ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 38 (2007) 82–94

Balancing bias, reliability, noise properties and the need for parametric maps in quantitative ligand PET: [^{11}C]diprenorphine test–retest data

Alexander Hammers,[a,b,1,2] Marie-Claude Asselin,[c,d,2] Federico E. Turkheimer,[a] Rainer Hinz,[c,d] Safiye Osman,[c] Gary Hotton,[a] David J. Brooks,[a,c] John S. Duncan,[b] and Matthias J. Koepp[a,b,*]

[a]*MRC Clinical Sciences Centre and Division of Neuroscience, Faculty of Medicine, Imperial College, Hammersmith Hospital, DuCane Road, London, W12 0NN, UK*
[b]*Department of Clinical and Experimental Epilepsy, Institute of Neurology, UCL, Queen Square, London, and National Society for Epilepsy MRI Unit, Chalfont St Peter, UK*
[c]*Hammersmith Imanet, GE Healthcare, London, UK*
[d]*The University of Manchester Wolfson Molecular Imaging Centre, Manchester, UK*

Received 31 December 2006; revised 22 May 2007; accepted 18 June 2007
Available online 24 July 2007

[^{11}C]diprenorphine (DPN) is a non-subtype selective opioid receptor PET ligand with slow kinetics and no region devoid of specific binding. Parametric maps are desirable but have to overcome high noise at the voxel level. We obtained parameter values, parametric map image quality, test–retest reproducibility and reliability (using intraclass correlation coefficients (ICCs)) for conventional spectral analysis and a derived method (rank shaping), compared them with values obtained through sampling of volumes of interest (VOIs) on the dynamic data sets and tested whether smaller amounts of radioactivity injected maintained reliability. Ten subjects were injected twice with either ~185 MBq or ~135 MBq of [^{11}C]DPN, followed by dynamic PET for 90 min. Data were movement corrected with a frame-to-frame co-registration method. Arterial plasma input functions corrected for radiolabelled metabolites were created. There was no overall effect of movement correction except for one subject with substantial movement whose test–retest differences decreased by ~50%. Actual parametric values depended heavily on the cutoff for slow frequencies (between 0.0008 s^{-1} and 0.00063 s^{-1}). Image quality was satisfactory for restricted base ranges when using conventional spectral analysis. The rank shaping method allowed maximising of this range but had similar bias. VOI-based methods had the widest dynamic range between regions. Average percentage test–retest differences were smallest for the parametric maps with restricted base ranges; similarly ICCs were highest for these (up to 0.86) but unacceptably low for VOI-derived VD estimates at the low doses of injected radioactivity (0.24/0.04). Our data can inform the choice of methodology for a given biological problem.

Keywords: Opioid receptors; Positron emission tomography; Diprenorphine; Reproducibility; Spectral analysis

Introduction

Ligand PET studies are performed either as one single study per subject, used for comparisons of groups or single subjects against a control group, or as paired studies where the same subject is studied longitudinally or under two or more conditions. Transformation of the time course of regional tissue radioactivity measured by the PET scanner into physiologically or biochemically relevant parameters requires the accurate absolute quantitation of image counts, the acquisition of an input function, often measured by arterial sampling, the application of a number of corrections for nuisance factors affecting images and blood data and the use of a mathematical model to describe the kinetics of tracer uptake in relation to the physiological or biochemical process of interest. Absolute quantification is usually a prerequisite for PET in research applications (Cunningham et al., 2004).

The issue of how to assess the clinical validity of a specific PET method has been debated in the past (Costa, 2002; Hallett, 2004; Schmidt, 2002). While evaluation of sensitivity and specificity for a given condition requires sampling of a clinical population, a necessary prior step is the acquisition of test–retest data in controls

* Corresponding author. Department of Clinical and Experimental Epilepsy, Institute of Neurology, 33 Queen Square, London WC1N 3BG, UK. Fax: +44 1494 875 666.
E-mail addresses: alexander.hammers@csc.mrc.ac.uk (A. Hammers), mkoepp@ion.ucl.ac.uk (M.J. Koepp).
[1] MRC Clinical Sciences Centre and Division of Neuroscience, Imperial College, DuCane Road, London W12 0NN, UK. Fax: +44 20 8383 1783.
[2] These authors equally contributed to the work.
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2007.06.035

for the assessment of reliability. Test–retest studies allow for an estimation of reproducibility, i.e. within-subject differences. Reproducibility, however, represents only one aspect of a measurement (Laruelle, 1999). A method can conceivably be very reproducible at the expense of not reflecting parameters of interest. A more accurate assessment of the value of a method is therefore given by its reliability, estimated from the intraclass correlation coefficient (ICC) for repeated test–retest experiments:

$$\mathrm{ICC} = \frac{\mathrm{MS(bs)} - \mathrm{MS(ws)}}{\mathrm{MS(bs)} + \mathrm{MS(ws)}} \quad (1)$$

where MS = mean sum of squares, bs = between subjects and ws = within subjects.

Intra- and intersubject differences may be desirable when they are due to true biological variability, but are undesirable when caused by nuisance variables. For test–retest data acquired under similar conditions, true intrasubject differences will be small, and the strategy yielding the highest ICC will be preferable (Laruelle, 1999).

Traditionally, mathematical modelling of PET dynamic data has adopted compartmental models (Mintun et al., 1984) in which exchanges between hypothetical bodily compartments (e.g. arterial plasma, free and non-specific binding and specifically bound tracer) are described by systems of first-order differential equations with constant coefficients (reviewed e.g. by Slifstein and Laruelle, 2001). The volume of distribution (VD) of the target tissue is frequently the parameter of interest for PET studies using reversibly bound radiotracers as it is a linear function of the available receptor concentration (Logan et al., 1990) and, as a macroparameter, its estimate is relatively insensitive to noise compared to individual rate constants (Jones et al., 1994).

Parameters can be obtained either for a volume of interest (VOI) or alternatively for each voxel in the image, yielding parametric (functional) maps or images (Blomqvist, 1984). VOI-derived time–activity measurements contain less noise than those derived from single voxels. Successful creation of parametric maps, however, allows to harness the power of functional imaging by enabling very fast visual surveying of entire imaging data sets without having to rely on prespecified VOIs and allows the use of voxel-based image analysis methods for the statistical comparison of groups of subjects and single subjects versus controls (Koepp et al., 1996).

Spectral analysis (Cunningham and Jones, 1993) is a very versatile method for the derivation of VDs from VOI or voxel data of reversibly bound radiotracers. Like in compartmental modelling, spectral analysis describes the tissue time course by a sum of decaying real exponentials convolved with an input function, but the non-linear estimation of the exponential functions is replaced by the linear estimation of the coefficients of a predefined set of biologically plausible exponential basis functions under the non-negativity constraint. Mathematically, VD corresponds to the area under the curve of the impulse response function (IRF) extrapolated to infinity. Spectral analysis does not rely on a prior decision on the most appropriate model structure, and it returns the number of tissue compartments distinguished from the data (Gunn et al., 2002). The fastest bases correspond to the passage of radiotracer in and out of tissue within the time of a short time frame. For data that has not been corrected for radioisotope decay, the slowest bases theoretically possible correspond to the half-life of the radioisotope used; this corresponds biologically to irreversible trapping. For compounds which are expected to be reversibly bound, the cutoff frequency of the bases is set faster than the half-life of the radioisotope to reduce noise. This will lead to an underestimation of VD, since the slowest components of the IRF are not allowed to participate in the definition of VD and will reduce contrast in the data.

Other approaches to addressing noise concerns have been proposed. For example, rank shaping (RS) regularisation uses singular value decomposition of the exponential bases to define an appropriate regularisation over an unconstrained least squares solution, with shrinkage parameters conditioned on the expected signal-to-noise ratio (Turkheimer et al., 2003); it is expected that the resulting maps will have Gaussian characteristics which is desirable if ultimately use of statistical parametric mapping techniques is intended. Another alternative approach for reducing the number of possible bases and equally expanding the use of spectral analysis to reference region models by allowing negative coefficients is the sparse selection of bases by basis pursuit denoising (DEPICT; Gunn et al., 2002).

[^{11}C]diprenorphine (DPN) is a non-subtype selective antagonist at opioid receptors with similar affinities for μ-, κ- and δ-subtypes, used mainly in pain, neurodegeneration, epilepsy and addiction research (Hammers and Lingford-Hughes, 2006). Relative to the duration of the PET study, [^{11}C]DPN has slow kinetics. For such radiotracers, there have been attempts to reduce the dependency on the large area under the curve resulting from interpolation to infinity the by the use of values of the IRF at predefined times, e.g. 60 min in the case of [^{11}C]DPN (Cunningham and Jones, 1993; Koepp et al., 1998a). Only limited [^{11}C]DPN test–retest data have so far been available for one control subject and one patient with rheumatoid arthritis (Jones et al., 1994). In addition, the influence of the choice of cutoff frequency for the exponential bases on parametric map characteristics has never been formally explored for this tracer.

Signal-to-noise ratios (SNRs) depend on many factors including the dose of radiotracer injected. The upper limit for the amount of radiotracer injected is on the one hand given by the tracer kinetics assumption, i.e. the mass of co-injected unlabelled radiotracer should lead to a receptor occupancy of ideally under 1% (Hume et al., 1998), but increasingly so by radiation protection considerations. We therefore investigated whether quantification methods judged optimal for data acquired after injection of ~185 MBq of [^{11}C]DPN still led to acceptable test–retest results after injection of the reduced dose of ~135 MBq of [^{11}C]DPN.

Materials and methods

Subjects

Complete data sets from a total of ten healthy control subjects (Table 1) were available. They were scanned twice as control subjects as part of various clinical research studies over 6 years. Inclusion criteria were age 35–65 years and good health. Exclusion criteria included neurological or psychiatric disease, regular medication, positive pregnancy test and intake of any opiate containing medication in the last 2 weeks. Ethical permission and permission by ARSAC (radiation protection agency) were obtained, and all subjects gave informed written consent. All underwent paired 95-min dynamic [^{11}C]DPN PET baseline (test–retest) scans on the same camera (see PET data acquisition). Five