CHAPTER 7

Estimation of Local Receptor Density, $B'_\text{max}$, and Other Parameters via Multiple-Injection Positron Emission Tomography Experiments

Evan D. Morris,* Bradley T. Christian,† Karmen K. Yoder,‡ and Raymond F. Muzic, Jr.§

*Departments of Diagnostic Radiology and Biomedical Engineering
Yale PET Center
Yale School of Medicine
New Haven, Connecticut 06510

†PET Physics
Waisman Laboratory for Brain imaging and Behavior
Departments of Medical Physics and Psychiatry
University of Wisconsin-Madison, Madison

‡Center for Neuroimaging
Stark Neurosciences Research Institute, and
Department of Radiology
Indiana University School of Medicine, Indianapolis, Indiana

§Departments of Radiology
Biomedical Engineering, and Oncology
Case Western Reserve University, Cleveland, Ohio

I. Update
II. Introduction
III. Theory
   A. Need for Models
   B. Compartmental Models
   C. Standard Model Equations
   D. M-I Model Equations
In the last few years, interest in “multiple-injection” (M-I) paradigms, related PET experiments, and the models required to describe them has proliferated. This development has been driven by three factors: (1) desire of PET users to study a broader range of physiological processes with imaging; (2) ongoing improvement in PET scanner sensitivity; and (3) greater availability of new tracers requiring kinetic characterization. With greater use of M-I experiments has come increased scrutiny of the details of model implementation (Salinas et al., 2007). Interestingly, a number of mathematically distinct, but practically equivalent, implementations have been developed by different researchers (model implementation is discussed later in the chapter).

Recently, Vandehey et al. made use of the M-I approach to do a complete characterization of the kinetic parameters of two high-affinity dopamine D₂/D₃ receptor tracers ([¹⁸F]fallypride and [¹¹C]FLB). Because M-I studies have the power to make precise estimates of all the individual rate constants of the model, the authors were able to identify subtle kinetic differences (specifically, the interplay between binding rate constants and blood flow parameters (Morris and Yoder, 2007)) that might make one tracer better suited to imaging competition with endogenous dopamine. Gallezot et al. (2008) used two different M-I protocols in baboons to estimate all of the model parameters for a popular nicotinic ligand, 2-[¹⁸F]fluoro-A-85380. Because the nicotinic agonist tracer was given in nontrace amounts, it was necessary to include multiple values of the blood flow parameter, $K₁$, during the study. Thanks to the precise parameter estimates afforded by the
M-I approach, the authors were then able to determine the level of bias to be expected from more common methods of estimating distribution volume. Bottlaender et al. extended the models of Gallezot to include additional compartments for nicotine so that the effect of exogenous nicotine on tracer binding could be studied. With the assumption that nicotine was in fast equilibrium between plasma and free compartments (thereby reducing the number of model parameters), the authors used the M-I curves of A-85380 uptake (in the presence of nicotine) to derive a relationship between plasma nicotine levels and nicotinic-acetylcholine receptor (nAChR) occupancy. This method yielded less bias than if [11C]-nicotine had been used to estimate nAChR occupancy.

The authors note that one of their early papers on [11C]raclopride displacement by endogenous dopamine (Morris et al., 1995) presaged the combined use of kinetic compartments for both multiple injections and multiple ligand species, such as was used by Bottlaender et al. The work of Morris et al. (2008) has taken a different direction of late, leading to the estimation of minute-to-minute variation in endogenous dopamine from M-I experiments with [11C]-raclopride. Although in this work the injections of tracer did not occur within a single scan session (as is the case for the M-I experiments discussed in this chapter), the data were analyzed simultaneously by a compound model that included compartments for tracer and endogenous neurotransmitter, separately. Finally, another twist on multiple injections was recently introduced by Converse et al. (2004) who have pioneered PET studies that include injections of both blood flow ([18F]fluoromethane) and receptor ligand ([18F]fallypride) tracers in a single scan session. Unlike the other M-I studies described herein, the approach of Converse does not require a coupling of parallel compartmental models since there is no competition between [18F]fluoromethane and [18F]fallypride for receptor sites. Nevertheless, this new method is quite innovative in its use of a complicated dual-tracer protocol which allows for easy distinction of the contributions of each tracer to the PET data.

The M-I approach to kinetic characterization of a system is based on two concepts: (1) perturbations of the system help to reveal its inner workings and (2) parallel, coupled compartmental models are necessary to describe the system’s behavior. In the following section, we outline the theory and practical application of the basic M-I protocol, which serves as the foundation for the recent advances in imaging with M-I described in Section I.

II. Introduction

Positron emission tomography (PET) is a functional imaging technique that allows an investigator to probe the biochemistry of an organism noninvasively. Because every PET study involves the injection of a radioactive molecule—a radiotracer—the practice of quantitative PET is tightly intertwined with the theory of tracer kinetics. The mission of tracer kinetics, in turn, is to estimate
physiologically relevant parameters (e.g., blood flow rate, local cerebral metabolic rate, binding or dissociation rate constants) by modeling the uptake of a labeled molecule that mimics (“traces”) the behavior of an endogenous or physiologically relevant exogenous chemical substance. Tracer kinetics merges experimentation and modeling. The experimental process involves both injection of a radioactive tracer and observation of local concentrations of said tracer over a period of time. (In PET, “observations” take the form of individual pixels in a time sequence of images, but this is secondary to our discussion.) Mathematically, what is needed is a model of the tracer uptake and its sequestration into various species and/or compartments. A formal comparison of model predictions with experimental observations yields estimates of model parameters; these parameters often represent the speed or magnitude of a physiological process.

Mere construction of a mathematical model, however, does not assure that each and every parameter of a model can be estimated from data. This goes to the identifiability of the parameters. Identifiability is hindered by noise in data or by ambiguity in the model structure. Identifiability is achieved through a combination of modeling parsimony and experimental design optimization. For example, no amount of optimization of an experiment could help identify the parameters $n$ and $m$ in the model, $Y = n \times m \times X$, where observed values of $Y$ are related linearly to measured values of $X$; $X$ is an independent and $Y$ a dependent variable. However, as discussed later, if the model is something akin to $Y = n(m - X_1)X_2$, where both $X_1$ and $X_2$ are independent variables, then by suitably varying each of them, it may be possible to collect data that will enable an investigator to identify the parameter $m$ as distinct from the parameter $n$. Roughly speaking, an appropriate experiment design would modulate the value of $X_1$ over a sufficient operating range, which is significant compared to $m$, so that the $(m - X_1)$ term did not behave effectively like the constant $m$.

The beauty of M-I PET experiments is that they are used to methodically perturb—and then observe—the system in question over a range of operating points so that the resultant data contain information that will differentiate the effects of parameters from one another. These complicated but elegant experiments (first conceived by Delforge et al., 1989, 1990) can enable the investigator to dissect out effects of otherwise highly correlated kinetic parameters such as those that describe the kinetics of a PET tracer.

What parameters do we want to distinguish? Let us focus on PET studies with receptor-ligand tracers. One can imagine that the amount of ligand bound to target receptors (specific binding) will be dependent on at least two physiological factors that have direct correlates in parameters of a kinetic model. Net receptor-mediated uptake of a tracer will be dependent on the speed of interaction of the tracer with the receptor (association and dissociation rate constants), as well as the number of the receptors in a given volume of tissue (receptor density). Not surprisingly, in the standard single-injection experimental design, the respective parameters that represent speed and number of binding sites are highly correlated. In many PET studies, it may not be necessary to distinguish the binding rate constant from the
receptor density. In such cases, it would not be necessary to mount the demanding experiments or data analysis described in this chapter. However, in those specialized situations where it is important to identify the receptor density as distinct from binding rate constants for a tracer, M-I PET studies are the only way to do so.

M-I PET experiments have been carried out to closely examine the kinetics of various ligands that bind to receptors in the brain and heart (Christian et al., 2004; Costes et al., 2002; Delforge et al., 1991, 1993, 1995, 1999; Gregoire et al., 2000; Morris et al., 1996; Muzic, et al., 1996, 2000; Poyot et al., 2001) What specialized situations might require identification of individual rate constants and receptor densities from PET data? Experiments that are intended for any of the following would be appropriate specialized applications of the techniques described herein:

- Fully evaluate the kinetics of a new tracer; determine if new tracer is different because it binds to different populations of receptors or binds to same population more avidly than established tracers.
- Accurately determine the regional variation in receptor density.
- Assess the validity of using a particular brain region as a reference region (i.e., test the assumption of no receptors but otherwise identical kinetics as a target region).
- Differentiate possible diseases of receptor (or neuronal) loss from diseases of receptor dysfunction.

This chapter is intended as a guide to graduate students, postdocs, and principal investigators who want to quickly get up to speed on key theoretical and experimental aspects of the M-I PET technique and begin to appreciate the attendant sensitivity analysis that gives the technique its power. The following discussion covers (1) the basics of the theory behind M-I PET studies, (2) practical considerations in planning and executing a successful M-I study, (3) key elements of the numerical implementation of the kinetic model and data-fitting algorithms, (4) an approach to interpretation of the parameter estimates once data are fitted, and (5) an examination of the sensitivity of PET data to the parameters and how that information can be used to improve the design of subsequent experiments.

### III. Theory

#### A. Need for Models

PET is an imaging technique that measures radioactivity indiscriminately. No distinction can be made at measurement time between radioactivity (actually the detection of two simultaneously emitted photons) that emanates from a tracer molecule flowing with the blood, free in the extracellular space, bound to a cell protein in the intracellular space, or even from radioactivity that comes from a radionuclide attached to a metabolic product of the injected tracer. All of the aforementioned
sources of radioactivity are detected and logged by the PET scanner, and all contribute to the reconstruction of a PET image. However, not all of these sources of detected signal are of equal importance to the investigator. In fact, in the case of a receptor-binding tracer molecule, the primary signal of interest is the radioactivity associated with tracer bound to a target molecule—typically a receptor or enzyme. To discern the wheat of the bound tracer signal from the chaff of the free and metabolized sources of radioactivity, the investigator must rely on a mathematical model.

B. Compartmental Models

The models used to describe PET data are usually compartmental. That is, they do not take account of spatial gradients in tracer concentration but rather assume that tissue concentrations can be properly described as well-mixed compartments. In fact, compartment models have been compared rigorously to distributed models and have been found to be satisfactory to describe PET data (Muzic and Saidel, 2003). In PET, the volume of the compartment might correspond to the volume of the voxel (if the model is being applied on a pixel-by-pixel basis) or to a larger region of interest (ROI) (if applied on the ROI level). Compartmental models are described mathematically by a series of ordinary differential equations (ODEs); one ODE is required for each compartment. Compartments typically correspond to distinct kinetic states taken on by the radiolabeled tracer. These compartments can be distinct entities physically (e.g., intra- vs. extracellular pools), distinct kinetically (e.g., bound to enzyme and bound to cell surface receptor), or distinct chemically (e.g., native vs. metabolized tracer). As long as the states represent radioactive species, they must be included in the model of the PET measurements. Sometimes, as shown later, it is necessary to model nonradioactive species as well. Usually, the tracer is introduced into the organism via bolus injection(s) and so the uptake, retention, and eventual efflux of tracer from the tissue region of interest are transient phenomena that never reach steady states. That is, the concentrations of tracer in tissue or plasma do not achieve a constant level. If the system (tracer and tissues of interest) were to reach equilibrium, the system of ODEs would reduce to a set of algebraic equations that could be solved analytically. Since this is often not the case, and certainly not true with M-I experiments, the differential equations must be solved—either analytically or numerically—to solve the model for the predicted PET activity over time in a given region (details of this procedure are given in Section V).

In most PET models of receptor-ligand interactions, we hypothesize three kinetically distinct compartments and an arterial plasma pool of tracer—all of which contribute to the measured PET radioactivity. The arterial pool is not a compartment in the mathematical sense, although it is physically distinct from tracer in tissue. Because the arterial plasma concentration in most PET studies is a measured (i.e., applied to the model as a known) quantity, its depiction does not require a differential equation. In fact, the plasma concentration (or some other input function) must exist to drive the model. If no activity is introduced into the
plasma, none is ever taken up into the tissue of interest. A version of the compartmental model corresponding to free, specifically (i.e., receptor-) bound, and nonspecifically bound tracer is shown schematically in Fig. 1 (arrows between compartments connote rate constants). \( K_1 \) and \( k_2 \) are first-order constants that are related to blood flow. The term \( k_{on} \) is a second-order rate constant describing the association of tracer and receptor; \( k_{off} \) is the dissociation rate constant. \( B_{max} \) represents the concentration of receptors available for binding. Terms \( k_5 \) and \( k_6 \) are first-order constants that measure the rates of forward and reverse nonspecific binding.

**C. Standard Model Equations**

In the language of mathematics, the “boxes” in Fig. 1 represent three unknown concentrations whose time-varying functions are encoded in their respective mass balances. The mass balances state that the change in concentration with time of species \( x \) (where \( x = F, B, \) or NS) is a function of those processes that contribute to an increase of \( x \) minus those processes that cause a loss of \( x \)

\[
\frac{dF}{dt} = K_1 C_p - (k_2 + k_5)F - k_{on}[B_{max}' - (B + B_c)]F + k_{off}B + k_6NS \\
\frac{dF}{dt} = k_{on}[B_{max}' - (B + B_c)]F - k_{off}B \\
\frac{dNS}{dt} = k_5F - k_6NS
\]

where \( C_p \) is the time-varying plasma radioactivity associated only with labeled native tracer. \( C_p \) is measured via blood samples. The state variables of the model, \( F, B, \) and NS, represent the time-varying concentrations of tracer (in pmol/ml) in

**Fig. 1** Standard compartmental model used to describe dynamic PET data. The PET pixel is indicated to show that the measured quantity is a weighted sum of radioactivity in the compartments \( F \) (free), \( B \) (bound), NS (nonspecific), and some amount in the blood. \( P \) indicates that the metabolite-corrected plasma concentration is not a compartment because it is measured separately from the PET images and is assumed to be known.
free, bound, and nonspecifically bound states, respectively. \( B_c \) is the concentration of unlabeled (or “cold”) tracer bound to receptors.

The part of the model of greatest interest to investigators of receptor binding is
\[
k_{on} \left[ B_{\text{max}} - (B + B_c) \right] F.
\]
This expression describes binding of free tracer to available receptors. It states that the concentration of available receptors is the difference between available receptors at steady state, \( B'_{\text{max}} \) (a constant), and receptors bound to either labeled, \( B \), or unlabeled tracer, \( B_c \) (time-varying functions). Note: In a single-injection experiment, there is always a known relationship between labeled and unlabeled bound ligand. The ratio of the labeled to the unlabeled is the specific activity (SA is given in \( \mu \text{Ci}/\text{pmol} \) or \( \text{Bq}/\text{pmol} \), ratios of radioactivity to mass of ligand). Thus, the expression for available receptors is often written as
\[
(B'_{\text{max}} - B/SA).
\] It will be clear why this is not adequate for modeling M-I PET data (Morris et al., 1996).

Because the binding of a ligand to a receptor is a bimolecular process, it depends on available receptors, the presence of free ligand, \( F \), and a bimolecular rate constant, \( k_{on} \). In conventional single-injection experiments, which are predicated on injecting only a tiny (“trace”) amount of radioligand, the amount of bound tracer (labeled or unlabeled) never rivals the available sites at steady state and so the term of interest, \( k_{on} \left[ B'_{\text{max}} - (B + B_c) \right] \), reduces to \( k_{on} B_{\text{max}} \). In this case, the model is analogous to the example described in Section II. Namely, the parameters \( k_{on} \) and \( B'_{\text{max}} \) are not uniquely identifiable and the parameter estimation problem is reduced, of necessity, to finding an effective first-order rate constant \( k_3 = k_{on} B_{\text{max}} \).

The raison d’etre of M-I PET experiments is specifically to overcome the problem of \( k_{on} \) and \( B'_{\text{max}} \) being irretrievably correlated (i.e., unidentifiable). Why do we want to identify these parameters separately? For one, the equilibrium dissociation constant (affinity constant) \( K_D \) is the ratio of the rate constants \( k_{off} \) and \( k_{on} \). Thus, estimation of the in vivo \( K_D \) for a PET ligand is effectively dependent on the estimation of \( k_{off} \) and \( k_{on} \). Estimation of these two constants—or their ratio—is not possible from a single-injection PET study. However, if the injected mass of tracer is modulated sufficiently in the course of multiple bolus injections of tracer such that the occupancy of receptors varies over a large enough range, then the term \( (B'_{\text{max}} - B - B_c) \) must be retained explicitly in the model. It then becomes possible to identify the unique roles of the association rate constant and the concentration of available receptors in the uptake and binding of a tracer. In doing so, we move toward being able to estimate the receptor number and the affinity constant separately and possibly toward using PET to distinguish a defect of receptor function from a defect of receptor number.

D. M-I Model Equations

How do we adapt the standard model equations to accommodate the description of M-I data? One approach is to treat the separate injections as separate species that compete for the same receptor sites. Figure 2 diagrams the case of three
separate bolus injections of tracer. The important thing to recognize about this extension for M-I data is that the specific activity of each injection is intentionally different. Therefore, it will be necessary to somehow track the individual inputs over the entire course of the study. This is one of the subtleties of the M-I PET technique. A blood sample taken shortly after a third bolus injection will contain radioactivity that originates with each of the three injections (assuming that all three injections contain radioactivity). Figure 3 depicts the multiple injections in terms of measured radioactivity and in terms of molar quantities needed for solving the model. One approach to \textit{a posteriori} dissection of the measured blood radioactivity is discussed at length by Morris \textit{et al.} (1996). In brief, we might assume that all input functions have the same shape but different scales. Thus, the observed plasma radioactivity can be described as

\[
C_p(t) = \sum_j S^j C^j_p(t - T^j) U(t - T^j)
\]  

\[\text{Eq. (4)}\]

where \(S^j\) is a scale factor related to injected dose, \(C_p(t)\) is an analytical expression of exponentials, and \(U(t - T)\) is the unit step function at time \(T\) (i.e., \(U = 0, t < T; U = 1, t \geq T\)). From Eq. (4), it is possible to recover separate \(C^j_p\) curves for each injection from the measured plasma radioactivity.
Fig. 3  (A) Input functions for each injection in terms of total ligand concentration in pmol/ml (all species (solid) and metabolite-corrected (dotted)). Metabolite-corrected molar concentrations are used to construct the input function (see text for details). (B) Input functions in terms of radioactivity concentration (nCi/ml). The third injection consisted of unlabeled ligand only; therefore, there is no peak of radioactivity at the time of the third injection. There is input to the system, however, that must be measured somehow or modeled based on the shape of the other input functions as described in the text. Injection times are indicated by vertical arrows.
The general model equations for multiple injections take on the following form:

\[
\frac{dF^j}{dt} = K_1 C_p^j - (k_2 + k_5)F^j - k_{on} \left( B_{\text{max}}^j - \sum_l B_l^j \right) F^j + k_{off} B^j + k_6 N S^j \tag{5}
\]

\[
\frac{dB^j}{dt} = k_{on} \left( B_{\text{max}}^j - \sum_l B_l^j \right) F^j - k_{off} B^j \tag{6}
\]

\[
\frac{dN S^j}{dt} = k_5 F^j - k_6 N S^j \tag{7}
\]

where \( j \) is the index over injection number; \( B_l^j \) is either \( B_l^i \) or \( B_c^i \).

The assumption of Eqs. (5)–(7) is that the kinetic parameters \( (K_1, k_2, k_{on}, k_{off}, B_{\text{max}}, k_5, k_6) \) are unaffected by the injection of either a high- or a low-specific activity tracer.

From the mass balance equations in Eqs. (5)–(7), we can construct the instantaneous output equation to describe the total radioactivity, \( T(t) \), measured in the tissue at any moment in time as a result of one or more injections:

\[
T(t) = \sum_j S A^j(t)(1 - F_v)[F^j(t) + B^j(t) + N S^j(t)] + F_v C_{wb}(t) \tag{8}
\]

where \( S A^j \) converts the concentration associated with each injection in the tissue \( (F, B, \text{and NS are in pmol/ml}) \) into radioactivity, \( F_v \) is the blood volume fraction, and \( C_{wb} \) is the radioactivity concentration (nCi/ml) in whole blood. \textit{Note:} The concentration of tracer in the arterial plasma (pmol/ml) is the driving force for uptake of the tracer into the tissue and, hence, the appropriate input function. However, any radioactivity in the microvasculature in the ROI contributes to the PET signal, so \( C_{wb} \) is the appropriate term for the output equation, which is in the units of the PET measurement (nCi/ml).

---

**IV. Experimental Protocol and Considerations**

M-I PET studies are sophisticated experiments, both in design and in implementation. The appropriate duration for the experiment is dependent on both the kinetics of the ligand and the half-life of the positron-emitting nuclide. In nuclear-counting experiments, the noise in the data and hence, the parameter precision are determined by the amount of radioactivity (i.e., photons) collected. From this point of view, it would be desirable to acquire PET data for as long as possible. Unfortunately, due to the expenses associated with reserving PET scanner time, veterinarian staff, and anesthetization of the animal, one of the primary considerations when designing an M-I experiment is to minimize the duration of the experiment. On the other hand, because we use these experiments to maximize precision
of the parameter estimates, there is a trade-off between convenience and precision. Successful experimental optimization can help balance this and other trade-offs and achieve a desired level of parameter precision.

A. Animal Preparation

Conducting M-I studies requires a small team of personnel to ensure a successful experiment. Input from all of the team members is needed to carefully plan and design an experiment that follows the guidelines of the Institutional Animal Care and Use Committee (IACUC). A skilled veterinary staff is needed to anesthetize the animal and to insert the catheter lines for ligand injection (venous) and blood withdrawal (arterial). In general, the choice of anesthetic is determined by the investigator and the veterinary staff based on their familiarity with the anesthetic agent, ease of use, and animal safety considerations. Because M-I PET experiments measure tiny (subnanomolar) concentrations, care must be taken that any biochemical effects of the anesthetic drugs do not perturb the biochemical system under study. For example, ketamine is a widely used preanesthetic known to interact with the dopaminergic system of the brain (Smith et al., 1998); therefore, a M-I study targeting the dopaminergic system should allow adequate time (>1 h) for the effects of ketamine to subside before administration of the PET ligand. Most M-I studies require a minimum of 3 h of animal anesthetization, including animal preparation (see, e.g., timeline in Table I). The entire experiment can last up to 12 h. The anesthetic must provide a stable physiological system throughout this time course, it must minimize changes in regional blood flow (which affects ligand delivery), and it must withstand possible drug-induced stimulation. For the safety of the animals, typically 1–2 weeks must be allowed between experiments for the animal to recover from the effects of the anesthesia.

The insertion of two catheters is needed for M-I studies: a venous port for the administration of ligand and an arterial port for the temporal sampling of plasma radioactivity. The ligand is generally administered into a vein as a bolus infusion (5–30 s in duration) in several milliliters of saline. A bolus infusion is needed to accurately identify the ligand delivery parameter ($K_1$). As nearly as possible, all injections of ligand for each experiment should be given in an identical fashion so that ligand delivery is consistent throughout each epoch of the experiment (see Section IV.B).

B. Measuring Blood Activity and Constructing Input Curves

Accurate measurement of the radioligand concentration over time in the arterial plasma is essential to precise estimation of model parameters. The input function provides the essential time-varying details of radioligand delivery to the tissue of interest. An example of three arterial plasma input functions (in terms of both molar concentration (Fig. 3A) and radioactivity (Fig. 3B)) and the plasma radioactivity curves from which they are derived are depicted in Fig. 3. The graph in
Table I
Example of a Timeline for a Multiple-Injection Pet Experiment

<table>
<thead>
<tr>
<th>Time</th>
<th>Event Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00:00</td>
<td>Preanesthesia with glycopyrolate (0.01 mg/kg)</td>
</tr>
<tr>
<td>00:30:00</td>
<td>Anesthesia with ketamine (10 mg/kg)/xylazine (0.5 mg/kg)</td>
</tr>
<tr>
<td>00:45:00</td>
<td>Intubate monkey and maintain with 1–2% isoflurane</td>
</tr>
<tr>
<td>00:60:00</td>
<td>Insert venous (saphenous) and arterial (femoral) catheters</td>
</tr>
<tr>
<td>01:30:00</td>
<td>Position monkey in PET scanner, monitor vitals</td>
</tr>
<tr>
<td>01:45:00</td>
<td>Acquire 10-min transmission scan for attenuation correction</td>
</tr>
<tr>
<td>02:00:00</td>
<td>Begin PET data acquisition</td>
</tr>
<tr>
<td>02:00:00</td>
<td>Injection #1, “tracer study” with high-specific activity injectate</td>
</tr>
<tr>
<td></td>
<td>(SA1 ~2000 mCi/μmol); withdraw blood samples periodically for analysis (~1 ml each)</td>
</tr>
<tr>
<td>02:54:00</td>
<td>Injection #2, “partial saturation” with low-SA injectate</td>
</tr>
<tr>
<td></td>
<td>(SA2 ~100 mCi/μmol); take blood samples</td>
</tr>
<tr>
<td>03:38:00</td>
<td>Injection #3, “saturation” with unlabeled fallypride only</td>
</tr>
<tr>
<td></td>
<td>(SA3 = 0 mCi/μmol); take blood samples</td>
</tr>
<tr>
<td>05:00:00</td>
<td>Terminate PET acquisition; remove anesthesia</td>
</tr>
<tr>
<td>05:15:00</td>
<td>Remove intubation when gag reflex is recovered</td>
</tr>
<tr>
<td>06:00:00</td>
<td>Monitor monkey during recovery from anesthesia</td>
</tr>
</tbody>
</table>

*The protocol was designed to elicit a precise estimate of \( B_{\text{max}} \), available receptors in the thalamus that bind \([^{18}\text{F}]\)fallypride. In this particular design, the last injection contains only unlabeled fallypride. Note the absence of a “hot” peak at the corresponding third injection in Fig. 3B.*

Fig. 3B shows measurable quantities of radioactivity (in nCi/ml) in the plasma. The plot displays two different measured quantities: total radioactivity in arterial plasma (solid curve) and metabolite-corrected arterial radioactivity (dash-dot line). The latter, corrected plasma concentration, is data needed for each of three input functions that drive the model, whereas whole blood radioactivity measurement (not shown) is needed for solution of the output equation (see Section III). With knowledge of the specific activity, blood data corresponding to each of the injections can be converted to input functions for the total (molar) ligand concentration, as shown in Fig. 3A. In the case of an experiment that includes a “saturation” component, the contribution of the third injection to the total ligand input function, \( C_p(t) \), cannot be measured directly from the (radioactivity) blood curve, as no additional radioactivity is injected (see earlier discussion). Instead, the shape of the curve must be inferred from the previous injections. The scale factor, \( S \), for the injection of unlabeled material can be determined from the ratio of the doses.

The shape of the input function is, in part, determined by the speed of the venous injection (e.g., a rapid bolus injection will result in a sharply peaked input function (blurred by dispersion as the bolus travels through the vasculature)). The blood curve is measured by withdrawing blood samples from the arterial port. The frequency of withdrawal must be matched to the anticipated shape of the input function. Blood samples are usually drawn every 5–10 s for the first several minutes following ligand injection. The samples can be drawn less frequently as the ligand begins to equilibrate between plasma and tissue(s).
As a general rule of thumb, the total volume of blood withdrawn for an experiment should not exceed 10% of the blood volume of the animal. Sampling of a 10-kg rhesus monkey would be limited to roughly 70 ml (assuming 7% of body weight is blood volume). This volume should be replaced by an iv drip of saline over the course of the experiment. The arterial blood samples are centrifuged to separate the plasma from the red blood cells. The plasma samples can be assayed further to separate the native ligand from the radiolabeled metabolic by-products. The volume of each arterial plasma sample must be large enough to yield an accurate measurement of radioactivity in the final plasma fraction (as gauged by radioactive counting statistics). In the case of primates, the volume of each arterial sample is typically 1 ml.

C. Generation of Regional Time-Activity Curves (TACs) from PET Images

As in most dynamic PET experiments, the M-I model is fitted to tissue TACs derived from PET scans yielding estimates of the kinetic parameters of interest (see Section V). Data at each time point are based on investigator-defined regions of interests (ROIs) placed on the PET images at each time. Although single-injection PET studies are sometimes analyzed in a pixel-by-pixel manner (to generate parametric maps), fitting the M-I model to data would be too demanding computationally to do so.

In most cases, ROI analysis requires that high-resolution structural image (typically a T1-weighted MRI) is acquired and coregistered to a PET image for each PET subject. The preferred PET image (for use in registration only) is usually an image of the summed (or averaged) radioactivity over the entire duration of the PET study. The MRI allows the investigator to precisely define the exact anatomical location of ROIs, and the coregistration then gives coordinates that allow the ROIs to be applied in proper spatial orientation to each frame of PET data. Taking the average radioactivity in the ROI at each time point generates the desired TAC. Typically, PET images are not suitable as the basis for ROI templates for two reasons: (1) specific anatomy may not be resolved easily and (2) hot spots in the PET image may induce unintentional bias in the ROI placement by the investigator. Several intermodality registration algorithms (e.g., automated image registration (AIR) by Woods et al., 1992, 1993) are available both in commercial medical image analysis packages (e.g., MEDx, Medical Numerics, Germantown, MD) and as stand-alone code for free download (http://bishopw.loni.ucla.edu/AIR5). The standard is now to use registration software based on a mutual information algorithm (available, e.g., in recent versions of SPM, see http://www.fil.ion.ucl.ac.uk/spm/software/).

The ROIs selected will probably depend on the characteristics of the radiotracer used. \([^{18}F]Fallypride\) is a highly selective D2/D3 receptor ligand with a high-PET signal-to-noise ratio and excellent resolution in areas with low-to-moderate concentrations of D2 receptors (e.g., cortex, thalamus) (Mukherjee et al., 1995). In investigations with this ligand, several brain regions are available for analysis (in contrast, other ligands may not provide reliable data outside of brain regions with extremely high numbers of D2/D3 receptors (e.g., striatum) and ROIs of interest may range from very large volumes (e.g., whole striatum, whole thalamus)
to smaller, more specific volumes (e.g., cortical areas, amygdala, nucleus accumbens, individual thalamic areas). An important point to consider in selecting ROIs is their volume. Larger volumes (composed of many voxels) result in TACs with higher signal-to-noise ratios. These low-noise curves typically lead to successful data fitting and thus precise parameter estimation. Unfortunately, larger ROIs are also more prone to be heterogeneous in tissue composition and therefore lead to data that reflect an average of kinetically different regions. The investigator must also be aware that small ROIs may suffer from partial volume (PV) effect error if the structures they circumscribe are small relative to the resolution of the scanner. PV error will lead to nonlinear underestimation of the true radioactivity in the ROI and, subsequently, bias in the parameter estimates. For a review of PV error and for various approaches to correcting for it, see Kessler et al. (1984), Meltzer et al. (1996), Morris et al. (1999), Muller-Gartner et al. (1992), Muzic et al. (1998), Rousset et al. (1993, 1996, 1998), Strul and Bendriem (1999).

The resulting TACs from three brain regions generated by placing ROIs on images made in the example protocol given earlier (Table I) are shown in Fig. 4. Specific binding is highest in the striatum, moderate in the thalamus, and nearly absent in the cerebellum. We can tell this, in part, from observing (1) similarity of the decline in tracer concentration following the first and second injections and (2) the absence of deflection from the descending curve of the cerebellum at the time of the third injection. That is, injection of high- or low-specific activity does

![Fig. 4](image-url)
not change the shape of the curve because of (1) the absence of receptors in the region and (2) the injection of cold tracer does not accentuate displacement because there is no specific binding to be displaced. The next sections address the fitting of models to data and present a resulting fit to one of the curves in Fig. 4.

V. Models and Data Fitting

A. Implementing Model Equations

This section describes the framework for relating the theoretical models (described earlier) to data that are collected (as described earlier). Since a PET scanner measures radioactivity concentration, the output equation, Eq. (8), is used to relate the radioactivity concentration (e.g., Bq/ml or μCi/ml) to the molar concentration (e.g., pmol/ml) in various compartments. However, a modification to Eq. (8) should be used in practice, as the PET scanner does not measure instantaneous radioactivity concentration. Rather, it measures concentration averaged over acquisition time intervals commonly referred to as frames. Thus, we define model$_i$ as the time-averaged concentration over frame $i$ and compute this quantity as

$$\text{model}_i = \frac{1}{d_i} \int_{t_i}^{t_i+d_i} T(\tau) d\tau$$  \hspace{1cm} (9)

where $t_i$ and $d_i$ are the start time and durations of frame $i$ and $T$ is from Eq. (8).

To implement such an equation, it is convenient to express it in a form that can be solved with an ODE solver such as is used to solve the state equations (Eqs. (5)–(7)). Accordingly, we introduce a new expression for the integrand in Eq. (9):

$$\frac{dh}{dt} = T$$  \hspace{1cm} (10)

When this differential equation, Eq. (10), is solved with initial condition $h(0) = 0$ (no radioactivity in the system at time zero), the expression for the model-predicted PET signal in time frame $i$ is simply

$$\text{model}_i = \frac{1}{d_i} [h(t_i + d_i) - h(t_i)]$$  \hspace{1cm} (11)

B. Parameter Estimation

In M-I studies, estimating values of model parameters as precisely as possible is often the primary goal. Parameter values provide quantitative assessments of receptor concentration, affinity, blood flow, etc. To accomplish the goal, one typically “fits a model” to data. This entails finding the values of model parameters that are most consistent with data. Mathematically, the problem is equivalent to
adjusting the values of the model parameters in order to minimize the difference between the model prediction and the actual measurement of tissue radioactivity (“data”). Consider what is called the weighted least-squares objective function, which measures this difference:

\[ o(p) = \sum_i \left[ w_i (\text{model}_i(p) - \text{data}_i)^2 \right] \]  

(12)

where \( \text{data}_i \) represents data from frame “\( i \)” and \( \text{model}_i(p) \) represents the model output, which is intended to predict \( \text{data}_i \). Model output depends on the values of the parameter vector \( p \). The task is then to adjust values of the components of \( p \) to make model output most closely agree with data. Mathematically, we minimize \( o(\cdot) \) with respect to \( p \). We include weights, \( w_i \), because we do not expect to achieve perfect agreement between model and data and because we do not have uniform confidence in the measurements. The value of \( w_i \) may be specified as the reciprocal of (an estimate of) the variance of \( \text{data}_i \), in which case the value of \( p \) that minimizes \( o(\cdot) \) is a maximum likelihood estimate.\(^1\)

While one could adjust the values of model parameters manually to find the ones that best explain data, such a search can often be done more efficiently and more objectively by a mathematical algorithm implemented on a computer. The Levenberg-Marquardt algorithm is popular for this application (Levenberg, 1944; Marquardt, 1963; More, 1977) Because the value of \( o(p) \) has a complex dependence on values of the parameters, a closed form solution, which minimizes \( o(p) \) in the general case, is not available. Consequently, an iterative approach must be used. One starts with an initial guess of the parameter values and then adjusts the components of the parameter vector \( p \) in order to reduce the value of \( o(p) \). The efficiency of this process depends on having a means to predict values of \( o(p) \) as values of \( p \) are altered because this provides a basis for adjusting parameter values. For this purpose, algorithms often require an estimate of the derivative of the objective function \( o(\cdot) \) with respect to the parameter vector \( p \). By differentiating Eq. (12) with respect to component \( j \) of the parameter vector, we obtain the expression

\[ \frac{d o}{d p_j} = 2 \sum_i \left[ w_i (\text{model}_i - \text{data}_i) \frac{d \text{model}_i}{d p_j} \right] \]  

(13)

Notably, this expression contains a term for the derivative of the model output with respect to component \( j \) of the parameter vector. These derivatives have particular significance and are given the name sensitivity functions.

One numerical approach to evaluating the sensitivity functions is to use finite differences. This approach is attractive because it is conceptually very simple: solve the model equations at one value of \( p \), change the value of the \( j \)th component of \( p \) by a small amount denoted here as \( \Delta p_j \), solve the model equations again, and then estimate the derivative as the difference in model output divided by \( \Delta p_j \).\(^1\)

\(^1\) Under certain assumptions about the data.
Unfortunately, in practice, this approach is not very robust. It is not at all trivial to pick a value of $\Delta p_j$ small enough so that the finite differences approximate the desired derivative but not so small that the differences are dominated by “noise” or numerical imprecision.

A more robust—and recommended—approach to evaluating the sensitivity functions can be obtained by differentiating the state equations. To describe this approach, we have to take a step back and define notation for the composite set of differential equations for the state and output equations (Eqs. (5)–(7) and (10)). Recall that the equations were all of the form $dx/dt = y$ with accompanying specified initial conditions. We can group these together by defining a vector $c$ and a vector-valued function $f(\ )$ that have components corresponding to the state and output equations and their variables. In the example given earlier, the vector $c$ would have components

$$c = [F \ B \ NS \ h]^T$$  \hspace{1cm} (14)

and the function $f(\ )$ would be defined as

$$\frac{dc}{dt} = f(c, t, p) = \left[ \frac{dF}{dt} \ \frac{dB}{dt} \ \frac{dNS}{dt} \ h \right]^T$$  \hspace{1cm} (15)

with

$$p = [K_1 \ k_2 \ k_{\text{on}} \ k_{\text{off}} \ B_{\text{max}} \ ...]^T$$  \hspace{1cm} (16)

The superscript $T$ connotes the transpose; Eqs. (14)–(16) describe column vectors. The initial condition for $c$, called $c_0$, is a column vector of the initial conditions of each of the state equations.

With the state and output equations expressed in this framework, we now obtain the equations needed for a robust approach to evaluating the sensitivity functions.

Specifically, by differentiating Eq. (15) and its initial condition with respect to the parameter vector $p$ we obtain

$$\frac{dS}{dt} = \frac{\partial f}{\partial c} S + \frac{\partial f}{\partial p} \quad \text{with} \quad S_0 = \frac{\partial c_0}{\partial p},$$  \hspace{1cm} (17)

which is an initial value problem like Eq. (15) except that $S$ is a matrix. The rows of $S$ correspond to those of $c$, whereas the columns of $S$ correspond to different components of the derivatives. For example, the element in row 2, column 3 of $S$ would be the derivative of $B$ (element 2 of $c$) with respect to $k_{\text{on}}$ (element 3 of $p$).

C. Numerical Solution of Differential Equations

Having presented a formalism of how we relate a model to experimental measurements, we next turn to the details of the numerical implementation of the solution of state, Eq. (15), and sensitivity equations, Eq. (17). These are both
considered initial value problems. Numerically solving state equations entails programming Eq. (15) and selecting an appropriate ODE solver. For example, in MATLAB one could use a solver from Shampine’s ODEsuite (Shampine and Reichelt, 1997), whereas in C or FORTRAN one might use a member of the LSODE family of solvers (Hindmarsh, 1983; Hindmarsh and Serban, 2002; Leis and Kramer, 1988).

Conceptually, algorithms for solving these initial value problems begin with the initial value and use the Euler formula to approximate the solution at the next time step. For example,

\[ c(t + \Delta t) \approx c(t) + f(c, t, p) \cdot \Delta t. \]  

(18)

Details of the implementation must include a strategy for selecting \( \Delta t \) to achieve a specified accuracy in the solution without requiring an excessive amount of computation. Fortunately, problems of this form are common and a number of algorithms are available. Generally, algorithms are classified as being designed for “stiff” or “nonstiff” equations. Details of these designations are beyond the scope of this chapter, but suffice it to say that “stiff” equations are “hard” problems to solve in that the solver is forced to take very small steps in \( \Delta t \). Special algorithms have been designed for stiff equations. In comparison to nonstiff solvers, stiff solvers trade-off more complex algorithms and evaluations in each step for the ability to take larger steps.

How does one determine if equations are stiff in any given case? The pragmatic approach is to try both stiff and nonstiff solvers. Well-written solvers have built-in methods to select step size (\( \Delta t \)) and still keep errors in the solution within a specified range. Under such conditions, using a nonstiff solver with stiff equations (and vice versa) would lead to computationally inefficient solutions.

We alert the reader to the availability of a MATLAB-based software package that implements methods for setting up and solving models such as those used to analyze dynamic PET data. The package includes implementations of state and sensitivity equations and functions for fitting models to data in order to estimate parameters. COMKAT (Muzic and Cornelius, 2001) can be downloaded from http://comkat.case.edu. It was written by one of the authors of this chapter (R.F.M.) and is presently used by each of the authors in their research. COMKAT takes into account the details described in the preceding section so that its users do not have to be experts in numerical analysis.

D. Parameter Estimation Considerations

1. Selection of the Initial Guess

As mentioned in the previous section, algorithms estimate parameters by starting with an initial guess of the parameter values and adjusting them to minimize the value of the objective function. Care should be taken in selection of the initial guess. Algorithms often converge to the true parameter values only when the initial
guess is “close enough” to the true values. How close is “close enough” is difficult to quantify in practice because it depends on the true parameter values, which are unknown, and also on the information content of data. In practice, as one is developing the fitting strategies for a particular application, one should try the estimation procedure with a range of initial guesses. Analysis of the resultant parameter estimates will provide insight into how close is “close enough.”

2. Validating Parameter Estimates

When the optimization algorithm has converged to parameter values and the model output and data are in close agreement, one might assume that the parameter estimates are valid and even precise. This is not always the case. For example, there could be more than one set of parameter values that produce a model output that agrees well with data. One possibility is that there are multiple local minima in the objective function $o()$. Another possibility is that the parameters are correlated, meaning that different combinations of parameter values will lead to essentially the same model output. Consider plotting $o()$ as a function of values of two parameters with the height of the surface indicating the value of $o()$. If the surface is relatively flat, then a large change in the parameter values would give rise to a small change in $o()$. To achieve good precision in the parameter estimates, we would like to design the experiments to make the surface of $o()$ steep. A steep objective function means that data are very sensitive to the model parameters. Moreover, we want the surface to be steep in all directions. Consider an alternative case wherein the surface is shaped like a long narrow valley aligned with the parameter axes. Changes along the valley floor make hardly any difference in the value of $o()$. An experiment that yielded such an objective function would be insensitive to the parameter aligned with the valley, and it would not be possible to make reliable estimates for this parameter.

To investigate these possibilities, it is important to conduct simulation studies a priori. The basic steps of the study are as follows. (1) Create data; using representative parameter values, solve the model equations to create “perfect” data. (2) Add “noise” to perfect data to emulate the expected imprecision in the experimental data. (3) Fit simulated data with the proposed parameter estimation method. This process must be repeated numerous times with different noise realizations. Parameter estimates are then compared to the known values used to create data. In particular, one might calculate the error in the parameter estimates by subtracting the true values from each estimate and then summarize data in terms of the bias and precision of the estimates by calculating the mean and standard deviation of the error.

While the aforementioned techniques are important in establishing the validity of the parameter estimates, they are not necessarily complete. Simulation is but one component of validation. The next section describes another component: careful examination of the model fit to measured data.
VI. Results and Interpretation

A fit to the TAC for an ROI drawn on the thalamus (middle curve in Fig. 4) is shown in Fig. 5. This fit results in estimates of the “best” parameter values that can explain data, but how do we know if these estimates are good? There are tests that must be done. One was suggested in the previous section; namely, if fitting the model to simulated data sets reveals that multiple choices of parameters would result in equally good fits, there is little hope that fits to experimental data (which may not be strictly consistent with the model) will yield more identifiable parameters. However, assuming that the model appears to fit simulated data well and that minimizing the objective function yields unique parameters, what are the basic steps that must be followed to evaluate the quality of the results?

A. Examination of Residuals

Figure 6 shows a plot of normalized residuals derived from the fit to M-I data shown in Fig. 5. Normalized residuals, calculated as \( \frac{\text{model}_i - \text{data}_i}{\text{S.D.}(\text{data}_i)} \), are a good way of determining the quality of the fit. Ideally, these residuals should be distributed normally with zero mean and unit standard deviation. Both of these conditions appear to be met in Fig. 6. The mean of the residuals will be obvious...
from their plot. Any order in the pattern of the residuals, however, may indicate that the model is deficient. To determine the nonrandomness of the residuals it is useful to perform a “runs” test. A run is defined as a series of adjacent residuals that are either all positive or all negative. The fewer the number of runs, the more likely that the fit is poor and that either the model or the fitting algorithm is suspect (for examples using the runs test, see Bard, 1974).

B. Parameter Precision

If fits to data are acceptable, then the investigator will want to report his/her findings in terms of the estimated parameter values and their approximate uncertainties (variance, standard deviation, confidence intervals, correlation, etc.). To report intratrial variance, it is necessary to use an estimate of parameter variance because one fit to a data set yields only one estimate of each parameter. Many search algorithms (such as the Levenberg-Marquardt mentioned previously) will return a covariance matrix along with the optimal parameter set. The covariance matrix is usually approximated by the inverse of the weighted product of the sensitivity matrix (mentioned in the previous section) with its transpose (\( \text{Cov}(\mathbf{p}) = [S^T WS]^{-1} \), where \( W \) is an \( n \times n \) diagonal matrix whose elements are related inversely to the \( n \) data points in the TAC). The approximation is valid when the parameter values \( \mathbf{p} \) are close to the optimal point. The diagonal elements of the

Fig. 6  Plot of the normalized residuals derived from fitted M-I thalamus data (shown in Fig. 5). Note the apparent zero-mean behavior of the residuals. See text for details.
resulting covariance matrix are the variances of the respective parameters. Parameters should be reported plus or minus a standard deviation (±S.D.) (square root of the variance).

Normalization of the covariance matrix by its diagonal elements yields the correlation matrix. It is prudent to examine this matrix, whose diagonal elements are unity and whose off-diagonal elements are the covariances between parameters. Highly correlated parameters (e.g., correlation >0.95) are not separable. See Table II for an example correlation matrix produced from a six-parameter fit to [18F]fallypride data shown in Fig. 5. If two parameters \( a \) and \( b \) are highly correlated, then, in practical terms, only their product (or their ratio) is identified. Neither of their values individually should be trusted because an increase in one could be completely offset by a comparable decrease (or increase) in the other with no decrement to the quality of the fit and no basis for choosing one combination of parameters over another with the same product. Consider the following practical scenario. If \( B'_{\text{max}} \) is highly correlated with \( k_{\text{on}} \) (as is often the case), then there will be multiple pairs of these parameters that will be equally plausible choices to explain the acquired data. Imagine further that we are trying to compare the on rate \((k_{\text{on}})\) of a tracer at the serotonin transporter site in two groups of subjects, who are known to express different genetic variants of the transporter, to test the hypothesis that the binding rate will be different. If one of the groups also tends to have fewer available receptors at steady state (smaller \( B'_{\text{max}} \)) because of medication that blocks these sites (e.g., Prozac), then the medication will be a confound and the population on medication may be seen, artifactually, to have faster binding because the higher \( k_{\text{on}} \) merely balances a lower \( B'_{\text{max}} \) when data are fitted. The correlation matrix in Table II confirms that, thanks to the M-I experiment, correlations among \( k_{\text{on}}, B'_{\text{max}}, \) and \( k_{\text{off}} \) have all been minimized. In contrast, in a single-injection experiment, the correlation between \( k_{\text{on}} \) and \( B'_{\text{max}} \) would be nearly 1.

### Table II
Correlation Matrix for the Data Fit Depicted in Fig. 5\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>( K_1 )</th>
<th>( k_2 )</th>
<th>( k_{\text{on}} )</th>
<th>( B'_{\text{max}} )</th>
<th>( k_5 )</th>
<th>( k_{\text{off}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K_1 )</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>( k_2 )</td>
<td>0.809</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>( k_{\text{on}} )</td>
<td>–0.632</td>
<td>–0.181</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>( B'_{\text{max}} )</td>
<td>0.612</td>
<td>0.751</td>
<td>–0.197</td>
<td>1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>( k_5 )</td>
<td>0.13</td>
<td>0.496</td>
<td>0.429</td>
<td>0.604</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>( k_{\text{off}} )</td>
<td>–0.374</td>
<td>–0.211</td>
<td>0.416</td>
<td>0.162</td>
<td>0.635</td>
<td>1</td>
</tr>
</tbody>
</table>

\(^a\)Each element in the matrix is Corr\((a, b)\). Diagonal elements are 1 because each parameter is completely correlated with itself. Thanks to the M-I experiment, there is very little correlation among any of the parameters \( B'_{\text{max}}, k_{\text{on}}, \) and \( k_{\text{off}} \) (see italic values). The correlation matrix is symmetric, so the top half of the matrix has not been shown; Corr\((a, b) = Corr(b, a)\).
C. Model Selection/Goodness of Fit

Often, even well-designed experiments produce data that do not justify the use of models of the desired complexity. That is, not all parameters of the model can be identified. In these cases, it may be necessary to opt for a simpler model by fixing some parameters and not estimating them. How do we know that the simpler model is appropriate? There are a number of popular criteria that gauge “goodness of fit.” One such determinant of goodness of fit is the $F$ statistic (Landaw and DiStefano, 1984). As in all statistical testing, it is conducted with reference to the question of whether to accept or reject the null hypothesis. In the case of model selection, the null hypothesis is that the simpler model is adequate to describe data. Another popular index is the Akaike criterion (Akaike, 1976), $AIC = \ln(SS) + 2P$, where SS is the weighted sum of squares that result from the fit to data and $P$ is the number of parameters in the model. Thus, a “good fit” will correspond to a low-$AIC$ value, but $AIC$ will be penalized if the fit is achieved through the use of extraneous parameters. In the case of the fit to $[^{18}F]$fallypride data shown for the thalamus in Fig. 5, data did not support use of both a $k_5$ and a $k_6$ parameter. It was found that setting $k_6$ identically to zero and estimating only $k_5$ was necessary and sufficient to fit data. To confirm that this was the appropriate model, the Akaike criterion was calculated for both six-parameter ($k_6$ set to 0) and seven-parameter fits, and the six-parameter fit was shown to be better.

VII. Understanding and Designing M-I Experiments

A. Sensitivity Functions

The sensitivity functions described earlier are key to the procedure of minimizing the least-squares objective function. They are also central to understanding how data fitting in general and the analysis of M-I data in particular work. Once we understand what the sensitivity functions tell us, we can use them to improve the design of our experiments. For an example, consider the sensitivity functions plotted in Fig. 7. These curves correspond to the derivatives of the model with respect to the six parameters, $p$, that were estimated by fitting the model to data from thalamus (shown in Fig. 5). The sensitivity equations have been solved at the value of the parameters that minimized the objective function given in Eq. (12). First, we noted that the sensitivities are time-varying functions as we would expect from looking at Eq. (17). In other words, the sensitivity of the observed PET signal to any model parameter rises and falls throughout the course of the experiment. The PET signal may be most sensitive to one parameter at one moment and to another at the next. Early time data are usually the most sensitive to blood flow parameters (i.e., $K_1$, $k_2$), whereas late-time data are sensitive to receptor binding. In fact, the independent, time-varying status of each of the sensitivity functions is at the heart of parameter identifiability. In Fig. 7A and B, one can observe that the sensitivities to the $K_1$ and $k_2$ parameters are very nearly identical except that one is
always positive and the other always negative. We explain this behavior by noting that these two parameters are both dependent on blood flow ($K_1$ = extraction fraction × flow; $k_2 = K_1$/volume of distribution). As blood flow increases, more tracer is delivered to the tissue and, hence, the effect on the measurable signal is a positive one. An increase in blood flow also means that $k_2$, the rate at which tracer leaves the tissue, will increase and we would expect the PET signal to be diminished. The fact that these time courses mirror each other so closely means that they are not independent; in fact, they are nearly linearly dependent and so the parameters are highly correlated (Table II, Corr($K_1$, $k_2$) = 0.809). Thus, these parameters are not identified easily from the type of experiment that was performed with [$^{18}$F]allypride to estimate receptor binding in regions of moderate binding. Luckily, identification of the blood flow parameters is not the goal of the experiment. $K_1$ and $k_2$ are much better and more easily identified by an experiment involving a very sharp injection of tracer and rapid blood sampling to catch the fine detail of the input function.

---

**Fig. 7** Sensitivity curves for all six parameters of the model depicted in Fig. 2 evaluated over time and at the optimal parameter vector resulting from the fit to data shown in Fig. 5. Each curve is on its own scale. The vertical dotted line in D, E, and F corresponds to the time of the second injection. Everything on those plots to the left of the dotted line is the equivalent of sensitivity curves for a single injection study. See text for details.
More interesting from the standpoint of potential receptor-ligand characterization are the curves in Fig. 7D–F. Figure 7D–F shows the time-varying derivatives of the PET with respect to \( k_{on} \), \( B_{max} \), and \( k_{off} \), respectively. If we consider just the first epoch in each curve (to the left of the dotted line), it is very hard to distinguish the role that is played by any of these parameters. Certainly, there would be no difference in effect between raising \( B_{max} \) or raising \( k_{on} \) during the first epoch. Recall that the first epoch is merely a single-injection experiment (with a high-SA tracer), and it is well known that \( k_{on} \) and \( B_{max} \) are not identifiable from such a limited experiment. The effect of lowering \( k_{off} \) during this period would also be hard to differentiate from a concomitant rise in either of the other two parameters. As mentioned earlier in such cases, modelers must fall back to an identifiable parameter and not try to estimate both the “\( m \)” and the “\( n \)” as discussed in Section II.

If we look over the entire study duration at the sensitivity curves for \( k_{on} \), \( B_{max} \), and \( k_{off} \), we can begin to appreciate that (1) they are each distinguishable from each other and (2) it takes a sufficiently complicated experiment that manipulates occupancy to draw out differences in the processes represented by the three separate parameters. Recall that the second and third injections were termed “partial saturation” and “saturation” (see Table I). In fact, Christian et al. (2004) observed that there is apparently a narrow range of partial saturations that, if achieved during the second phase of the M-I experiment, yield TACs for \([^{18}\text{F}]\text{fallypride} \) experiments, which produce estimates of \( k_{on} \) and \( B_{max} \) that are uncoupled. If the target level of occupancy is under- or overshot in these experiments, interestingly, the parameters remain correlated in the fitting. The potential success of M-I experiments has been explained previously in terms of the sensitivity coefficients (Morris et al., 1999).

B. Using the Sensitivity Information for Design

How can we use this information that appears to be contained in the sensitivity curves objectively? This is the subject of what is known as sensitivity analysis and optimal experiment design. As learned in Section VI.B, the sensitivity matrix can be used to approximate the variances of each parameter estimate. Many scalar quantities can be derived from this matrix and used to compare different experimental designs. A classical index for optimization of an experiment is the \( D \)-optimal criterion. “\( D \)” refers to the determinant of the Hessian matrix \( (H \approx S^TWS) \) or, equivalently, to the determinant of the inverse of the covariance matrix. In either case, to achieve a \( D \)-optimal design, we seek to maximize the value of the determinant of the matrix. The matrix, in turn, contains information about the collective variances of the parameters. In a physical sense, the confidence region surrounding the optimal choice of parameters in parameter space is a \( n_p \)-dimensional ellipsoid (where \( n_p \) is number of parameters) whose axes are the eigenvectors of \( H \). Maximizing the determinant of the Hessian matrix is equivalent to minimizing the volume of this confidence region and thus reducing the possible choices of the parameter vectors that yield an equally good fit to data. That is,
maximizing $\text{det}(H)$ is equivalent to minimizing overall variance of the parameters. Many other quantities can be derived from the Hessian matrix and used as design criteria to maximize or minimize some other aspect of a parameter or parameters (for examples related to PET experiments, see Muzic et al., 1994, 1996, 2000).

Because the Hessian matrix is a function of both the parameters and the experimental protocol, there are two points to consider: (1) optimal design of experiments is iterative—it must be repeated as more becomes known about parameter values upon which Hessian-based criteria depend and (2) the variances of the parameter estimates can be improved by the best choice of protocols, that is, by optimizing over a set of design variables.

1. Design Variables

What are the design variables in the typical M-I PET experiment? There are two design variables. First, the specific activities of the respective injections can be varied by mixing differing amounts of labeled and unlabeled ligand for each injection. Second, the time between injections can be varied. Because the specific activity (or equivalently the mass for a given radioactivity dose) will determine the occupancy level of receptors at a given time, we can appreciate that specific activity is the experimenter’s tool for manipulating the receptor-ligand system to achieve decreased parameter correlation and increased parameter precision. In some circumstances, it may be necessary to put constraints on the design. For instance, if the total time of the experiment must be limited for reasons of convenience or safety, this will act as a constraint on the combination of times between injections. If the synthesis of the radiopharmaceutical is very difficult, it may be practically necessary to limit the design of the M-I study to a one synthesis. If so, then we constrain the choice of specific activities. In particular, the second and third injections will be limited to lower SA than the high-SA material available for the first. To investigate more about this technique, the reader is directed elsewhere for uses of optimal design in PET and tracer kinetics (Bard, 1974; Beck and Arnold, 1977; Carson et al., 1983; Christian et al., 2004; Delforge et al., 1989, 1991; Feng et al., 1999; Jacquez, 1988; Morris et al., 1991; Muzic, et al., 1996, 2000).

VIII. Conclusion

M-I PET studies are labor-intensive undertakings that demand not only experimental acumen but also a synthesis of mathematical and numerical expertise as well. Despite the overhead associated with such experiments, they may be the only practical means of extracting certain kinetic information about tracer uptake and behavior \textit{in vivo} from dynamic PET images. To the extent that it may be helpful and illuminating to determine precisely the values of all the \textit{in vivo} kinetic parameters of a tracer, it is hoped that this chapter served part as an introductory review, part as a tutorial, and part as an operating guide to a useful technique that merges functional imaging with tracer kinetics and optimal experiment design.
References


