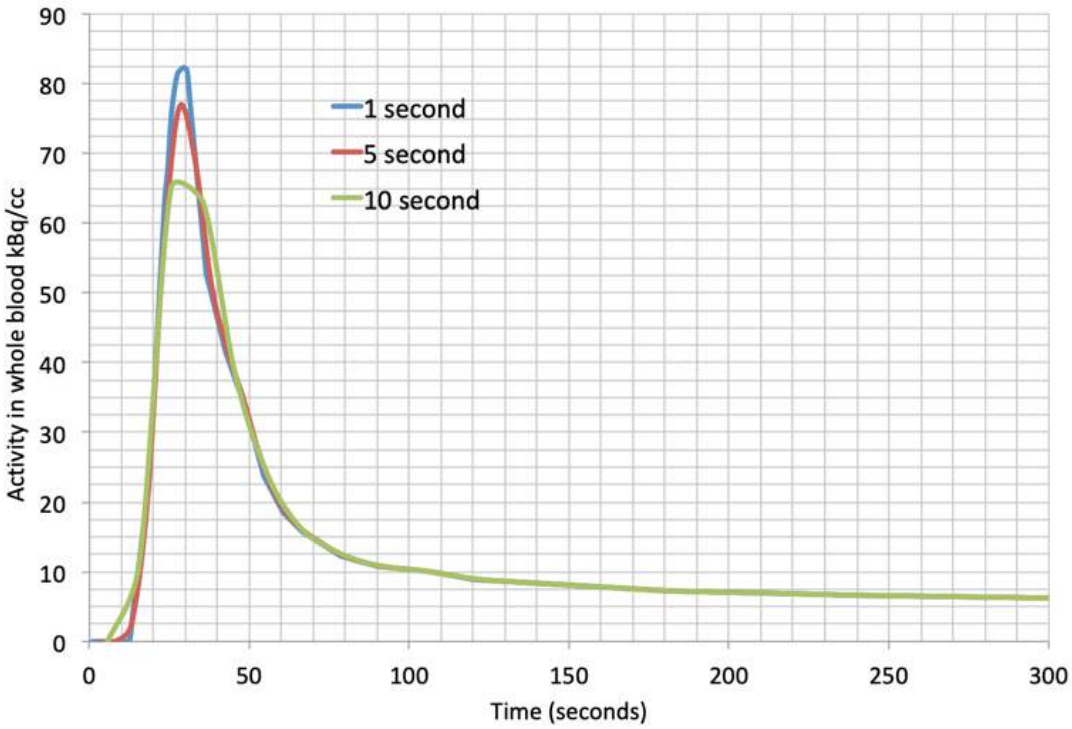
1.3.4 Example Whole-blood Activity Curve

An example of whole blood activity data collected with the swisstrace twilite after injection of 45 MBq FDG in a rat is shown below. In this raw data background counts and noise from the LYSO crystals are not suppressed, seen as non-zero counts before tracer injection and noise in the tail of the input function. The raw data is shown in red, and a bi-exponential model fit in black.



Sampling Density

The issue of limited temporal resolution sampling on a simulated blood curve is illustrated below. As the temporal resolution worsens, the peak concentration is successively underestimated.



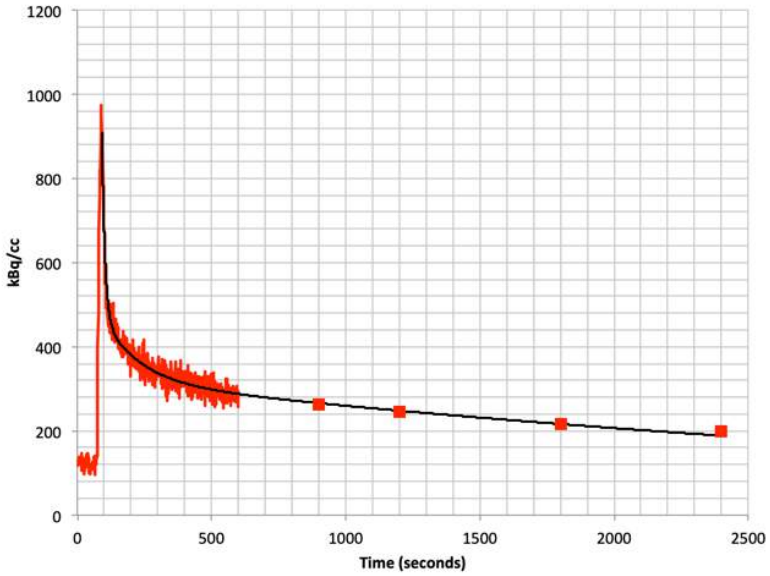
1.3.5 Pitfalls

Continuous measurement avoids some of the problems of manual blood sampling, but has other constraints. It is reliant on the blood flowing without clotting or catheter failure for the duration of measurement. Maintaining flow with minimal pauses is generally sufficient to maintain flow in human subjects, and heparin may be used in animal subjects. Poor positioning of the hand in

human subjects with radial artery catheter can lead to impeded flow and/or clotting. In order to maintain flow during the peak of activity concentration, concurrent manual sampling is not recommended. It can therefore be difficult to measure plasma/WB ratio during the first minutes after tracer injection. During this time the plasma/WB ratio may change rapidly for some tracers. A solution is to perform a separate experiment outside the PET scanner, focusing on manual blood sampling specifically for the plasma/WB ration and parent fraction. This approach is discussed later.

Where limited blood volumes may be withdrawn in human subjects and larger animals, a combination of continuous measurement and manual sampling can be used. In order to capture the peak as accurately as possible, the continuous measurement system is favorable. In human subjects, 5 minutes of acquisition at a blood withdrawal rate of 5 mL/min is generally sufficient to capture the peak. Thus 25 mL of blood is withdrawn. Swisstrace recommends a protocol that continues for the next 5 min at 2.5 mL/min, thus another 12.5 mL blood. Thereafter, a low flow rate of 0.33 mL/min can be used for a slow moving average of the input function tail, while keeping the catheter patent. Assuming a 90-minute total study, the remaining 80 minutes at 0.33 mL/hr would withdraw another 26.4 mL blood. In addition to this 65 mL blood (rounded up to allow for starting the pump slightly before tracer injection), a number of manual samples would still be taken for plasma/WB ratio and parent fraction. 8 samples of 5 mL would thus bring the total blood withdrawal to 105 mL.

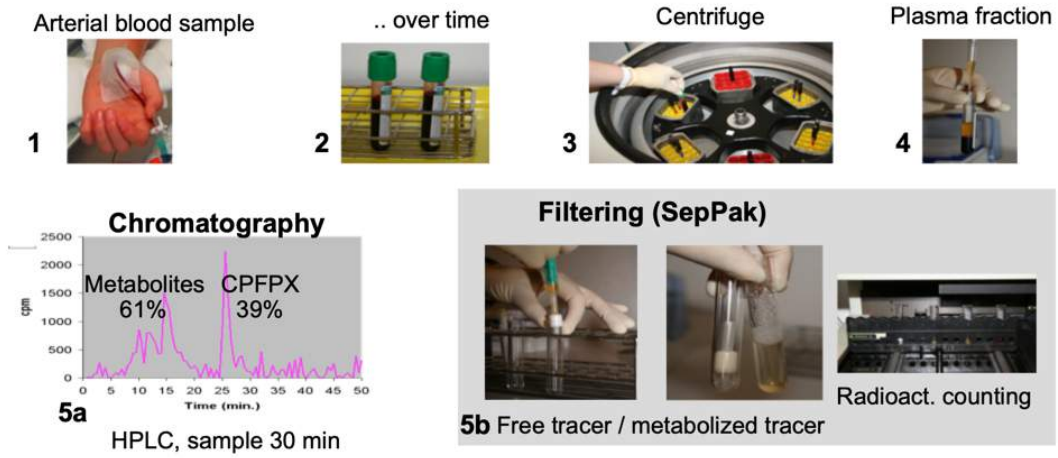
A limited protocol could stop continuous measurement after 10 minutes, having taken concurrent 5 mL manual samples at 5 and 10 minutes. Subsequently, 5 mL manual samples could be taken at 15, 20, 30 and 40 minutes. An aliquot of whole blood can be used to measure the activity in a gamma counter (calibrated to yield kBq/cc), and plasma/WB ratio and parent fraction measured in the same samples. This protocol would withdraw 67.5 mL blood. The data can be merged, and a model fitted to the data to recreate a continuous whole blood curve, as illustrated below. The initial period of continuous measurement could be further shortened to save additional blood volume.



1.4 Blood Sample Analysis

1.4.1 Overview

The blood analysis procedure is visually summarized below.



Blood Sample Handling

Manual blood samples should be rapidly transferred to heparinized or EDTA-treated sample tubes to prevent clotting, and stored on ice. Nevertheless, processing the samples as soon as possible is recommended. Care should be taken to gently mix the blood with heparin/EDTA. Shaking the sample may cause lysis of the red blood cells resulting in an inaccurate plasma/WB ratio.

Whole Blood Pipetting

An Eppendorf-style pipette should be used to extract an aliquot, or duplicate, of whole blood from the sample. Typically 20 or 50 uL is sufficient per aliquot, but correct pipetting technique must be observed to keep the volume reproducible. The entire pipetting tip is usually ejected into the tube for gamma counting to avoid losing activity sticking inside the tip.

Centrifuge to get Plasma

The remaining whole blood sample should then be centrifuged to separate the plasma. A typical centrifuge protocol is 5 minutes at 1100 g (or RCF, relative centrifugal force), 4 Celsius. Note that this is not RPM, revolutions per minute. Consult your centrifuge manual for the conversion between RPM and RCF for the rotor installed.

Visually inspect the sample after centrifugation. The plasma should be clear and red blood cells in a clear layer at the bottom of the tube. The sample may be centrifuged a second time if the separation is poor.

Plasma Pipetting

Pipette aliquots from the plasma in exactly the same way as for the whole blood. Continue to observe good pipetting technique and the precise volume used for the whole blood aliquots. This aliquot is used for the plasma/WB fraction.

Gamma Counting

The whole blood and plasma aliquots from the sample should then be transferred to the gamma counter and a counting protocol appropriate to the isotope started. Use the results for calculation of the Plasma/WB fraction as described below 14.

Metabolite Analysis

The remaining plasma is pipetted off without disturbing the red blood cell layer and used for metabolite analysis as described below 16. Indeed, the same samples may be used to derive both plasma/WB ratio and the parent fraction.

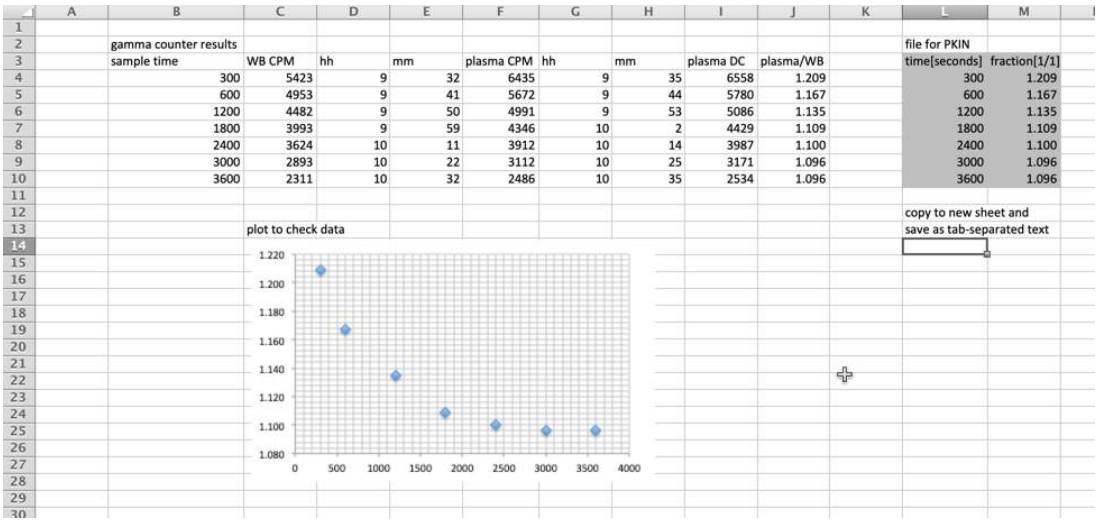
1.4.2 Plasma/Whole Blood Ratio

The plasma/WB ratio converts the activity concentration measured in whole blood to the expected concentration in plasma. Given whole blood activity units of kBq/cc, the resulting units for plasma activity will thus also be kBq/cc. Essentially, the activity in a fixed volume of plasma from a sample at a given time point should be divided by the activity measured in the same volume of whole blood.

1.4.2.1 Plasma/WB Ratio Calculation

When calculating the plasma/WB ratio it is not generally necessary to calibrate the results to kBq/cc. The ratio between CPM for plasma and whole blood is sufficient.

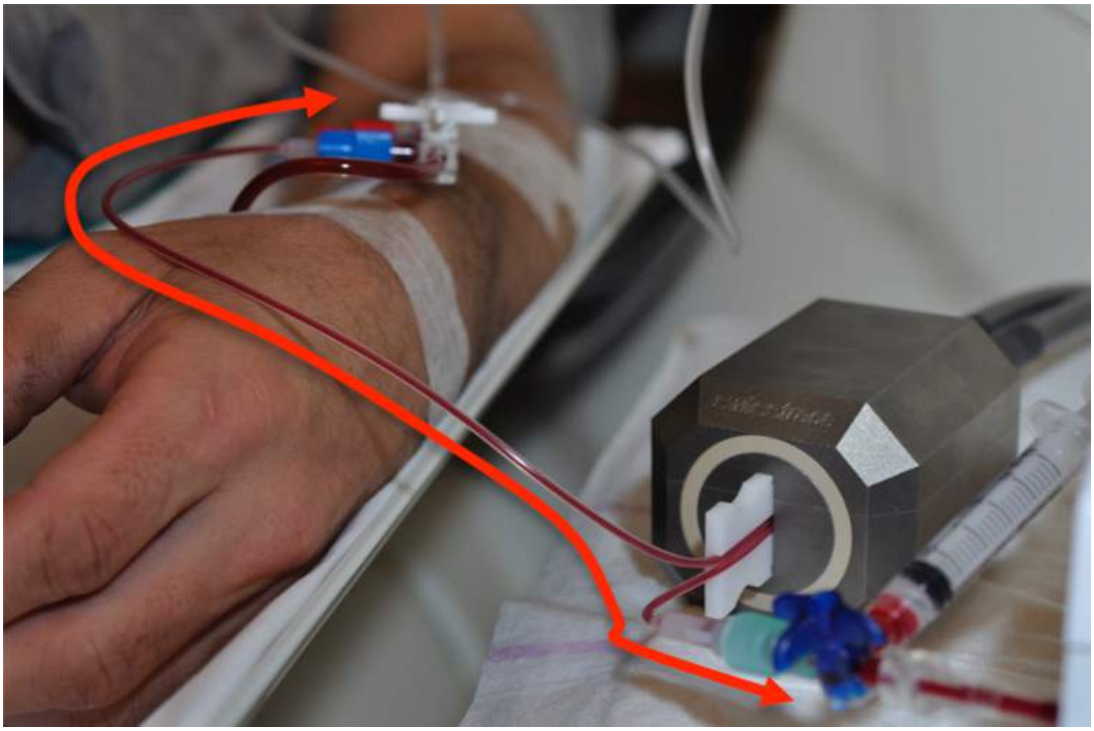
The resulting CPM for plasma and whole blood should be transferred to a spreadsheet, and the ratio plasma divided by whole blood calculated for each time point. A simple text file indicating the time of each sample, and this ratio, can be prepared for importing the Plasma/WB ratio 29 into PMOD's PKIN tool. The corresponding data structure is indicated below.



1.4.2.2 Pitfalls

Dead Volume

The volume of blood in the catheter/system between the subject and the point of sample collection is known as the "dead volume" (the image below indicates the dead volume). In the case of discrete sampling for the arterial input function, the catheter may be flushed with saline between samples. In such a case it is clear that directly collecting a sample would lead to dilution of the blood with this saline "dead volume". This would falsely increase the apparent volume of plasma and bias the plasma/WB ratio. During continuous measurement with flowing blood this issue is less important, as the dead volume is constantly refilled.



Timing

It is important that the timing of the manual samples relative to the time of tracer injection and/or start of PET acquisition is recorded accurately. This is particularly important in early samples where the plasma/WB ratio may be changing rapidly. In the case of tracers with a uniform/fixed plasma/WB ratio the effect of errors will be minimal (e.g. 18F-PSS-232, Warnock et al., 2018).

Flow Disturbance

Where samples for plasma/WB ratio are taken in addition to measurement of activity with continuous blood flow, it is important that the sampling procedure does not interfere with the flow of blood through the activity detector. This is particularly critical in the early stage of acquisition when the activity concentration is changing rapidly. For later samples the flow may be completely interrupted.

Pipetting

Since small aliquots of blood and plasma will likely be used to calculate the plasma/WB ratio, and the opportunity for replicate measurements is limited, accurate pipetting is essential. A 1 uL error on a 20 uL sample could introduce 5 % error into that point of your plasma/WB ratio. When using Eppendorf-style pipettes, the smallest volume possible pipette should be used. E.g. for a 10 uL sample, use a 20 uL max. volume pipette; for a 50 uL sample, use a 100 or 200 uL max. volume pipette, not a 1000 uL max. volume pipette. Carefully follow the manufacturer's recommendations for accurate pipetting. Generally these include: only setting the volume by winding the adjuster "down", keeping the pipette close to vertical during operation, during the sample in a slow, steady, motion. Eject the entire sample and tip into the tubes used for gamma counting to avoid leaving a small volume in a discarded tip.

Plasma must be accurately separated following centrifugation. If red blood cells are accidentally pipetted, the aliquot should be discarded. If the volume of plasma available is too small to discard any, pipette the contaminated aliquot back into the sample, and centrifuge again. As mentioned above, excessive shaking of the sample before centrifugation can lead to lysis of the red blood cells. In this case any tracer inside the red blood cells would be falsely measured as part of the plasma. If the plasma appears orange in color, consider discarding the entire sample.

Sampling Frequency

A sufficient number of samples must be taken to accurately describe the plasma/WB ratio over the time of PET acquisition. For the first subjects of a study this indicates many samples, e.g. at 2, 5, 10, 15, 20, 30, 40, 60, 70, 80, 90 minutes. If the data from these initial subjects reveals a slowly changing, or static ratio, the number of samples could be reduced to two or three for subsequent subjects.

Gamma Counters

Typically the activity in each sample is measured in a rack gamma counter. Such counters are capable of accurately measuring a low number of counts from small volumes, and they facilitate counting of duplicates to improve accuracy. The output of these counters is usually counts per minute (CPM).

Gamma counters are generally quite slow devices. The time delay between counting the whole blood and plasma aliquots from a given sample may be significant. E.g. gamma counters usually count for 60-120 seconds, and require several seconds to change to the next counting tube (for O-15- and C-11-labeled tracers this is relevant). Assess the delay between samples, and consider applying decay correction in your spreadsheet.

The accuracy and reliability of the gamma counter used to quantify activity in the whole blood and plasma aliquots should be verified. Your local radiochemist is usually a good source of expertise in using the gamma counter.

1.4.3 Metabolite Analysis of Plasma

An example procedure using a cartridge-based method for parent/metabolite separation will be described below. Consult your local radiochemist when designing your own protocol. HPLC may be required if cartridges are not suitable. 1-2 ml plasma is normally sufficient for metabolite analysis (a 5 ml blood sample will likely yield 2-2.5 ml plasma).

1.4.3.1 Cartridge Analysis

The cartridges used for separation of the metabolites (e.g. Sep-Pak C-18) should be pre-conditioned according to the manufacturer's recommendations. E.g. wash with 5 ml ethanol, then 10 ml distilled water.

An equal volume of cold (bottle on ice) acetonitrile is then added to the plasma to denature the proteins.

The plasma-acetonitrile mixture is then centrifuged using the same settings as for separation of plasma. This removes the denatured proteins, forming a pellet.

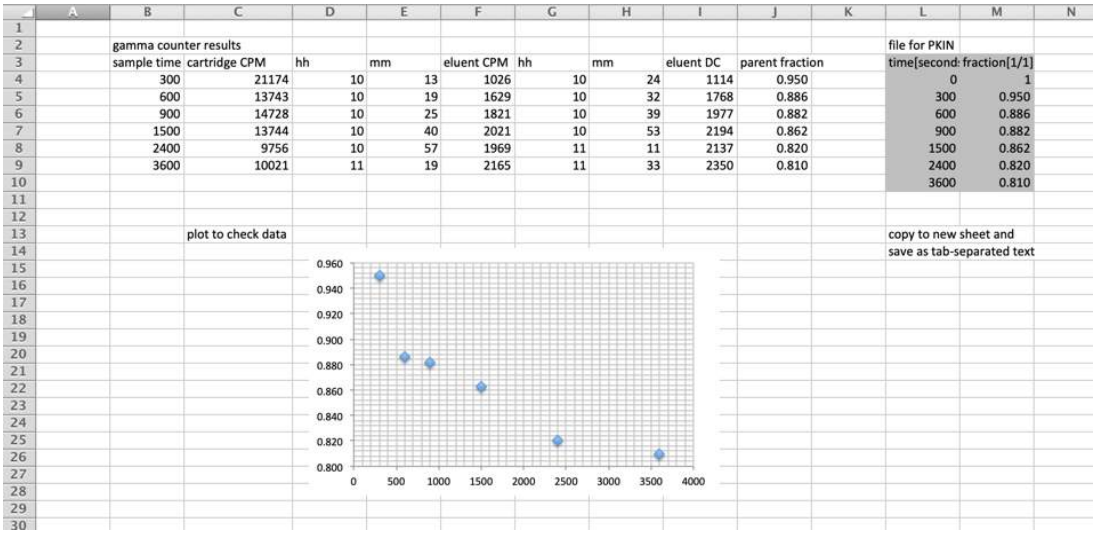
The supernatant is then pipetted off, and diluted in a 1:5 ratio with distilled water. The diluted supernatant is then passed through the pre-conditioned cartridge into a collection tube using a syringe. If the tracer is lipophilic, the intact compound should be trapped in the cartridge, and the polar metabolites remain in the water phase. Elute any remaining metabolites (not fully flushed through the cartridge in the first pass) from the cartridge into the same collection tube with 5 ml distilled water.

Count the eluent and cartridge using the gamma counter.

1.4.3.2 Parent Fraction Calculation

The ratio between the counts in the cartridge and the total (cartridge plus eluent) yields the parent fraction.

The calculation, and format for a text file to be loaded in the PMOD PKIN tool, is illustrated below. A simple text file indicating the time of each sample, and the parent fraction, can be prepared for importing the Parent Fraction 30 into PMOD's PKIN tool.



		gamma counter results									file for PKIN		
		sample time	cartridge CPM	hh	mm	eluent CPM	hh	mm	eluent DC	parent fraction	time[second	fraction[1/1]	
		300	21174	10	13	1026	10	24	1114	0.950	0	1	
		600	13743	10	19	1629	10	32	1768	0.886	300	0.950	
		900	14728	10	25	1821	10	39	1977	0.882	600	0.886	
		1500	13744	10	40	2021	10	53	2194	0.862	900	0.882	
		2400	9756	10	57	1969	11	11	2137	0.820	1500	0.862	
		3600	10021	11	19	2165	11	33	2350	0.810	2400	0.820	
											3600	0.810	
			plot to check data								copy to new sheet and		
											save as tab-separated text		

1.4.3.3 Pitfalls

The pitfalls in preparation of the parent fraction are similar to those discussed above for the plasma/WB ratio. However, a few additional issues should be considered.

- It is not always possible to use a cartridge method as described above. The suitability of cartridges from a given manufacturer for the tracer in question must be assessed. If cartridges are not suitable, an HPLC workflow is necessitated. However, sufficient knowledge of the HPLC is required. A column should be selected that optimally separates the parent tracer and the metabolites. If the HPLC system has a detector for radioactivity, is needs to be sensitive enough for the low activities present in late blood samples from studies with short half-life isotopes.
- An alternative to such a detector is fractionation of the column output into tubes for gamma counting. For example, the output may be collected in aliquots of 30 seconds, and these counted in a gamma counter, yielding a coarse description of the compound peaks.

1.5 Special Considerations

1.5.1 Limitations for Small Animals

The components of the blood data required for studies in small animals such as mice and rats are identical to those for studies in human subjects and larger animals. The major difference to be considered is the small volume of blood allowed for mice and rats.

Human

For comparison, for a 1-hour study in a human subject with a combination of continuous measurement and 10 manual samples for plasma/WB and parent fraction, approximately 130 mL blood would be used with the following protocol:

- Flow rate(s) for continuous measurement: 300 mL/hr, 6 min; 150 mL/hr, 5 min; 20 mL/hr, 49 min
- Manual samples: 2 mL dead volume, 5 mL sample
- Total volume withdrawn: ± 130 mL blood (based on a blood volume of 77 mL/kg in humans, a 75 kg subject would have 5.8 L blood)

Rat

As illustrated above 8, an arteriovenous shunt system can be used in mice and rats to externalize arterial blood for counting, but avoid blood loss by returning the blood directly to a major vein.

However, is there even sufficient blood for manual sampling and calculation of the plasma/WB ratio and metabolite analysis to get the parent fraction?

Assuming a blood volume of 64 mL/kg, a 300 g rat would have approximately 19 mL of blood. As a general rule, no more than 10 % of the blood volume should be withdrawn to avoid drastically altering the animal's physiology, and this only in an acute study. Hence the maximal blood volume available in our 300 g rat is approximately 1.9 mL. For the plasma/WB ratio alone, this may be sufficient for 8-10 samples, assuming careful blood handling and experience in pipetting small volumes (e.g. 200ul sample, 20 uL for WB measurement, 20 uL from approx. 100 uL plasma after centrifugation). However, metabolite analysis typically requires at least 300 uL plasma, necessitating 700-800 uL samples. Hence, a maximum of three data points is available for the parent fraction!

Mouse

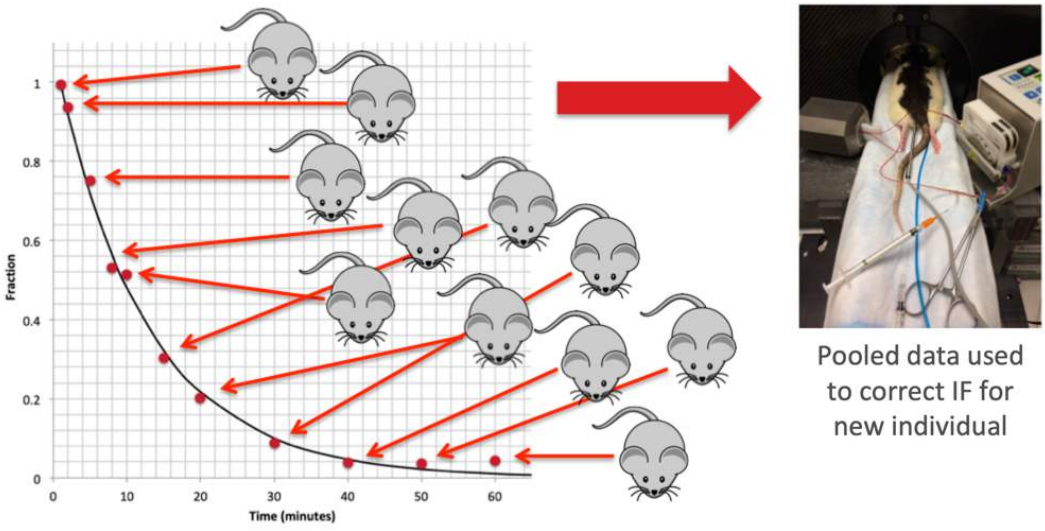
In mice, the situation is even more critical. Based on a blood volume of 79 mL/kg, a 30 g mouse would only have 2.4 mL of blood. The same 10% limit yields 240 uL blood for sampling, not even enough for a single metabolite analysis. How can an experimenter work around this problem?

Population Data

A common solution is to perform a separate study, outside the PET scanner, in which multiple animals are used to build up a "population" data set for plasma/WB and parent fraction. In rats, arterial catheters may still be used to get multiple samples per animal, while terminal blood sampling is typical in mice. In this way, a single rat may provide samples at 1, 5, and 15 minutes, another at 10, 20, and 40 minutes, and a third at 30, 50 and 60 minutes. Additional animals may be used to fill in the gaps further, or replicate measurements. The plasma/WB and parent fraction data from these animals is then pooled, providing the "population".

Assuming a homogeneous group of animals for the PET study (same strain, source, similar body weight, similar tracer injection and imaging protocol) the population plasma/WB ratio and parent fraction are used for kinetic modeling in all animals.

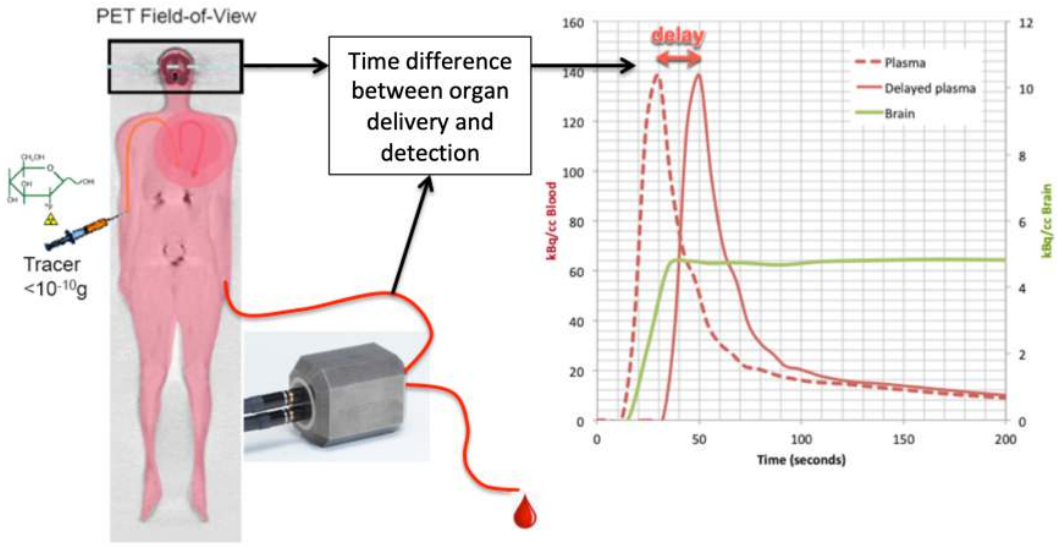
In disease models, or for example with transgenic mice, it may be beneficial to check that the plasma/WB ratio and parent fraction are consistent. The image below illustrates the process of pooling metabolite data as described.



1.5.2 Blood Timing Delay

In the context of PET kinetic modeling, "delay" is a time shift between tissue time-activity curves measured from the PET data and the blood data (i.e. whole blood curve). Such a delay occurs despite using synchronized timing for PET and blood sampling/detector systems.

When a tracer is injected intravenously, e.g. via the median cubital vein, the activity must first travel to the heart. Blood returning to the heart is pumped through the lungs before circulating to the rest of the body via the left ventricle. Taking a brain PET study as an example, the bolus of activity will then rapidly reach the target organ via the carotid arteries. If arterial blood is sampled from the radial artery, with a catheter (to a detector), the bolus must travel via the brachial artery to the radial artery, to the detector. Thus, the detector may see the bolus some seconds later than the target organ.



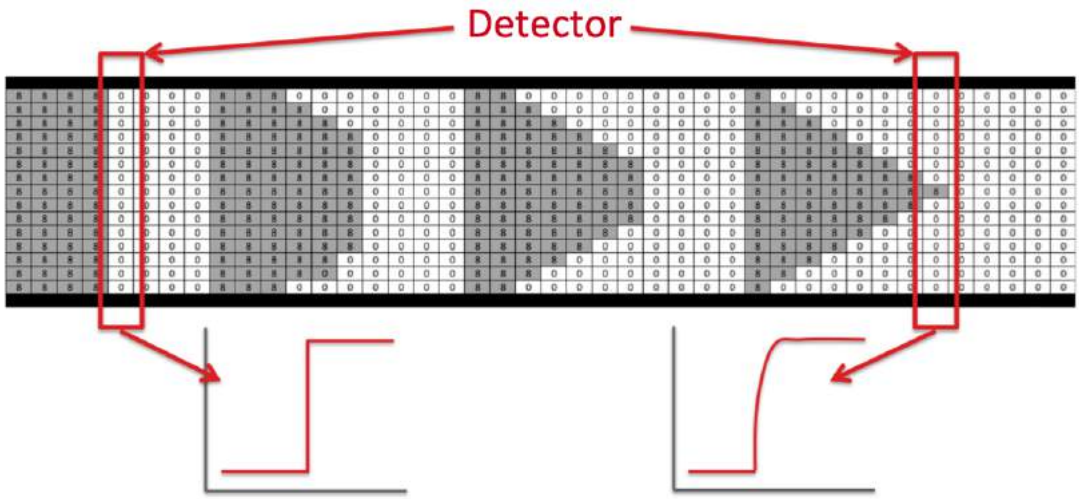
This delay is significant in kinetic modeling, as the initial up-slope of the input function should describe the initial tracer uptake in the target organ. Without delay correction, the rise in the target organ activity concentration before a rise in the blood would be unexplained, and a poor model fit would result.

With a standardized imaging protocol and subject setup, the delay may be reproducible and could thus be corrected during preparation of the blood data. Alternatively, the delay can be handled as part of kinetic model fitting [32], as is possible in PMOD's PKIN tool.

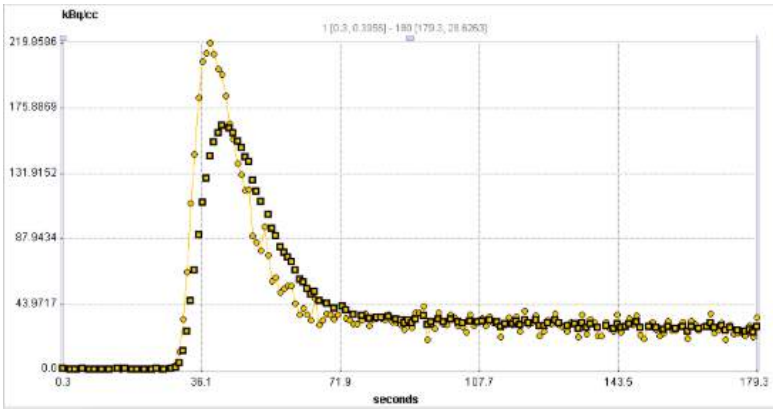
1.5.3 Bolus Dispersion

In addition to delay as described above, fluid (blood) flow in the blood-sampling catheter can introduce "dispersion" to the whole blood curve. Dispersion leads to a smoothing of fast concentration changes in the catheter, and is a consequence of laminar flow in the catheter. The flow velocity in the catheter ranges from zero at the walls, to a maximum at the cross-sectional center. The degree of dispersion is governed by the flow rate, length of the catheter and its internal diameter. With sufficient flow rate, minimum catheter length and diameter the effect of dispersion may be negligible.

The effect of dispersion on a step change in the blood activity concentration can be visualized as follows.

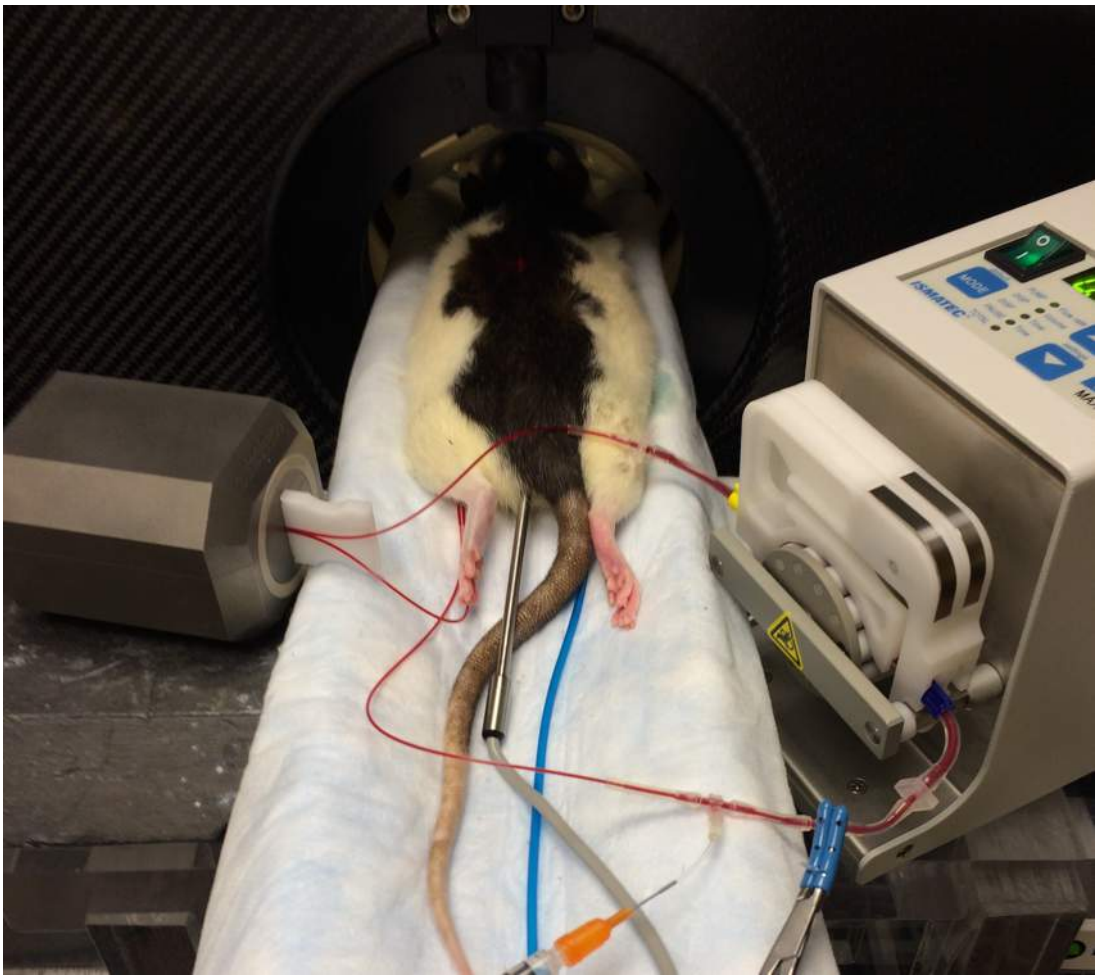


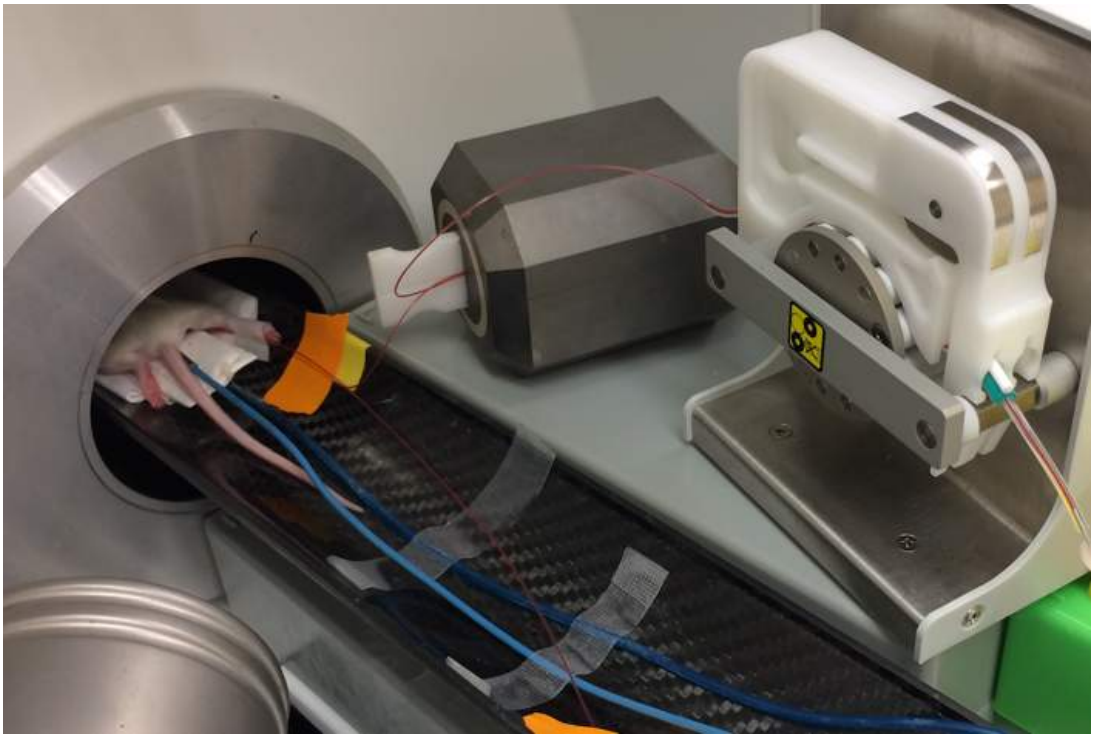
The effect of dispersion on the whole blood curve is illustrated in the figure below.



The degree of dispersion introduced by the sampling system can be quantified using bench-top experiments and a "dispersion kernel" described by a time constant calculated (see below). In experiments using a continuous flow measurement system, a dispersion time constant of 10 seconds lead to 33% error in measured cerebral blood flow, while a 1.3 second dispersion constant only lead to 0.3% error (Votaw & Shulman, 1998).

Dispersion can be minimized by keeping the blood sampling catheter as short as possible, or the detector as close to the subject as possible, using the highest possible flow rate (during the expected peak in activity concentration) and using a catheter with a small internal diameter (Convert et al., 2007). These factors are limited by the experimental setup, subject physiology and detector sensitivity. With the swisstrace twilite the detector is placed as close to the subject as possible, illustrated below, using a 1 mm internal diameter catheter with flow rate 5 mL/min for human subjects and large animals. In rats, a PE50 catheter with 0.58 mm internal diameter and 0.32 mL/min flow rate is used. In mice, a PE10 catheter with 0.28 mm internal diameter and 0.12 mL/min flow rate is used.

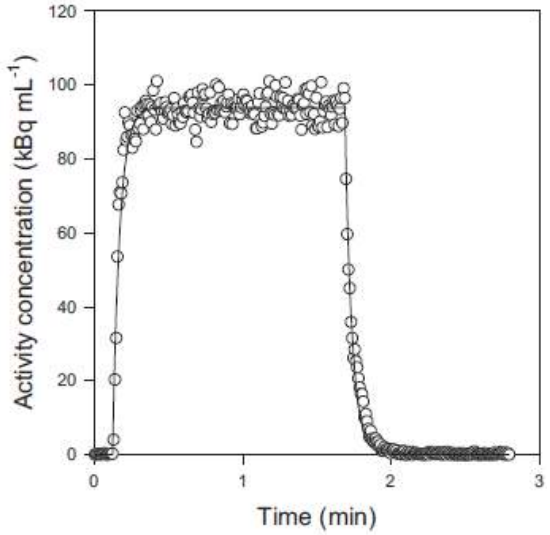




Dispersion Correction

Several dispersion correction methods have been developed. Most of them describe dispersion as a convolution of the true concentration curve with a dispersion function, and correction amounts to a numerical deconvolution that produces results suffering from excessive noise. Munk et al. (2008) developed an alternative approach that circumvents numerical deconvolution. It describes transport of blood through a catheter by a "transmission-dispersion" model which includes two components: molecules which travel undisturbed in the inner of the catheter (convective flow), and molecules close to the catheter wall to which sticking occurs (stagnant flow).

The model parameters have to be obtained from a calibration measurement with exactly the same conditions as the live experiment, i.e. with same catheter length, withdrawal speed and tracer as described in Munk et al. (2008). Basically, two beakers are prepared: one beaker with blood only, the other with blood and tracer (taken 1 min after tracer infusion). A three-way tap is used with one catheter leading to the blood sampling system, and the other two into the beakers. Blood sampling is started with connection of the blood-only line to the detector to measure the baseline with no radioactivity. Then the tap is switched to the catheter with blood and tracer. Correspondingly, the measured radioactivity rises to a constant level. Sometime later the tap is switched back to the blood-only line, and the radioactivity falls back to the baseline. Instead of the true rectangular shape of the concentration at the tap a smoothed shape is measured.



A workflow to calculate the system dispersion is available in the PKIN tool using such input data and the method described by Munk et al.

2 Alternatives to Measured Arterial Input Function

2.1 Population Input Function

In the same way that a population plasma/WB ratio and parent fraction can be used in subsequent groups of subjects (animals), as described above, researchers have attempted to use a common whole blood curve in all subjects after an initial group with blood sampling (Treyer et al., 2003; Meyer et al., 2006).

A pilot study is performed with PET imaging protocol identical to that planned for the main study. A standardized protocol for the tracer injection is highly recommended, for example using a syringe pump. In this pilot study, all of the blood data components are collected in several subjects and an average input function calculated. In order to achieve this, each individual input function must be scaled, for example to its own peak, to account for differences in the injected tracer dose.

With a standardized injection protocol and subjects with similar physiology/size/weight, the shape of the input function may be similar. However, as for the population plasma/WB and parent fraction approaches described for small animals above, this must be verified in disease models or treatment groups.

Even when using a standardized injection protocol to make the shape of the input function more reproducible, direct measurement of the whole blood curve provides important scaling information (while in contrast the plasma/WB ratio and parent fraction are relative measures). Several approaches to scaling the average input function for subsequent individuals have been proposed. In acute studies in small animals this may be an end point blood sample (by this stage arterial and venous blood activity concentrations may be very similar). In the method described by Treyer et al. for cerebral blood flow measurements in human subjects, the relationship between tissue perfusion and tracer washout rate was leveraged to scale the final cerebral blood flow maps.

Population input function in PET kinetic modeling has not become common practice.

2.2 Image-derived Input Function

The most widely known, and most desirable, alternative to arterial blood sampling is the derivation of an input function from the PET data itself. Many studies have described methods to derive a reasonable image-derived input function with validation against blood sampling.

An image-derived input function relies on the presence of a major blood pool or vessel in the imaging field-of-view. The most commonly used regions, in human subjects, large and small animals, are the carotid arteries and heart left ventricle. In some cases, the abdominal aorta has also been used.

Even when a vessel or the ventricle can be clearly identified, an image-derived input function provides only the whole blood activity curve. Depending on the tracer, a correction strategy for plasma/WB ratio and parent fraction must be considered based on a study comparing image-derived input function and full blood sampling.

It should also be noted that an input function necessitates high temporal resolution in the first minutes of the reconstructed PET data. List mode acquisition is recommended to allow multiple reconstructions of the same data to find the optimal initial time framing (e.g. 2, 5, 10 seconds) with the problem that short image frames may be very noisy.

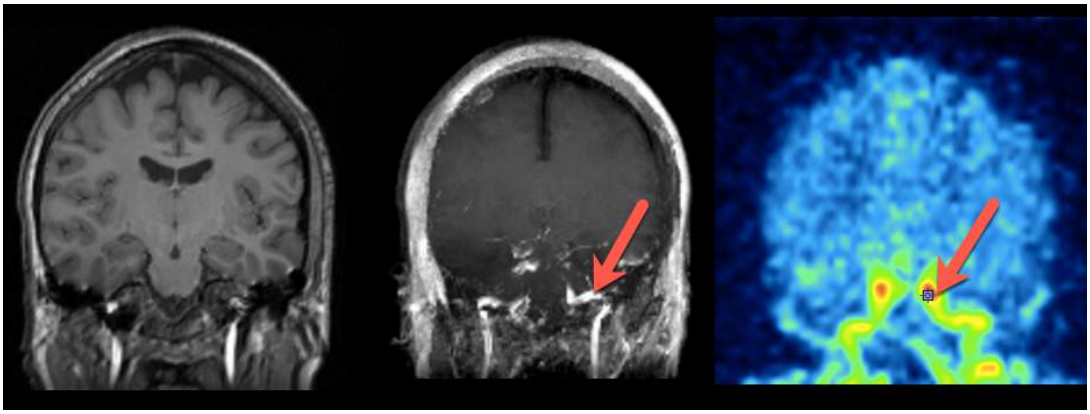
Some examples of image-derived input function in the literature and in practice are discussed below.

2.2.1 Carotid Arteries

In brain PET studies with sufficient field-of-view, the (internal) carotid arteries may yield a workable image-derived input function. For example, a method combining a "PET angiogram" from the early

image frames of O-15-water PET with an MR angiogram to delineate the internal carotid accurately has been described (Khalighi et al., 2018).

The quality of an input function derived from the carotid arteries should be assessed for individual PET scanners and image reconstruction algorithms. Without simultaneous PET/MR and suitable MR sequences in the acquisition protocol, accurately identifying the arteries may be difficult.



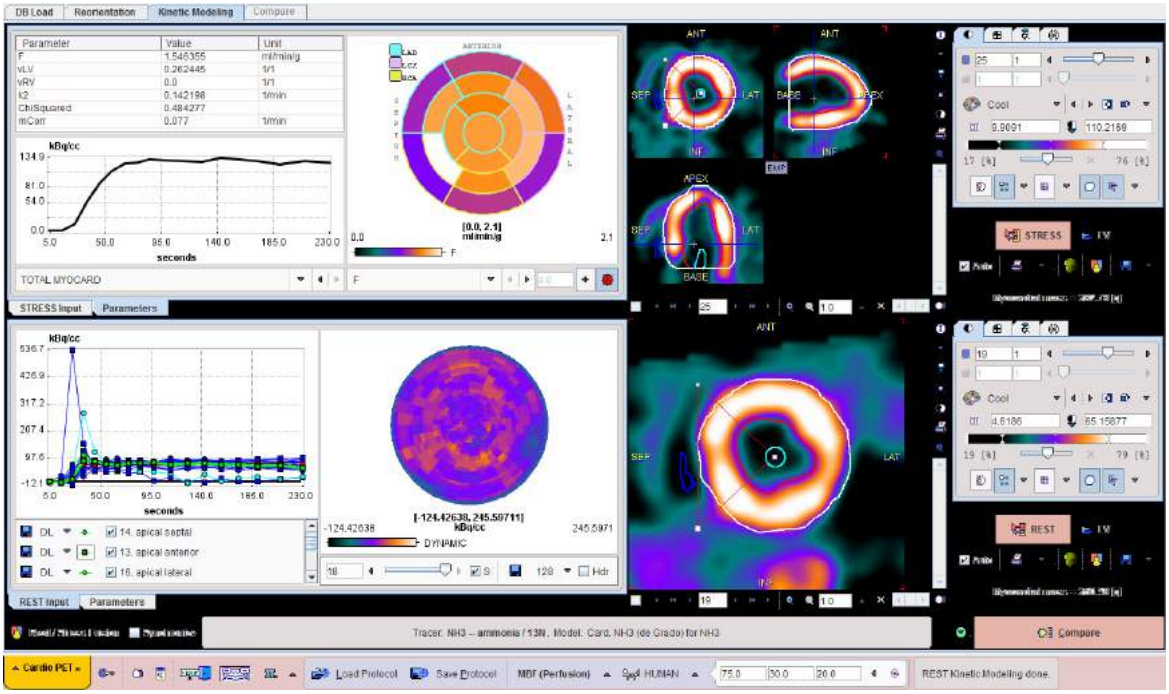
There are a number of limitations to this approach. Even in human subjects, the carotid artery is small compared to the spatial resolution of most PET scanners, and thus prone to poor recovery due to the partial volume effect.

Depending on the size of the field-of-view, only the internal carotid arteries may be included in a typical brain study, and then at the edge of the field-of-view. High activity in the blood may make it possible to identify the arteries in an early PET time frame, but in later frames it is often impossible to verify that subject motion has not invalidated the drawn VOI.

2.2.2 Heart Left Ventricle

Using the heart left ventricle is generally only possible during cardiac PET studies, or possibly in studies of lung/lung cancer. Advanced imaging protocols with rapid bed motion can circumvent this problem to some degree, and the advent of super-large field-of-view (human whole body) PET scanners could remove it entirely.

The most successful use of image-derived input function is the derivation of a blood curve for cardiac PET quantification in human subjects. In cardiac imaging the PET field-of-view is centered on the heart for the complete study, and the left ventricle provides a large blood pool in which to define a volume-of-interest. Additionally, the tracers and models used for cardiac kinetic modeling have been validated for use with only a whole blood curve, foregoing mathematical plasma/WB and metabolite correction. PMOD's Cardiac PET tool provides a convenient workflow for the analysis of myocardial perfusion in stress and rest studies, yielding the cardiac flow reserve.



The heart left ventricle has also been successfully used for FDG image-derived input function in mice on a PET scanner with sub-millimeter resolution (Alf et al., 2013).

2.2.3 Perspectives

Developments and improvements in PET technology will likely make image-derived input function more accessible in both human subjects and small animals.

Small animal PET systems with monolithic crystals and resolution recovery reconstruction algorithms yield sub-mm image resolution to rival depth-of-interaction dual-layer crystal designs, and have already been combined with high-field MRI systems. In synergy, such designs may make image-derived input function from the heart left ventricle more viable, especially for tracers other than FDG.

In human subjects, whole body voxel-wise kinetic modelling (parametric mapping) for FDG has been introduced by Siemens using rapid bed movement and regular measurement in the descending aorta. The recently developed Explorer PET from United Imaging with whole body field-of-view provides the blood pools of the heart left ventricle and the aorta for studies of all other organs.

However, a general solution to the issue of plasma/WB ratio and parent fraction determination for image-derived input function is not on the horizon and limits the approach to a few tracers including O-15 water and FDG.

3 Support for Blood Data in PMOD

As the AIF is an essential part of full PET quantification, it is used in several PMOD modules, namely PKIN (general kinetic modeling), PXMOD (parametric mapping), PNEURO (parametric mapping) and PCARD (cardiac quantification).

Two modules are involved in the actual creation of an AIF:

- PSAMPLE: This module is used to acquire the whole-blood activity curve via the twilite continuous blood sampling system of swisstrace.
- PKIN: This module allows loading the information of whole blood concentration, plasma/WB ratio and parent fraction, fitting smooth functions to the measurements, and combine them for calculating the AIF.

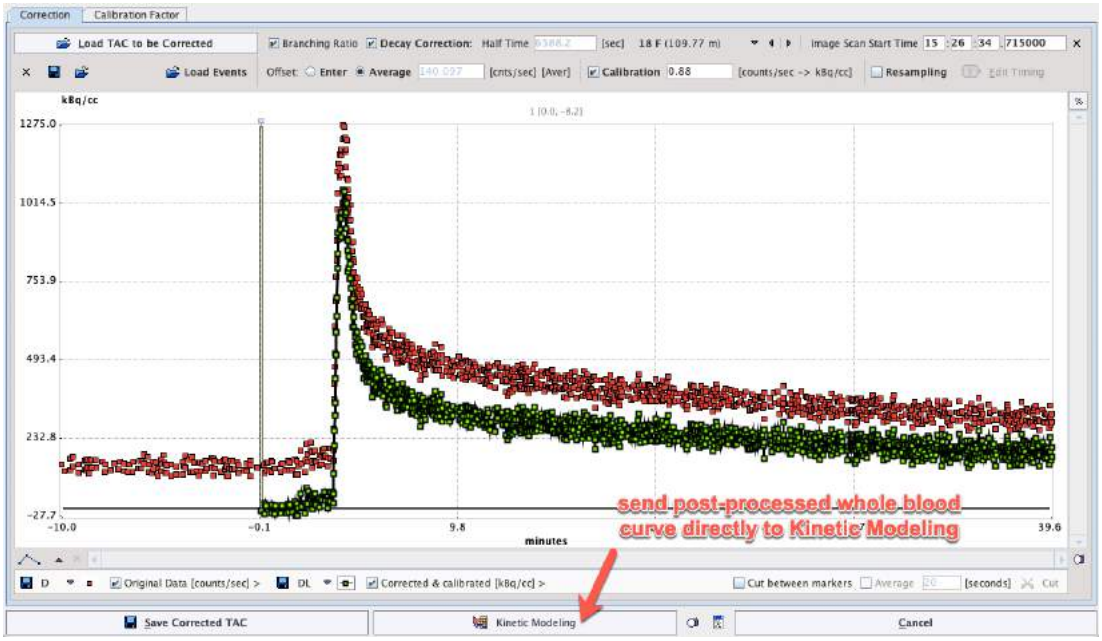
The final AIF can then be used for modeling in PKIN, but also exported for use in the parametric mapping modules (PXMOD, PNEURO).

The sections below provide a brief overview of PSAMPLE and the blood-related functionality of PKIN. Please refer to the PMOD Online Documentation for details.

3.1 PSAMPLE Tool

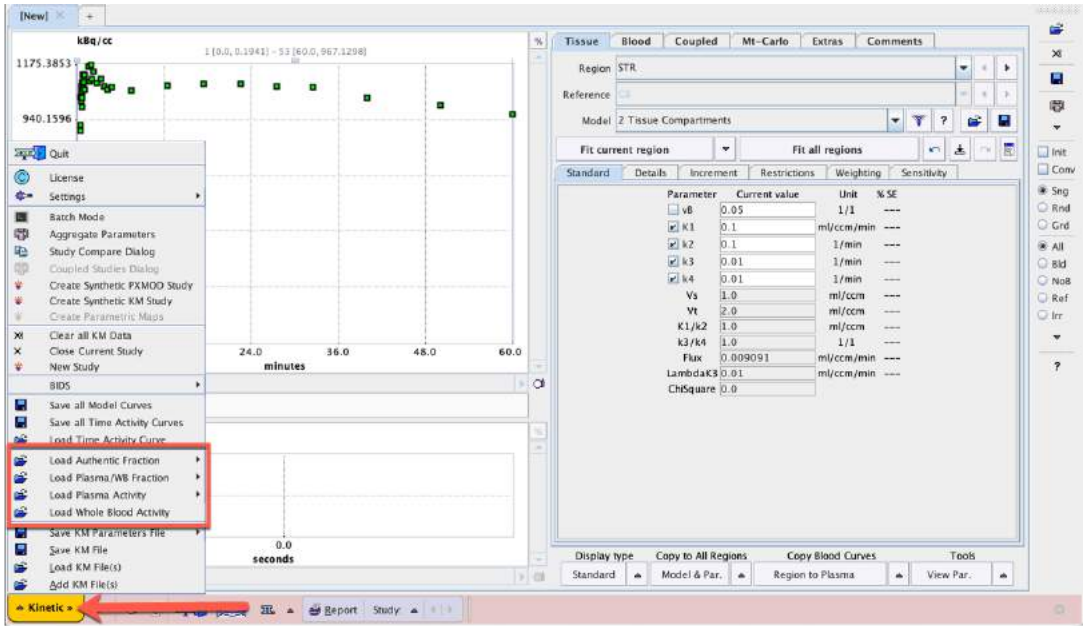
The swisstrace twilite blood sampling device is operated by the PMOD PSAMPLE tool. After data acquisition, correction and calibration procedures are applied to subtract the background, correct the radioactive decay, and calibrate the count rate data [counts/second] to kBq/cc.

The resulting whole blood activity concentration curve can directly be transferred to PKIN for further processing.



3.2 PKIN Tool

The blood components needed for calculating the AIF as discussed before are loaded using the related entries in the main tool menu.

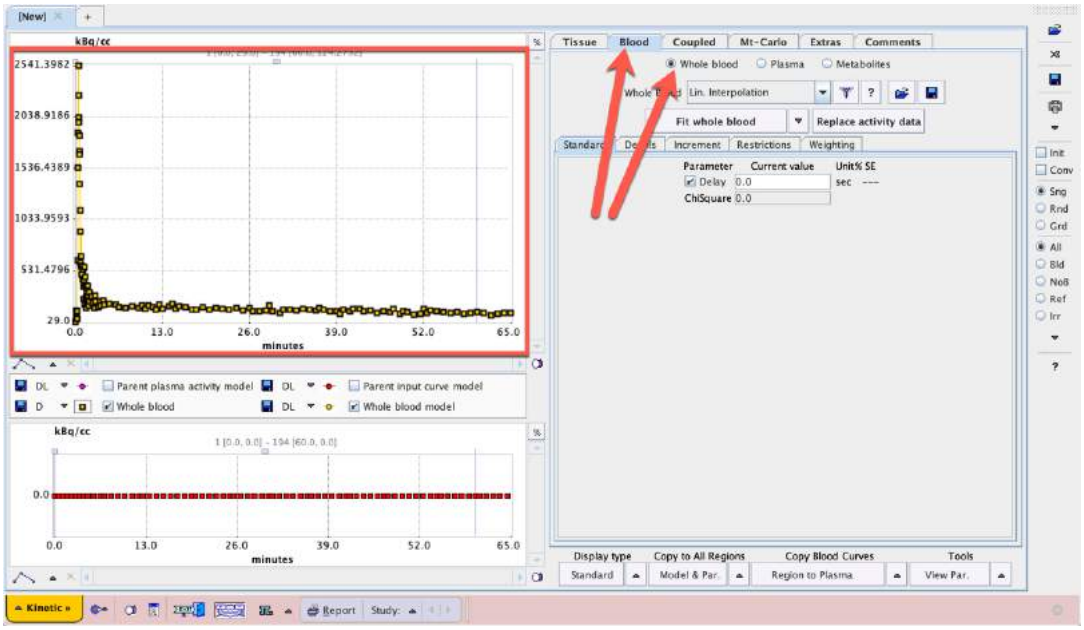


For the modeling, all components for the AIF need to be available at any arbitrary time. Per default, after loading of a measurement, the blood information in between the samples is calculated by interpolation. However, it is recommended and supported by PKIN to fit smooth and physiologically reasonable functions to the measurements. This is particularly relevant for the sparse results from blood analytics (plasma/WV, parent fraction). Adequate functions reduce measurement noise and support better extrapolation outside the measurement range.

The following sections briefly illustrate loading of the different blood components and fitting smooth functions to the measured curves.

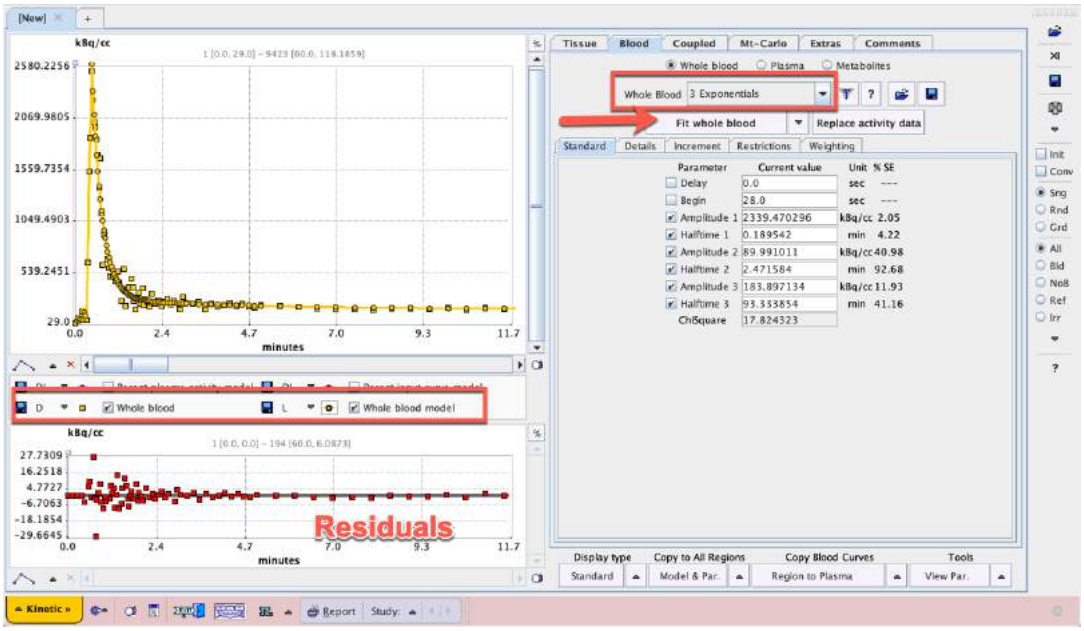
3.2.1 Whole Blood

The whole blood measurement should be loaded first with **Load Whole Blood Activity**. The data can be viewed via the **Blood** tab, **Whole blood** radio button.



On the same panel the interpolation function can be switched from **Lin. Interpolation** to a physiological alternative and the function fitted to the whole blood measurements. In the example below a **3 Exponentials** function was chosen. PKIN detected the peak, then only fitted the function

to the data from that point onwards. The fit can be assessed visually and using the residuals displayed in the lower plot area.

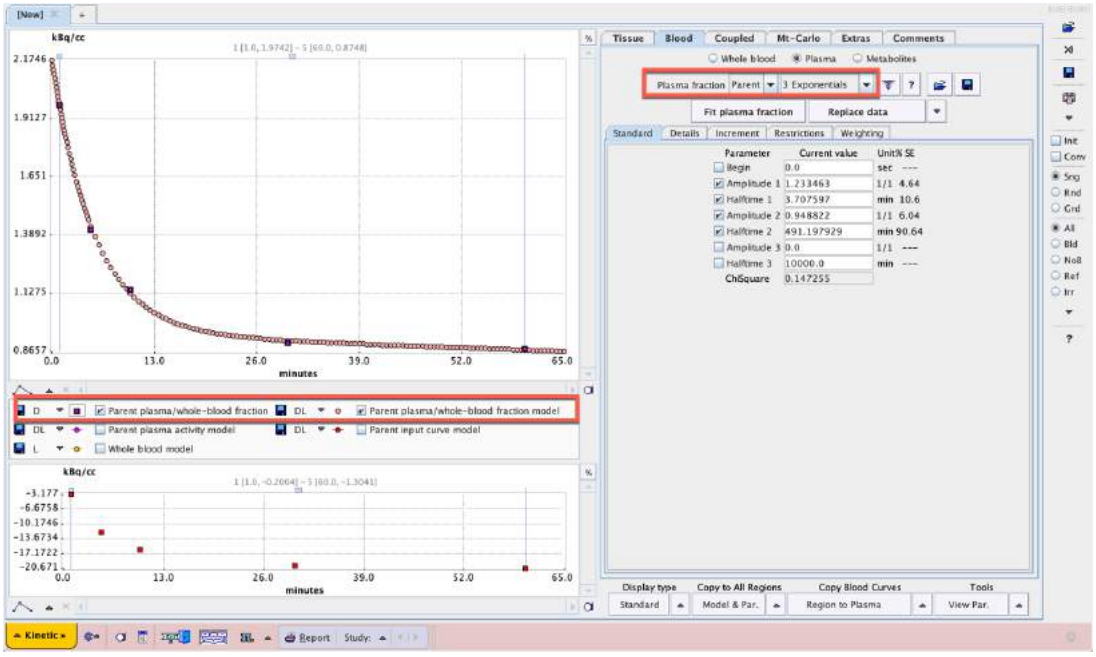


Note that the whole blood measurements will not used any more for calculations, but only the interpolation function. Further, as long as no other blood component is loaded, the whole blood interpolation function will directly be used as the AIF.

3.2.2 Plasma/WB Ratio

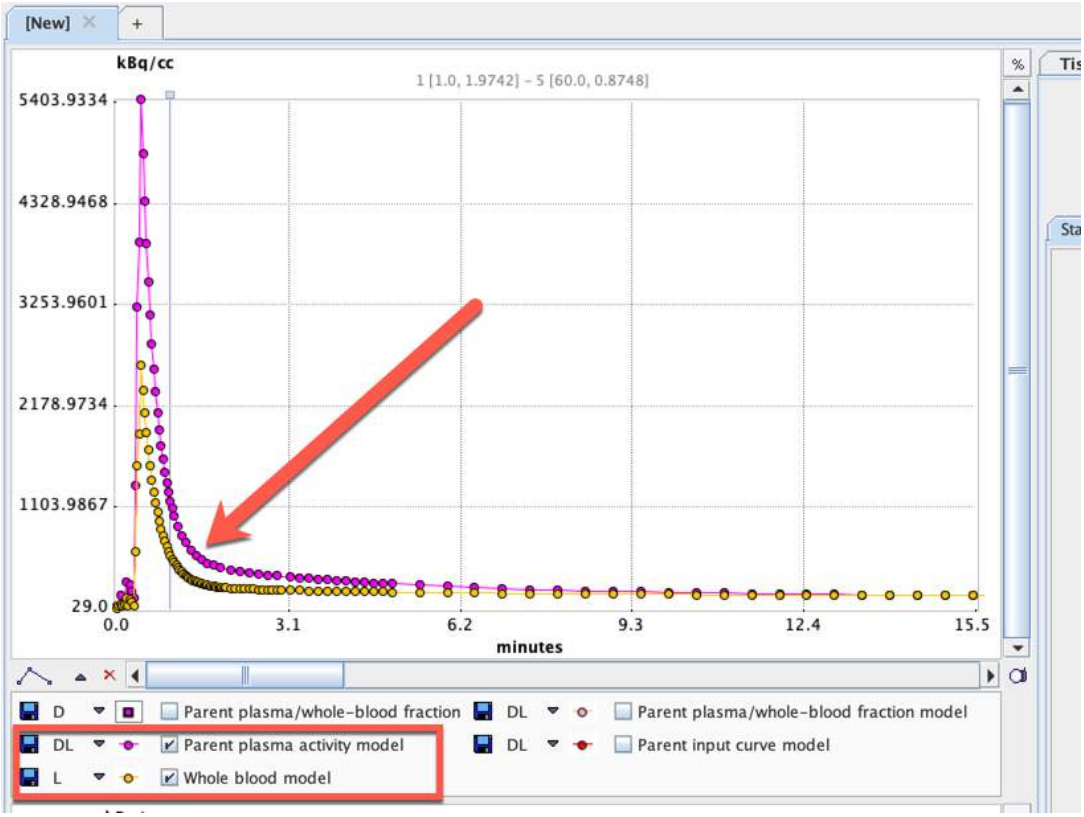
Next the plasma/WB ratio data is loaded via the menu **Load Plasma/WB ratio**. Remember that it is used to calculate the total activity in plasma by multiplication with the whole blood activity.

The example below shows a loaded plasma/WB ratio and a fitted **3 Exponentials** function similar to Whole Blood 28. Some of the curves should be toggled off to see only the measurement and the fit.



As soon as a Plasma/WB ratio has been loaded, PKIN multiplies its interpolation function with the whole blood function to calculate total plasma activity (called **Parent plasma activity model** below). The resulting curve can then be viewed by toggling on the appropriate checkboxes and

compared with the whole blood curve. As illustrated below, total plasma has a higher peak activity concentration than whole blood and approaches whole blood in the tail portion

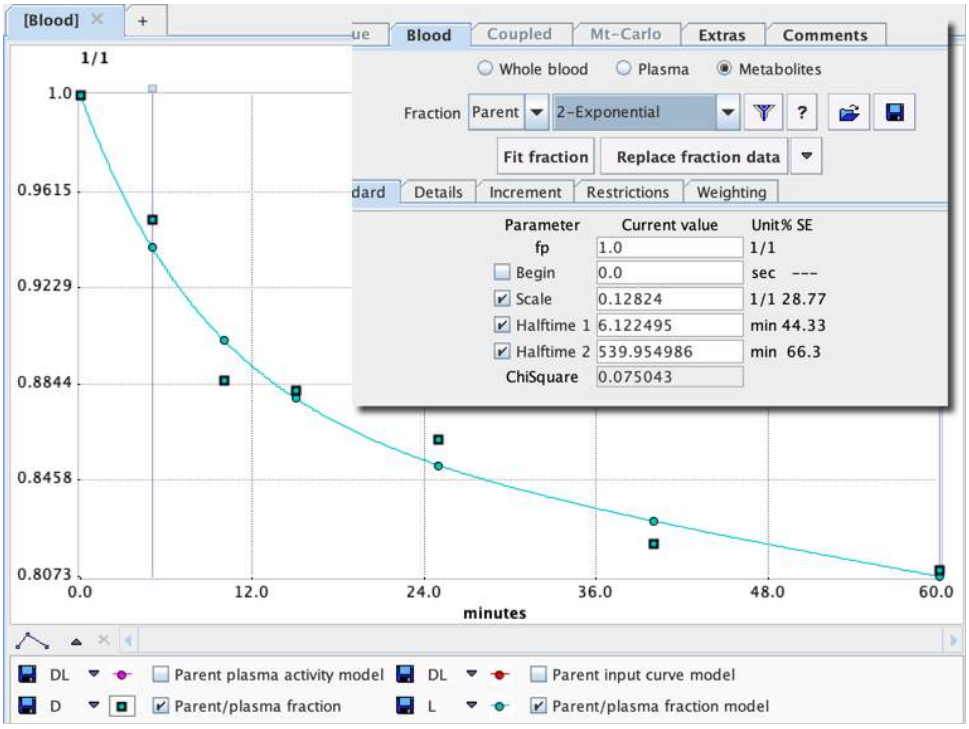


As an alternative, researchers may choose to prepare the total plasma activity and import it directly into PKIN via the **Load Plasma Activity** menu entry. In that case, no multiplication with whole blood is needed, and an interpolation function can directly be fitted to the total plasma measurements.

3.2.3 Parent Fraction

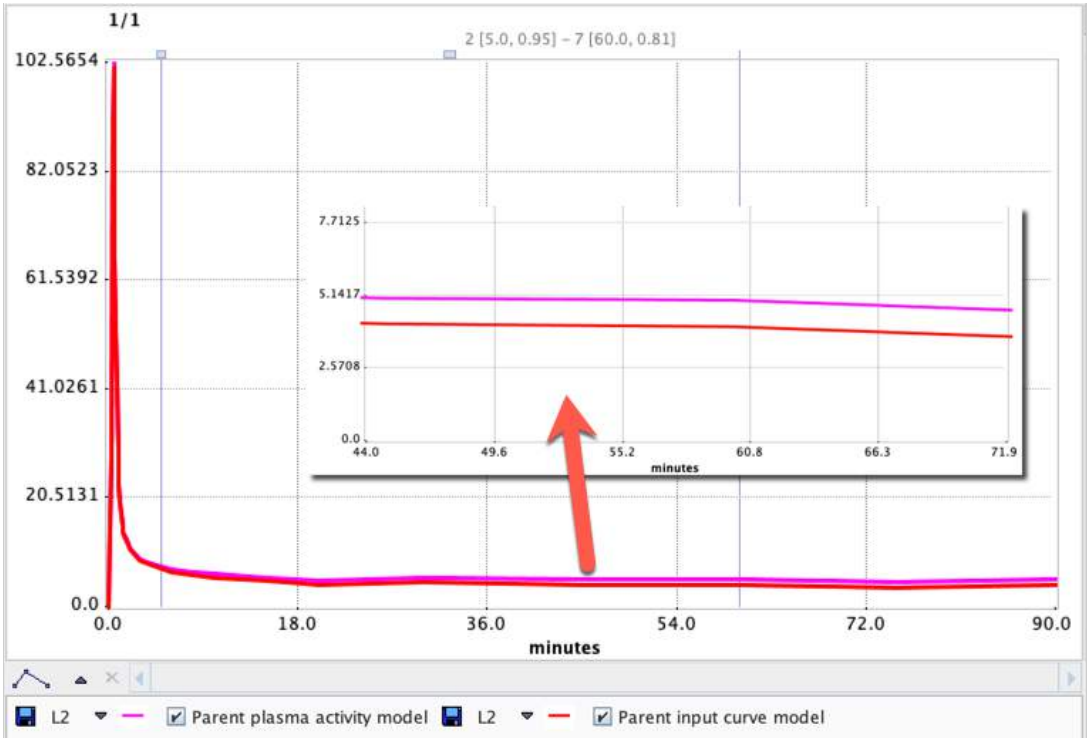
Next the parent fraction data is loaded via the menu **Load Authentic Fraction**. Remember that it is used to calculate the activity of unchanged tracer (parent, authentic) in plasma by multiplication with the total plasma activity.

The example below shows a loaded parent fraction and a fitted **2 Exponential** function similar to Whole Blood 28. It can be viewed via the **Metabolites** radio button on the **Blood** tab. Some of the curves should be toggled off to see only the measurement and the fit.



From then on, PKIN uses the product of this exponential parent fraction with the total plasma activity curve as the AIF.

As illustrated below, the resulting authentic tracer in plasma curve becomes lower over time to reflect the reduced contribution from authentic, non-metabolized, tracer.



3.2.4 Alternative Blood Representations

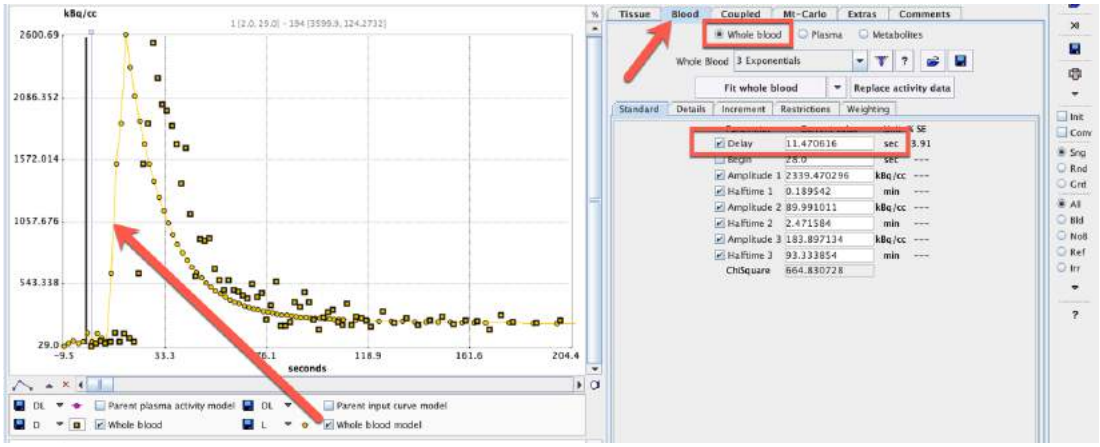
The three blood components discussed above represent only one of the possibilities how the data can be organized and handled in PKIN. Some alternatives are briefly described below.

- **Total Plasma Activity**: Instead of using the Plasma/WB ratio and multiplying it with whole blood to obtain total plasma activity, these measurements can directly be loaded from the menu via **Load Plasma Activity**. Metabolite correction will then directly be applied to the plasma interpolation function.
- **Metabolite-corrected Plasma Activity**: Metabolite correction can also be performed outside of PKIN. In that case the metabolite-corrected plasma activity can directly be loaded from the menu via **Load Plasma Activity**. Multiplication of the parent fraction with this plasma data will automatically be done, but because the default parent fraction is identical to 1 the loaded curve will act as the AIF.
- **Combination of Plasma/WB ratio and Parent Fraction**: It would be possible to directly calculate the ratio of authentic tracer to whole blood concentration. In this case, only a single function needs to be multiplied with whole blood for generating the AIF. In practice, the ratio would be loaded with **Load Plasma/WB ratio**, and the parent fraction left at identity to 1.
- **Interpolation function for AIF**: The AIF value at any time is calculated by multiplication of three functions at that time. In principle, it should again represent a smooth, physiologic shape which may be represented by a single function. This can be achieved by exporting the calculated AIF from PKIN, loading it again with **Load Plasma Activity**, and fitting an interpolation function.

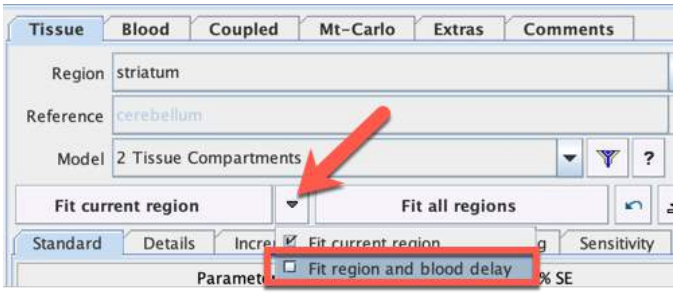
3.2.5 Blood Delay Handling

The PKIN tool handles the time difference between measurement of the blood and the tissue data as a **Delay** parameter in the blood interpolation functions. A positive delay is typical and means that the blood was measured after the same blood had been in the scanner field of view.

Illustrated below is the effect of the delay on the blood curve. As can be seen, the positive **Delay** of 11.47s shifts the blood interpolation curve left to earlier times, removing the delay between PET and blood measurements.



A delay value may be entered manually. But more conveniently, it can be estimated by including it in the model fitting process using **Fit region and blood delay**.



Note: A negative value may be entered in the **Delay** field as well, shifting the whole blood curve to the right. Although this does not correspond to the rationale for a delay between PET and blood data, a negative delay could represent error in the recorded timing. This is particularly possible for

manually sampled whole blood data, where samples take several seconds to be withdrawn. Additionally, the limited temporal resolution of manually sampled whole blood data may make the rising edge of the input function peak appear early or late. A negative delay value smaller than the time between blood samples may improve model fitting in this situation.

4 References

1. Weber B, Burger C, Biro P, Buck A. A femoral arteriovenous shunt facilitates arterial whole blood sampling in animals. Eur J Nucl Med Mol Imaging 2002; 29:319-323.
2. Warnock G, Sommerauer M, Mu L, Pla Gonzalez G, Geistlich S, Treyer V, Schibli R, Buck A, Krämer SD, Ametamey SM. A first-in-man PET study of [18F]PSS232, a fluorinated ABP688 derivative for imaging metabotropic glutamate receptor subtype 5. Eur J Nucl Med Mol Imaging. 2018; 45(6):1041-1051.
3. Votaw JR, Shulman SD. Performance evaluation of the Pico-Count flow-through detector for use in cerebral blood flow PET studies. J Nucl Med. 1998; 39:509–515.
4. Convert L, Morin-Brassard G, Cadorette J, Archambault M, Bentourkia M, Lecomte R. A new tool for molecular imaging: the microvolumetric beta blood counter. J Nucl Med. 2007; 48(7):1197-206.
5. Munk OL, Keiding S, Bass L: A method to estimate dispersion in sampling catheters and to calculate dispersion-free blood time-activity curves. Med Phys 2008, 35(8):3471-3481.
6. Treyer V, Jobin M, Burger C, Teneggi V, Buck A. Quantitative cerebral H2(15)O perfusion PET without arterial blood sampling, a method based on washout rate. Eur J Nucl Med Mol Imaging 2003; 30(4):572-80.
7. Meyer PT, Circiumaru V, Cardi CA, Thomas DH, Bal H, Acton PD. Simplified quantification of small animal [18F]FDG PET studies using a standard arterial input function. Eur J Nucl Med Mol Imaging 2006; 33(8): 948-954.
8. Khalighi MM, Deller TW, Fan AP, Gulaka PK, Shen B, Singh P, Park JH, Chin FT, Zaharchuk G. Image-derived input function estimation on a TOF-enabled PET/MR for cerebral blood flow mapping. J Cereb Blood Flow Metab 2018; 38(1):126-135.
9. Alf MF, Wyss MT, Buck A, Weber B, Schibli R, Kramer SD. Quantification of brain glucose metabolism by 18F-FDG PET with real-time arterial and image-derived input function in mice. J Nucl Med 2013; 54(1):132-8.

