TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Jarkko Johansson

Spatiotemporal Power of Positron Emission Tomography

Pushing the Limits of Poisson Statistics in High-Resolution Human Neurotransmission Studies

Julkaisu 1401 • Publication 1401

Tampere 2016

Tampereen teknillinen yliopisto. Julkaisu 1401
Tampere University of Technology. Publication 1401

Jarkko Johansson

Spatiotemporal Power of Positron Emission Tomography

Pushing the Limits of Poisson Statistics in High-Resolution Human Neurotransmission Studies

Thesis for the degree of Doctor of Philosophy to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 19th of August 2016, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2016

Supervisor:

Professor Ulla Ruotsalainen
Department of Signal Processing
Tampere University of Technology, Finland

Instructor:

Professor Juha Rinne
Turku PET Centre
University of Turku, Finland

Pre-examiners

Doctor Claude Comtat
Biomedical Imaging Institute
Alternative Energies and Atomic Energy Commission (CEA), France

Associate Professor Simon Cervenka
Department of Clinical Neuroscience
Karolinska Institutet, Sweden

Opponent

Professor Andrew Reader
Biomedical Engineering Department
King's College London, United Kingdom

Faculty of Computing and Electrical Engineering
Tampere University of Technology, Finland

ISBN 978-952-15-3778-3 (printed)
ISBN 978-952-15-3787-5 (PDF)
ISSN 1459-2045

To Elina and the Boys

Abstract

Brain disorders involving dysfunctions in neurotransmission constitute one of the most prevalent health problems. Subtle disruptions in human neurotransmission can result in significant dysfunction of cognition, locomotion, or practically any facet of human behaviour. In turn, homeostasis of a specific neurotransmitter system can often be retrieved through pharmacological or lifestyle interventions. At present, human neurotransmission can be best assayed using positron emission tomography (PET). To date, neurotransmitter-PET (nt-PET) has been employed to investigate neuroreceptor level phenomenon in human behavior/cognition as well as in treatment development. In the current work the goal was to explore and enhance the temporal capabilities of nt-PET, to allow better characterization of the temporal facets of neurotransmission.

Main obstacles limiting temporal characterization stem from the poor signal-to-noise-ratio of the PET measurement. In particular, the limitations related to image reconstruction algorithms and in turn the benefits obtained through regional analysis were in the focus of the investigations in this work. The main finding was that the best temporal resolution achieved using a commonly recommended iterative reconstruction method was insufficient for temporal characterization, while a newly developed algorithm allowing analytical reconstruction showed better temporal resolution without decreasing signal-to-noise-ratio. Furthermore, a novel atlas-based regional analysis method was found superior to the currently employed manual region-of-interest definition.

The findings made through this work will directly assist the planning of future neurotransmission studies, and it is wished that the observations in this work would spark new, more widespread interest on the application of nt-PET in e.g. cognitive stimulation studies.

Preface

This work was carried out in the Turku PET Centre at University of Turku and in the Department of Signal Processing at Tampere University of Technology.

I would like to express my sincere gratitude to my supervisor Professor Ulla Ruotsalainen for giving me the opportunity to work in the Methods and Models for Biological Signals (M^2oBSI) research group, and for giving me all the valuable guidance, in particular in the last phases of the work. Ulla's positive attitude and vast experience in science have been indispensable during this journey. Next, I want to thank my instructor Professor Juha Rinne for trusting the methods in a very ambitious memory training project in my hands, and for the patient guidance to avoid the pitfalls on the way. Juha's great scientific touch and remarkable knowledge in the field of neuroscience have been one of the driving forces during this work. I also sincerely thank the pre-examiners of my thesis, Doctor Claude Comtat and Associate Professor Simon Cervenka, for the careful assessment of my work and for the valuable comments.

I want to express my warmest gratitude to all my co-authors. I owe my deepest gratitude to Dr. Kati Alakurtti who has endured the long trip together with me. Without your positive example and encouragement I would have hesitated taking this challenge. It has been a pleasure working and learning together with you. I want to thank Vesa Oikonen for introducing me to the world of pharmacokinetic modelling, but also to the many facets of automated PET analysis. I thank Professor Mika Teräs, Tuula Tolvanen, and Dr. Virva Saunavaara, not only for introducing me to the basics and practicalities of PET physics, but also for preparation of the phantoms. I want to express my special thanks to Dr. Uygar Tuna and Assoc. Prof. Jussi Tohka from the M^2oBSI-team, it has been a pleasure working with such enthusiastic experts. My warmest thanks go also to

Dr. Terhi Tuokkola and Assoc. Prof. Juho Joutsa for lending their remarkable knowledge in the neurological and radiological fields. Jarmo Teuho and Jani Linden are warmly thanked for their efforts with the highly demanding technological and mathematical issues, and Dr. Kjell Någren for the preparation of raclopride. And last but not least, to Professor Matti Laine I owe my deepest gratitude for not only giving me the opportunity to work with a highly ambitious project but also for supporting me, both financially and scientifically, through this endeavour.

I would also like to thank all the great, fellow researchers and neuroscientists. In particular, Jouni Tuisku deserves my admiration for his neverending interest and unselfish attitude; without your contribution this work wouldn't exist in the present form. My cubicle mates in the "neurologists area": Eero Rissanen, Anna Bruck, Nina Kemppainen, Laura Ekblad, Nora Lindgren, Joonas Majuri and Marcus Sucksdorff; you are all appreciated for your enthusiasm towards neuroscience and for creating the inspiring scientific atmosphere at the Turku PET Centre. Anna Soveri and Otto Waris are acknowledged from outside the "box"; I owe my gratitude for patiently enduring my long and meandering accounts about PET methodology, and in return I have learned a great deal about psychology metrics from you. Professors Lars Nyberg and Lars Bäckman are warmly thanked for the opportunity to work in the memory training project, and for giving both financial support and superb scientific perspective. Professors Mika Scheinin and Harry Scheinin and Assoc. Prof. Valtteri Kaasinen are also warmly thanked for being such great scientists and still available for interesting discussions. I owe my gratitude to Professor Hidehiro Iida for providing the "Iida Phantom". The "PET boys and girls" group is warmly thanked for the inspiring conversations and cheerful moments outside the office: Henry Karlsson, Lauri Tuominen, Lauri Nummenmaa, Jussi Hirvonen, Jetro Tuulari, Henri Honka, Pekka

Jokinen, Joonas Eskelinen and Tiina Saanijoki; I'm grateful for your friendship. Pauliina Luoto, Heli Louhi, Gaber Komar, Tommi Kokki, and Henri Sipilä, you have become my friends at the PET Centre; I want to thank you for your friendship and support during these years.

I am priviledged to have been part of the Turku PET Centre personnel during these year. I express my sincere thanks to Professor Juhani Knuuti, Director of Turku PET Centre, for allowing me to use the facilities during this study. I thank Marko Tättäläinen and Rami Mikkola for their expert help with IT, and Hannu Sipilä for his craftsmanship. Mirja Jyrkinen, Sinikka Lehtola and Auli Kärpijoki are acknowledged for secretarial matters, and Elina Orava for coordinating me over the hurdles in the PhD process. I wish to express my thanks to radiographers Tarja Keskitalo, Marjo Tähti, Hannele Lehtinen, Anne-Mari Jokinen, and Minna Aatsinki, and to laboratory technologists Sanna Suominen and Heidi Lappalainen, for being there when needed.

I express my thanks to my financial supporters: the Instrumentarium Research Foundation, the Research Council of Southwestern Finland Hospital District, and Professors Ulla Ruotsalainen, Juha Rinne, Matti Laine, and Lars Bäckman through various research grants.

Last but not least, I thank all the close people for the support and understanding during this project. I owe my deepest gratitude to my parents Sanni and Veikko for their love and never-ending belief in my skills. My sister Pia and her family, you have shown me what is important in life. I also wish to thank my parents-in-law, Eija and Antti, for warmly accepting me as part of your family. I thank my sisters-in-law, Soila, Essi and Iida and brother-in-law Teemu, for all the laughs, and for letting me forget about my thesis project. Elina, my beloved spouse, I thank you for being in my life and for giving us our

amazing and talented sons Lenni, Linus and "?". I want to thank you for all the cherished moments together during these busy years; without your support and positive attitude towards this project it would simply not have happened.

Turku, May 2016 Jarkko Johansson

Contents

Abstract iii
Preface iv
Abbreviations xi
List of Publications xv
Overview of Publications xvi

1 Introduction **1**

2 Background **9**
2.1 PET receptor binding assay 9
2.1.1 Tissue-ratio analysis 12
2.1.2 Compartmental modelling 14
2.1.3 Simplified reference tissue model 16
2.1.4 Extended compartmental modelling . . . 17
2.2 Principles of positron emission tomography . . . 23
2.2.1 Data acquisition 25
2.2.2 Image signal-to-noise ratio 27
2.2.3 Image reconstruction 28
2.2.4 High resolution PET 35
2.2.5 Image reconstruction strategies for the HRRT 35
2.3 Region of interest analysis 40
2.3.1 Implications from the functional organization of the striatum 40
2.3.2 ROI delineation strategies in the striatum 43
2.4 Summary 46

3 Objectives **49**

4 Experiments **51**

5 Results and discussion **61**

6 Implications **77**

Bibliography **81**

Publications **103**

Abbreviations

2D two-dimensional

2DFT two-dimensional Fourier transform

3D three-dimensional

3DRP three-dimensional reprojection FBP

AUC area-under the curve

ASTR associative striatum

BP binding potential

Bq becquerel (radioactive decays per second)

CAU caudate nucleus

CB connectivity based

CERC cerebellar cortex

CFS constrained Fourier space

CT computed tomography X-ray computed tomography

DA dopamine

DCT discrete cosine transform

DIFT direct inverse fourier transform

DLPFC dorsolateral prefrontal cortex

DOI depth of interaction

DV distribution volume

ESRTM extended simplified reference tissue model

FBP filtered backprojection

fMRI functional magnetic resonance imaging

FWHM full width at half maximum

FOV field of view

FOM figure-of-merit

GABA gamma-aminobutyric acid

HRRT High Resolution Research Tomograph

ICC intraclass-correlation coefficient

keV kilo electron volt

LOA limits of agreement

LOR line of response

lp-ntPET linear parametric neurotransmitter-PET

LSTR limbic striatum

LSRRM linearized simplified reference region model

ML maximum likelihood

MLEM maximum likelihood expectation maximization

MNI Montreal Neurological Institute

MRF Markov random field

MRI magnetic resonance imaging

MRP median root prior

MT motion tracking

NAcc nucleus accumbens

NEC noise equivalent counts

NEMA National Electrical Manufacturers Association

nt-PET neurotransmitter-PET

OMPFC orbital and medial prefrontal cortex

OP-OSEM ordinary Poisson ordered subsets expectation maximization

OSEM ordered subsets expectation maximization

PET positron emission tomography

PSF point spread function

PUT putamen

RM resolution modeling

RM-OP-OSEM resolution modeling OP-OSEM

ROI region of interest

SD standard deviation

SEM standard error of measurement

SMST sensorimotor striatum

SNR signal-to-noise ratio

SPECT single-photon emission tomography

SPM Statistical Parametric Mapping

SSRI selective serotonin reuptake inhibition

SRTM simplified reference tissue model

STR striatum (whole)

TAC time activity curve

TRV test-retest variability

VTA ventral tegmental area

List of Publications

I Johansson, J., Oikonen, V., Teräs, M., "Quantitative brain imaging using the new, fast iterative histogram-mode reconstruction for the HRRT PET scanner," *Nuclear Science Symposium Conference Record.*, 2007, NSS '07. IEEE.

II Alakurtti, K., Aalto, S., Johansson, J., Någren, K., Tuokkola, T., Oikonen, V., Laine, M., Rinne, J. O., "Reproducibility of striatal and thalamic dopamine D2 receptor binding using [^{11}C]raclopride with high-resolution positron emission tomography.," *J Cereb Blood Flow Metab.*, 2011, 31, 1, 155-165.

III Johansson, J., Teuho, J., Linden, J., Tuna, U., Tolvanen, T., Saunavaara, V., Teräs, M., "Image quantification in high-resolution PET assessed with a new anthropomorphic brain phantom," *Nuclear Science Symposium Conference Record.*, 2013, NSS '13. IEEE.

IV Tuna, U., Johansson, J., Ruotsalainen, U., "Evaluation of analytical reconstruction with a new gap-filling method in comparison to iterative reconstruction in [^{11}C]-raclopride PET studies.", *Ann Nucl Med*, 2014, 28, 5, 417-429.

V Johansson, J., Alakurtti, K., Joutsa, J., Tohka, J., Ruotsalainen, U., Rinne, J.,O., "Comparison of manual and automatic techniques for sub-striatal segmentation in ^{11}C-Raclopride PET studies", *Nuclear Medicine Communications*, 2016, Accepted for publication.

Overview of Publications

This thesis work consists of five publications reporting original research. Throughout the thesis the Publications are referred by the roman numeral in List of Publications. Firstly, Publications I-II report the experiments made with iterative reconstruction and a predefined noise equivalent counts (NEC)-threshold to guarantee the quantitative accuracy in HRRT image reconstruction. Secondly, Publications III-IV report the further experiments made with modified iterative algorithms and using a novel gap-filling approach in tandem with three-dimensional reprojection FBP (3DRP) algorithm for analytical image reconstruction. And finally, Publication V reports the results obtained in an optimization study to enhance the methodology for striatal and cerebellar region of interest (ROI)-delineation in high-resolution dopamine studies.

In Publications I, III and V the author of this thesis was the first author and made the main contributions. In these publications the author of this thesis had the main responsibility of designing the experiments, implementing the new methods, analysing the results, and drafting the reports. While in Publication II the author of this thesis was responsible in designing the experiments related to reconstruction methods and writing the article in that part, assisted in all the analysis and reviewed the article. Thus, in Publication II the role of the author of this thesis was central in the parts involved with image reconstruction, highlighting the distinct roles of the authors in this particular study. In Publication IV the author of this thesis participated in the design of the experiments directed for assessing the reconstruction algorithms with real human data, assisted in the analysis and in the drafting of the report. The role of the author of this thesis was pivotal in the design phase when inclusion of the human data was discussed and decided.

1 Introduction

At present, positron emission tomography (PET) offers the best approach to assess human neuroreceptor binding in vivo. PET can be employed in the assessment of brain blood-flow and metabolism as well, but the current work focuses on the employment of neuroreceptor-binding radioligands, referred to as neurotransmitter-PET (nt-PET) throughout this work. In nt-PET a neuroreceptor binding assay can be employed in the investigation of e.g. pathophysiology of a brain disorder involving perturbations in specific neurotransmitter system.

It is hypothesized that relatively subtle alterations in human neurotransmission can result in significant dysfunction of cognition, locomotion, or practically any facet of human behaviour.

The human brain contains approximately 86 billion neurons [6] each connected through 1000-10000 synapses creating a huge neuronal network. Synapse is the connection point between neurons typically encompassing the axon terminal (pre-synaptic terminus), a small gap (20-40 nanometers across), and the adjacent axon or dendrite membrane (post-synaptic terminus) with appropriate receptor molecules. Neuronal signal travels from the axon hillock to the axon terminal electrically but synaptical signal transfer is achieved through chemical neurotransmission.

Arrival of the electrical signal (action potential) to the pre-synaptic terminus results in diffusion of neurotransmitter molecules into synaptic cleft and subsequently binding with the appropriate post-synaptic receptor molecule. Neurotransmitter binding

activates the receptor molecules in the post-synaptic membrane resulting in excitatory or inhibitory effect depending on receptor type and neurotransmitter. Excitatory couplings cause depolarization of the post-synaptic membrane resulting in action potential propagation, while the inhibitory couplings cause hyperpolarization and action potential is prevented from propagation.

There are currently over hundred neurotransmitters identified of which glutamate is the most prevalent followed by gamma-aminobutyric acid (GABA). Approximately 90% of neurons are glutamatergic, and 90% of those that are not glutamatergic are GABAergic. Glutamate is the major excitatory and GABA the main inhibitory neurotransmitter in the brain. Albeit the scarcity of the other neurotransmitters such as acetylcholine, dopamine, serotonin, norepinephrine and epinephrine, they have important roles in the normal brain function. Neurons expressing the rare neurotransmitters have been shown to form distinct neurotransmitter systems that regulate various important psychological and behavioural processes.

For instance dopaminergic neurons in the substantia nigra and their terminals in the striatum form the nigrostriatal pathway that has a major role in motor control, while dopaminergic neurons in the ventral tegmental area (VTA) that terminate in the ventral striatum form the mesolimbic pathway that is part of the limbic system and has a particularly important role in the human reward system. While the serotonergic neurons in the raphe nuclei and their terminals almost throughout the brain form the serotonergic pathways that have been associated with e.g. mood disorders.

Albeit neurotransmitter systems work together, and as a whole the brain circuits are heavily inter-connected it is often that brain disorders are associated with specific neurotransmitter systems. For instance Parkinson's disease has been convinc-

ingly connected with the cell death of dopaminergic neurons in the substantia nigra thus explaining the movement-related symptoms. Symptoms of Parkinson's disease can subsequently be treated using exogenous dopamine precursors (levodopa), dopamine agonists, or dopamine reuptake pathway inhibitors that aim to counteract the deficiency of dopamine induced by the loss of dopaminergic cells in the subtantia nigra.

Schizophrenia on the other hand has been associated with hyperactive dopaminergic signalling in the mesolimbic pathway [62] and its symptoms have thus been treated with dopamine antagonists.

While, depression, identified as a risk to public health by the Finnish Medical Society DUODECIM in their Current Care Guidelines, has been suggested to be caused by down-regulated serotonergic signalling and the first-line treatment recommendations include medication with antidepressant drugs, often acting via selective serotonin reuptake inhibition (SSRI).

Thus, several brain disorders are associated with disorders in specific neurotransmitter systems, and, consequently the molecular pathways are readily available for targeting treatments using e.g. psychoactive drugs. At present, nt-PET holds the greatest potential to provide insight into changes related to specific disorders and subsequently to pharmacological response to drugs.

Formulation of the concept of "receptive substances" to explain the pharmacological response to drugs dates back to the end of the 19^{th} and beginning of the 20^{th} century when Ramon y Cajal, Otto Loewi and others unveiled first the structure of a neuron and later the chemical neurotransmission. Receptor occupancy theory of drug action was coined by John Langley and Paul Ehrlich in the beginning of the 20^{th} century. The receptor occupancy theory aims to explain the pharmacological response via the proportion of receptors occupied directly or

indirectly by the drug, and it is at the moment used to describe the pharmacological response of a majority of psychoactive drugs.

For the first decades the receptor theory was employed to explain the drug actions without accurate knowledge of receptors. It was simply assumed that the neuron endings contain some molecular structures that recognize and bind specific chemical compounds named ligands (neurotransmitters or drugs) that subsequently elicit the effects in the receptive synapse.

Now the most abundant receptors are well characterized with regard to their molecular structure, binding mechanisms, and subsequent actions in the receptive cell membrane. Nevertheless, the true *in vivo* pharmacokinetics and further pharmacodynamics of a drug are best assessed using receptor binding assays with e.g. nt-PET.

Receptor binding has been extensively studied with *in vitro* and *in vivo* experiments using laboratory animals and radioactive ligands (radioligands) [68]. In the animal experiments radioligands have been most often radiolabeled using ^{3}H (tritium), a β-emitter. The low energy of β-radiation is however rapidly absorbed by the tissue, thus permitting its localization by external detectors. The experiments using tritiated ligands thus require killing the animals and assaying the radioactivity distribution using cryonic tissue sections and autoradiography.

The true *in vivo* radioligand assay can be achieved using PET. Short-living positron-emitters such as ^{11}C, ^{18}F and ^{15}O have been used to label radioligands that can be externally detected via annihilation energy. The annihilation energy arises from the positron-electron collision resulting in generation of two opposite-directional γ-rays.

A PET scanner is used in the detection of γ-particles using opposing detectors and simultaneous photons are thought to originate from the same annihilation event, thus allowing

localization of the event to within a hypothetical line of response (LOR). Acquisition of a large number of coincidence events within numerous LORs enable tomographic image formation via image reconstruction similar to X-ray computed tomography (CT) [52]. The true *in vivo* assay of radioligand distribution yields sequential measurement of the time-course of the ligand in living tissue, thus depicting the pharmacokinetics of the ligand.

Owing to the mild invasiveness and high number of potential radioligands PET has become the gold-standard in human *in vivo* receptor binding assays. The high specific activity of PET radioligands allow their usage in tracer doses, implying that they do not interfere with the molecular pathway under investigation. While single-photon emission tomography (SPECT) can also measure molecular pathways through radiotracer technique, it lacks the quantitative accuracy and sensitivity of PET. Furthermore, the quantitative, dynamic assay using PET allows modelling of the radioligand's pharmacokinetics yielding true physiological parameters, in contrast to mere radioactivity measurement.

Essentially dynamic PET receptor binding assay is sensitive to instantaneous alterations in the receptor availability [38, 63, 73]. Thus, in principle the receptor binding assay can reflect rapid alterations in receptor occupancy, induced by changes in exogenous drug or endogenous neurotransmitter concentrations. Neurochemical signaling *per se* is time sensitive, and the neurotransmitter release induced by e.g. prolonged cognitive activation can blunt relatively soon after the task initiation [5], whilst temporal characteristics of the activation response may be more important than the averaged amount of neurotransmitter release. Temporal characteristics in pharmacological stimulation may in addition reflect e.g. pathological state in the receptive site linked with for instance addictive behavior (c.f. [104]).

Considering the apparent potential of temporal characterization of the neuroreceptor response with PET its application has been very limited. It is anticipated that the implications from the stochastic data acquisition process in PET have been the major limiting factor in the implementation of these studies.

In the signal processing parlance counting process of the rare nuclear events in PET is termed a Poisson counting process, and thus the characteristics of Poisson random variables apply to the PET measurement. In particular the signal-to-noise ratio (SNR) of a PET measurement is governed by the count statistics, following from the essential equality between the expected mean and variance of a Poisson variable. The relative count statistics in turn are determined by the radioactivity of the source, the scanning time and the spatial sampling frequency, thus yielding a tradeoff between the spatial and temporal resolution.

Radioactivity of the source can not be largely increased due to tissue ionizing properties of the γ-particles, and thus the tracer doses administered to human subjects are strictly limited by the radiation safety authorities. In PET these limitations translate into relatively small radioligand dosing and thus weak measurable signal, but also to limited number of scannings per subject.

On the other hand, higher spatial sampling has only fairly recently become available in human PET imaging, partly due to the Poisson characteristics of the measurement process. Higher spatial sampling unavoidably results in relatively poorer count statistics and extensive development of scintillator materials and signal processing methodology has taken place before the first human-size high-resolution PET scanner was launched [108]. Despite the advances made, high-resolution PET imaging suffers from inherently poor count statistics to a degree that may compromize the temporal resolution, thus preventing the temporal characterization of the neuroreceptor response, while

the high-spatial sampling may become indispensable in the experiments probing the small, functionally specialized brain areas such as the human striatum.

Albeit at current PET provides the gold standard in *in vivo* receptor binding assay its relatively poor spatial resolution has been regarded as a significant hindrance in the study of small functionally specialized brain areas, while the temporal resolution of PET is governed by the length of the acquisition frame to yield sufficient count statistics. In contrast, functional magnetic resonance imaging (fMRI)-technique which is not dependent on Poisson characteristics can provide simultaneous spatiotemporal resolution, which in turn has allowed the implementation of various, important cognitive activation studies. Current magnetic resonance imaging (MRI)-based techniques however lack the molecular specificity of PET, thus not allowing direct inferences about the neurotransmitter level actions and consequently do not allow e.g. treatment targeting using neuroreceptor specific psychoactive drugs nor support their development.

Aim of the current work was to investigate and improve the prospects of high-resolution PET in the temporal characterization of neuroreceptor binding. At current, the potential of PET binding assay is not fully exploited, albeit a few pioneering studies on the temporal profiling have been made. It is anticipated that the Poisson characteristics of PET is the main obstacle on the way to successfully implement a temporally sensitive receptor binding assay. Thus, in the current work the focus of the investigations was on the particular implications from the poor count statistics of high-resolution PET.

In particular, performance of various existing and newly developed algorithms for high-resolution image reconstruction were evaluated in light of their quantitative accuracy and SNR performance. The role of image reconstruction algorithms is important since they can aggregate the statistical variation in

the measurement or in the worst case generate biased radioactivity estimates leading to erroneus interpretation of receptor binding. On the other hand, a commonly employed strategy in PET to improve the measurement SNR is to average the signal within functionally homogenous image elements. The functional region of interest (ROI)-definition in e.g. the human striatum is however prone to errors and the currently employed manual methodology does not fulfill the requirements in high-resolution PET. Thus, it was deemed necessary in the current work to revise the methodology in the striatal ROI-delineation and seek for automated approach in order to fully benefit from high-resolution PET.

2 Background

2.1 PET receptor binding assay

The PET receptor binding assay exploits radioligands (in tracer doses radiotracer) that can specifically bind to the target receptor of interest, while, the PET scanner can be used to depict the fate of the radiotracer in living human tissue in a timeline, thus depicting the tracer's kinetic profile.

Figure 2.1 illustrates typical time activity curve (TAC)-data from a PET measurement. TAC-data depicts the instantaneous radioactivity concentration (e.g. becquerel (radioactive decays per second) (Bq) per millileter) within a single image volume element (voxel) or within a ROI, representing a mixture of signals from specifically bound, nonspecifically bound and free tracer in brain tissue. Thus the TAC-data does not directly depict the physiological parameter of interest such as the binding potential (BP) (c.f. [71]), but is associated with nuisance component of nondisplaceable binding (free radiotracer in tissue plus nonspecific binding). Nondisplaceable binding can vary between study subjects and groups and its contribution should thus be removed from the measurements.

Calculation of the true physiological parameters can be achieved in various ways often resulting in similar or comparable outcomes. Often the calculation is based on interpretation of the tracer's kinetic profile. Analysis of the tracer's activity profile, i.e. the TAC-data, proceeds by distinguishing the signal

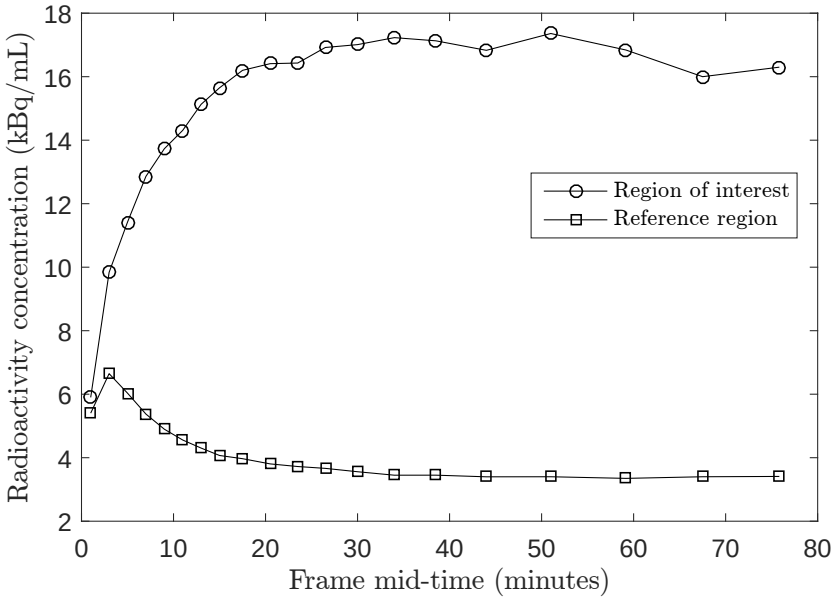


Figure 2.1: Graphs showing the instantaneous activity concentrations within a region of interest (ROI) and reference region as a function of time.

originating from the true receptor binding from those originating from nuisance components, such as tracer that is free or nonspecifically bound in brain tissue.

Often the desired outcome of nt-PET TAC-analysis is the receptor binding potential relative to the free fraction of the tracer in the nondisplaceable compartment (f_{ND}) and affinity ($1/K_D$), thus denoted as $\mathrm{BP_{ND}}$ [48]. $\mathrm{BP_{ND}}$, calculated using any of the methods, is of particular interest since it is thought to reflect the number of available receptors for tracer binding.

The number of available receptors for binding depends on the total density of the receptors and the amount of endogenous neurotransmitter or exogenous drug occupying the receptors. The true receptor occupancy due to intervention is defined as

$$RO_{true} = 100\% \times (1 - \frac{B_{avail}(intervention)}{B_{avail}(baseline)})$$

where $B_{avail}(baseline)$ is the density of available receptors at

baseline or at control condition, while $B_{avail}(intervention)$ the density of available receptors after intervention. The intervention can be pharmacological or non-pharmacological and the action can be direct receptor blocking by a drug or indirect blocking due to increased concentration of endogenous neurotransmitter.

Under ideal conditions assuming that f_{ND} and K_D do not change between the baseline and intervention conditions, the PET derived BP_{ND} can be used to assess the receptor occupancy. The PET derived receptor occupancy due to intervention is defined as

$$RO_{PET} = \Delta BP_{ND} = 100\% \times (1 - \frac{BP_{ND}(intervention)}{BP_{ND}(baseline)}). \tag{2.1}$$

Under the assumptions that the density of receptors (B_{max}), f_{ND} and K_D do not change between conditions, the RO_{PET} would be equal to RO_{true}.

Most commonly the BP_{ND}(intervention/baseline) in Equation 2.1 are assessed in a receptor blocking study, requiring separate PET measurements with multiple bolus injections of the radioligand often taking place on different visits to the imaging facility. Albeit short- and long-term test-retest repeatability of BP_{ND} with e.g. ^{11}C-raclopride are good [1, 3, 42, 67] it is often conceivable that additional biological variability is introduced if the scans are separated in time.

For drug induced alterations, but in particular for those due to behavioural stimulation a more recommendable approach would be using a single-injection displacement study. In the displacement study the control condition is typically initiated prior to the tracer injection and maintained until a reliable estimate of the BP_{ND}(baseline) can be calculated, followed by the intervention which is ideally maintained until reliable estimate of the BP_{ND}(intervention) can be calculated. Limitations due to

radiation safety also encourage employment of the displacement approach.

For a displacement study, binding of the tracer must be displaceable, that is reversible and specific. Not all PET tracers fulfill these requirements but for instance the ^{11}C-labelled raclopride (^{11}C-raclopride for short) a dopamine D_2 and D_3 receptor specific antagonist, has been qualified for this purpose. In [73] and [38] Morris, Fisher and colleagues used simulations to investigate the potential of ^{11}C-raclopride to measure alterations in the striatal dopamine concentration. First they approximated that the average synaptic concentration of dopamine is ≈100 nM and that it would be doubled during e.g. behavioural activation. Through consideration of the concentration and affinity of the radioligand, and that of the endogenous dopamine for dopamine receptors they inferred that it is likely that the changes are detectable using PET. Furthermore, they used mathematical simulation to seek for optimal ligand features and timing of the activation [73]. PET with ^{11}C-raclopride for cognitive activation was successfully employed first by Koepp and colleagues [58], although the findings have been partly challenged in later analysis [32].

Robust and sensitive methodology for the estimation of the BP_{ND}(intervention/baseline) has a crucial role in the successful implementation of the PET receptor binding assay. Thus, considerable effort have been devoted to the methodological development of PET pharmacokinetic modelling within the last few decades.

2.1.1 Tissue-ratio analysis

A straightforward approach to remove the nuisance components from the nt-PET measurement is to extract the TAC representing the nondisplaceable binding from that of total binding, resulting in a TAC of specifically bound tracer. This approach

requires that a reference region devoid of target receptors exists in the brain. In that case the TAC of specifically bound tracer is searched for peak-time and the instantaneous activity of bound tracer is divided by that of nondisplaceable tracer to yield BP_{ND}.

The peak-time of bound TAC was chosen because it is thought to define the transient or pseudo-equilibrium time of the tracer binding [35]. Typically the pseudo-equilibrium time is not easily determined from the noisy bound TAC data and prefixed time intervals around the equilibrium time [78] or late scan times [81] have been employed instead. In comparison to full kinetic modelling the straightforward tissue ratio method has been shown feasible for BP_{ND} calculation [49], with some limitations due to non-steady-state conditions in the tissue and plasma when rapid bolus administration of the radiotracer was employed.

The non-steady-state conditions can be overcome via administration of the radiotracer as a rapid bolus followed by continuous infusion until the true equilibrium is achieved [21]. A steady-state of radiotracer concentration in plasma, in target region and in the reference region can be obtained if the proportion of bolus (k_{bol}) and infusion can be optimized accurately [21], in which case the BP_{ND} can be calculated as bound over nondisplaceable uptake but now within a longer time-span, e.g. 50 minutes.

The bolus-infusion approach with optimized k_{bol} has been successfully employed in e.g. measurement of amphetamine-induced changes in ^{11}C-raclopride binding [22] showing similar outcome from the displacement study and from the more commonly performed blocking study. As a disadvantage, the k_{bol} parameter may show considerable variation between subjects [49] yielding non-steady state conditions in some subjects if global predetermined value is employed.

2.1.2 Compartmental modelling

Advanced analysis of the tracer's activity profile employ tools of pharmacokinetic modelling, in particular the mathematical compartmental model, that do not necessitate equilibrium condition. Figure 2.2 illustrates a compartmental model depicting transformations of a nt-ligand between free in arterial plasma, free in brain tissue and bound in brain tissue.

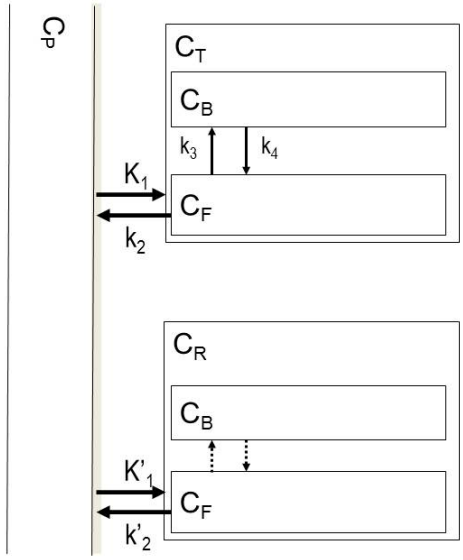


Figure 2.2: Compartmental model of a neurotransmitter ligand. Compartments C_P, C_T, and C_R represent measured radioactivities in arterial plasma (P), target tissue (T) and reference tissue (R), while compartments C_B and C_F represent calculated radioactivies of the free (F) and specifically bound (B) given the transfer constants k. Transfer constants K_1 and K'_1 depict the transfer from arterial plasma to the target and reference tissues respectively, while the k_2 and k'_2 depict the returning transfer. The in- and outbound transfer goes through the free compartment (C_F) from where the radiotracer can bind specifically at rate k_3 or to which it can return at rate k_4.

In a compartmental model the compartments are not physical places but states of the ligand and their unit is radioactivity concentration (e.g. Bq/mL), while the transfers between compartments are described by rate constants (k) that define the

relative transfer from compartment (C) in unit time, 0 reflecting no transfer, and the amount of transfer is kC per unit time.

In compartmental modelling the goal is to determine the values of the rate constants given the measurements, thus numerical implementation of the model is needed. Firstly, the appropriate compartments and transfers between the compartments are defined on the basis of the biochemical properties of the tracer, i.e. the number of compartments and connections between them depend on the specific characteristics of each tracer.

Secondly, the model is solved to provide equations of the instantaneous concentrations in each compartment. Let us denote the TAC of the radioactivity in compartment C as C(t), then using the notation in Figure 2.2 we have the following equations describing the radioactivity concentrations in compartments C_R, C_F and C_B

$$\begin{aligned} C_T(t) &= C_F(t) + C_B(t) \\ dC_F(t)/dt &= K_1 C_P(t) - k_2 C_F(t) - k_3 C_F(t) + k_4 C_B(t) \\ dC_B(t)/dt &= k_3 C_F(t) - k_4 C_B(t) \\ dC_R(t)/dt &= K_1' C_P(t) - k_2' C_R(t) \end{aligned}$$

where C_P is the radioactivity concentration in arterial plasma depicting the input function.

The solution of the compartmental model includes input function which can be externally measured from the arterial plasma samples, unknown concentrations in compartments of free and bound ligand and unknown rate constants. The solved model is used to simulate possible tissue TACs given the input function, and numerical optimization algorithms can subsequently be used to find the rate constants that yield best match with the measured TAC. Thus calculation of the rate constants requires definition of the model, closed form solution of the model and finally numerical optimization of the simulated data to find best fit with the measured data.

In the working example the rate constants k_3 and k_4 are of particular interest. Under certain reasonable assumptions they reflect the reaction rate constants k_{on} and k_{off} in the Michaelis-Menten equation, and we have

$$\begin{aligned} k_3 &= f_{ND} k_{on} B_{avail} \\ k_4 &= k_{off} \end{aligned}$$

where f_{ND} is the fraction of the tracer that is free in nondisplaceable compartment and B_{avail} is the density of receptors available to bind the radioligand [48].

Recall that the nondisplaceble compartment consists of the tracer that is free or nonspecifically bound in the brain tissue. The fraction $K_D = k_{off}/k_{on}$ is the dissociation constant in Michaelis-Menten model. Respectively the ratio of k_3 and k_4 coincides with the formulation of BP_{ND} (c.f. [71] and [48])

$$BP_{ND} = \frac{f_{ND} B_{avail}}{K_D} = \frac{k_3}{k_4}. \tag{2.2}$$

2.1.3 Simplified reference tissue model

The compartmental model presented in Figure 2.2, known as simplified reference tissue model (SRTM)[61], is commonly used to depict the pharmacokinetics of nt-tracers, e.g. that of ^{11}C-raclopride. Initially raclopride-PET (i.e. PET measurements using ^{11}C-raclopride) was modelled using three tissue compartments: free, non-specifically bound and receptor-bound [34], but it was noticed that a steady-state occurs rapidly between free and non-specifically bound tracer and it was hard to distinguish these two compartments in the model fitting, and they were thus pooled together [61]. Furthermore, it has been shown that the arterial input function measurement can be omitted if there exists a brain region devoid of the target receptors [61]. Measurement of the arterial input function is an invasive process

that is prone to errors, and it is avoided when possible. For raclopride a reference region is found in the cerebellum which is almost completely devoid of dopamine receptors according to human post mortem data.

Using the notation in Figure 2.2 we can present the solved SRTM equation in closed form as

$$C_T(t) = R_1 C_R(t) + k_2 \int_0^T C_R(t)dt - k_{2a} \int_0^T C_T(t)dt \qquad (2.3)$$

where $C_R(t)$ and $C_T(t)$ are the instantaneous activity concentrations in the reference and target tissue, respectively, and

$$R_1 = \frac{K_1}{K_1'} \quad \text{and} \quad k_{2a} = \frac{k_2}{1 + \frac{k_3}{k_4}} = \frac{k_2}{1 + BP_{ND}}. \qquad (2.4)$$

In Equation 2.3 $C_T(t)$ is the outcome of simulation while $C_R(t)$ is the measured activity concentration. Similarly as for the full compartmental model the SRTM parameters are estimated via optimization with measured PET data. The BP_{ND} in SRTM formulation corresponds to that obtained through full compartmental model fit with arterial input (c.f. Equation 2.2), but it can be assessed non-invasively using reference tissue.

The optimization can be made to the linear formulation of SRTM in Equation 2.3 and BP_{ND} can be solved from Equation 2.4, or k_{2a} can be substituted in Equation 2.4 when optimization of the non-linear formulation yields the BP_{ND} directly.

2.1.4 Extended compartmental modelling

Dynamic PET measurement lends itself for the measurement of instantaneous alterations in B_{avail} but the compartmental models discussed thus far do not accommodate the temporal changes. Instead, the bolus-infusion administration with tissue-ratio analysis has been almost exclusively employed in

the displacement studies (see [32] for review). In bolus-infusion studies the BP_{ND}(baseline) is typically estimated during initial part of the equilibrium condition followed by BP_{ND}(intervention) estimation after a new equilibrium level has been attained [22]. Considerable effort has taken place with regard to optimization of k_{bol} as well as the optimal timing of the baseline versus intervention. According to Watabe and colleagues [106] the optimal timings for ^{11}C-raclopride would be 0-50 minutes (control task) and 50-100 minutes (activation task), but there might be considerable variation in the optimal timings depending on the tracer and study setup.

A serious limitation of tissue ratio analysis in bolus-infusion studies is posed by the violation of equilibrium condition. As was discussed earlier, employment of predetermined k_{bol} may result in significant under- or over-shoot in the individual's TACs resulting in biased outcome in tissue ratio calculation. Furthermore, pharmacokinetic characteristics of other radiotracers than ^{11}C-raclopride may completely preclude their analysis using equilibrium techniques.

Extensions to SRTM have been proposed that can accommodate the temporal changes in receptor availability. Zhou and colleagues have proposed a method they call extended simplified reference tissue model (ESRTM) [110] that accommodates two BP_{ND} parameters in Eqs. 2.3-2.4 instead of one, one depicting BP_{ND} at baseline and the other at intervention. The advantage of ESRTM over tissue ratio method is that the equilibrium condition is not required. Indeed, they showed in an ^{11}C-raclopride study that the magnitude of amphetamine-induced ΔBP_{ND} from ESRTM was larger and more consistent with amphetamine dosing than that from tissue ratio analysis [110], and consequently its application was proposed instead of tissue ratio analysis.

Alpert and colleagues on the other hand have proposed a

method they call linearized simplified reference region model (LSRRM) [5], that can more flexibly accommodate the transient perturbations in radiotracer binding. In their proposal the time dependence of BP_{ND} is modeled via additive time dependent activation function.

In Equation 2.3 BP_{ND} acts through k_{2a}, thus the change in BP_{ND} should affect k_{2a}. Let us introduce time dependent activation function h(t) and enter the activation effect linearly to the model through $k_{2a}+\gamma h(t)$, where γ is the magnitude of the activation effect while h(t) gets values between 0 and 1. Now Equation 2.3 becomes

$$C_T(t) = R_1 C_R(t) + k_2 \int_0^T C_R(t)dt - k_{2a} \int_0^T C_T(t)dt - \gamma \int_0^T h(t) C_T(t)dt. \quad (2.5)$$

The activation function h(t) can have any shape but it should reflect the alterations in radiotracer's apparent dissociation constant k_{2a}. Figure 2.3 illustrates few possibilities for the activation function.

Equation 2.5 offers a flexible model for the assessment of transient perturbations in BP_{ND}. The approach has been applied in analysis of both pharmacological [5] and non-pharmacological [7–10] stimuli, and recently in generation of dynamic parametric maps of neurotransmission alterations [74] through optimization of the activation function peak-time voxel-by-voxel (c.f. Figure 2.3 for activation function peak-time). Furthermore, it has been suggested [5] that the t-value of the activation parameter γ as obtained through the linear optimization process ($t(\gamma) = \gamma / sd(\gamma)$) could be utilized to detect neurotransmitter release in single subject level, that is, to reject the null-hypothesis of no activation.

In the original model by Alpert and colleagues [5] time-dependent activation functions were inserted for all of the fit parameters: R_1, k_2 and k_{2a}. In their examination it was however found that model identifiability was limited to only one time-

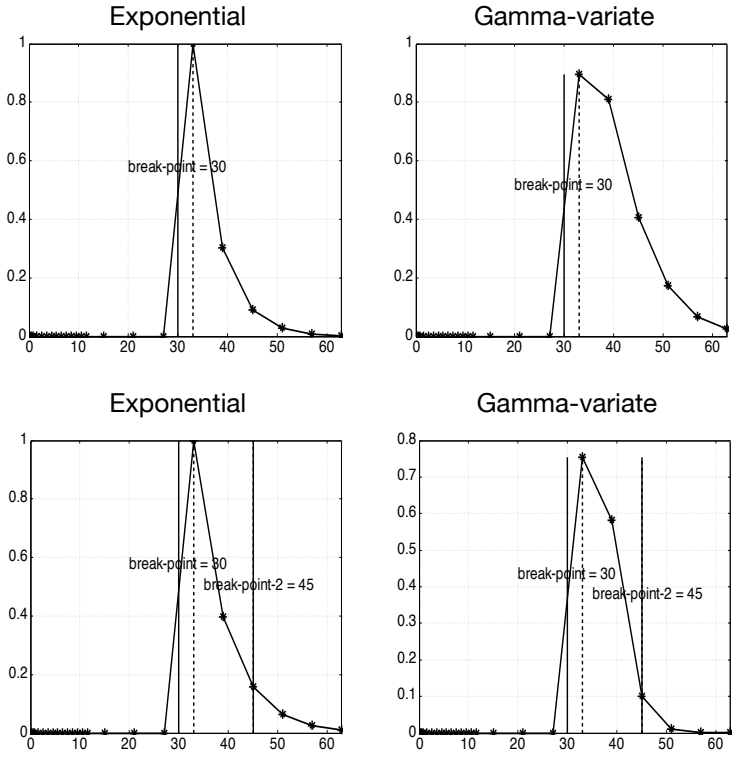


Figure 2.3: Graphs representing possible activation functions depicting dynamic alterations in radiotracer binding.

dependent parameter, due to the simple shape of the TACs. Thus only the factor γ related to the change in BP_{ND} was preserved and the others were omitted from the final model.

Furthermore, the ability to identify the time-course of the activation was acknowledged. In the original model by Alpert and colleagues [5] time-dependent activation functions had the shape of exponential function based on earlier studies [33], which peaks relatively soon after the task initiation and decays more or less rapidly depending on the decay parameter τ. Alpert and colleagues claimed that $\tau \approx 0.2$ min^{-1} yielded a good fit with true human data of spatial planning task, suggesting that the activation effect was nearly abolished after 10 minutes.

Nevertheless, in the model formulation of Alpert and colleagues [5] and later in the application of the model by others [8 – 10] focus has been on the detection of the activation rather than in the temporal characterization of the neurotransmitter response.

In a more recent work by Normandin and colleagues [76] LSRRM was extended to more readily incorporate the temporal characterization of the neurotransmitter fluctuations, which they called linear parametric neurotransmitter-PET (lp-ntPET). In contrast to original LSRRM where the neurotransmitter response onset coincides with the task initiation and the peak-time follows immediately, lp-ntPET permits varying onset and peak-time as well as different shapes for the decaying part.

The model of Normandin and colleagues was implemented using collections of simplified gamma-variate functions [64] as the basis activation functions. In their proposal the set of basis functions consisted of 897 gamma-variate functions depicting various different onset and peak-times and decay velocities. Their proposed approach is similar to the basis function method often employed in parametric image generation with SRTM [40], but differs in the parameter of interest.

The optimal basis function depicting the neurotransmitter fluctuation within LSRRM is identified as the one that yields the best fit with the measured data, that is, shows the smallest weighted least squares in linear optimization. Normandin and colleagues validated lp-ntPET using ^{11}C-raclopride data from rat experiments with intra-cranial infusion of methamphetamine [76]. Their finding was that lp-ntPET did not show any activation in a sham scan while in the true intervention the dopamine response was significant and agreed well with the invasive microdialysis measurements, thus the ^{11}C-raclopride measurement reflected well the changes in brain dopamine levels showing robustness of the novel method.

In a recent work by Morris and colleagues [74] lp-ntPET was

used to generate movies of cigarette smoking induced release of dopamine, showing feasibility of the novel approach. Through rigorous statistical treatment of the lp-ntPET data they were able to identify the voxels with significant dopamine release during smoking versus a control condition, and furthermore to represent the temporal characteristics of the activations in each voxel [74]. Their finding was that the cigarette smoking induced fluctuations in dopamine release may have different spatial and temporal profile between male and female smokers [74], thus indicating the importance of individual temporal profiling in neurotransmitter activation studies.

The LSRRM and lp-ntPET seek to detect the neurotransmitter activation through statistical analysis of the γ parameter. Albeit the elegance of the statistical treatment it would often be desirable to quantify the change in terms of $\mathrm{BP_{ND}}$, similarly as in ESRTM where two different $\mathrm{BP_{ND}}$ are explicitly sought. Disadvantage of ESRTM on the other hand is the inflexibility with regard to the timing of the neurotransmitter fluctuations that is inherent in LSRRM and lp-ntPET.

$\mathrm{BP_{ND}}$ of baseline in LSRRM is solved by reorganizing the parameters in Eq. 2.4:

$$BP_{ND} = \frac{k_2}{k_{2a}} - 1 \tag{2.6}$$

similarly the time-dependent $\mathrm{BP_{ND}(t)}$ is

$$BP_{ND}(t) = \frac{k_2}{k_{2a} + \gamma h(t)} - 1. \tag{2.7}$$

The time-dependent $\mathrm{BP_{ND}(t)}$ can be extracted at the activation peak-time or for instance the weighted average can be calculated to represent the mean $\mathrm{BP_{ND}}$(intervention). The mean $\mathrm{BP_{ND}}$(intervention) should coincide with the one calculated using ESRTM up to the differences in the accuracy of model fit.

The explicit presentation of the BP_{ND}(baseline/intervention) allow calculation of RO_{PET} and comparison with earlier studies.

2.2 Principles of positron emission tomography

Quantitative accuracy of PET is a unique feature that allows dynamic measurement of the radioactivity distribution and subsequently kinetic analysis of the data. Often quantification of the PET frames is attained through calibration measurements that provide a conversion factor between counts/sec and radioactivity concentration e.g. Bq/mL. Considering the acquisition and image reconstruction process in PET there must however exist a lower threshold $\gg 0$ of count statistics that is required for accurate image reconstruction.

Acquisition of a PET frame is ultimately a counting process of rare nuclear events - in the signal processing parlance a Poisson counting process. Thus, the probability of detecting an event approaches zero with decreasing detection time and/or observed area, and the total number of detections per sample point dictate the statistical quality of the measurement. The total number of detections (within a fixed time interval) can be increased to some degree by optimizing the sensitivity of the detectors. The total number of positron emissions in turn is intentionally kept low due to the tissue damaging properties of the annihilation energy (photons), and the tracer doses with healthy volunteers are strictly limited by radiation safety authorities. In PET these limitations translate into relatively small radioactivity doses and thus poor count statistics of the PET data.

Partial volume effect caused by poor spatial resolution of PET has on the other hand been regarded as a major limiting factor of PET brain studies [75]. In the classical simulated experiment by Mazziotta and colleagues [69] signal blurring by a 5 mm

kernel (full width at half maximum (FWHM)) corresponded with up to 20% underestimation in a thin high contrast region, while signal blurring by a 2.5 mm kernel (FWHM) halved the error to 10%.

The intrinsic spatial resolution of a PET scanner is defined by the surface area of the photon-detectors (detector crystals, scintillators), while engagement of small crystals to achieve improved spatial resolution has been hindered by a few obstacles. For instance, production of the small detector crystals is complicated and the stopping power of a small crystal may be insufficient. The scintillation cascade occurs when the arriving photon interacts with the crystal material resulting in photoemission. If the crystal size is however very small the arriving photon may pass through many crystals before the scintillation cascade occurs and the detection will be misplaced by a few millimeters, a phenomenon called crosstalk of scintillators. Furthermore, the signal from a single crystal is weak and it must be amplified prior to further processing, and until recently design and manufacturing of the amplifiers have significantly limited the application of small crystals.

The true reconstructed spatial resolution of PET is commonly measured using the procedures in the National Electrical Manufacturers Association (NEMA) standards. NEMA standards describe the spatial resolution of PET devices as the FWHM measured with an infitesimal point source in air. The purpose of the measurement is to characterize the point spread function (PSF) of compact sources in reconstructed image, and it provides the best-case comparison among scanners. Furthermore, the NEMA standards describe the point source measurements at different locations in the field of view (FOV), and the calculations in axial as well as transaxial directions. Table 2.1 summarizes the NEMA spatial resolution results for a brain-dedicated and a few whole-body PET scanners according to

the literature. An approximately 5 mm (FWHM) spatial resolution is achievable using the whole-body PET scanners while the brain-dedicated PET scanner is characterised by a nearly isotropic spatial resolution of approximately 2.5 mm (FWHM).

Table 2.1: National Electrical Manufacturers Association (NEMA) spatial resolution results from literature in full width at half maximum (FWHM).

	Transaxial (mm)		Axial (mm)			
Scanner	1 cm	10 cm	1 cm	10 cm	Avg.	Ref.
Siemens HRRT	2.30	2.90	2.50	3.40	2.78	[28]
Siemens HR+	4.40	5.00	4.00	5.50	4.73	[17]
Philips Ing.	4.70	5.15	4.60	5.00	4.86	[109]
GE D690	4.70	5.07	4.74	5.55	5.01	[13]
GE DVCT	5.12	5.68	5.18	5.86	5.46	[89]

Due to the Poisson characteristics, high-resolution PET is however associated with poorer statistical quality than whole-body PET and for e.g. diagnostic purposes visually appealing appearance of imaging data may be favored. Nevertheless, considering the experiments of Mazziotta and colleagues [69] the best achievable spatial resolution may often provide an indispensable advantage in nt-PET studies. Thus, in the light of advanced temporal characterization of neuroreceptor occupancy, using e.g. lp-ntPET derived analysis, careful optimization of the PET assessment with regard to both the temporal and the spatial resolution is required.

2.2.1 Data acquisition

PET data acquisition can be initiated after the radiotracer is transported to blood circulation of the imaged subject. Radiotracer travels through the blood stream (also) to the organ of interest which is positioned within the FOV of the PET scanner.

PET scanner consists of thousands of (typically) circularly arranged detectors capable of detecting low energy photons. Low energy photons arise from the annihilation process following a positron-electron collision. Annihilation occurs in the immediate vicinity of the positron-emission location and produces two opposite-direction gamma-quanta.

Gamma-quanta of known energy (511 kilo electron volt (keV)) travel at the speed of light to opposite directions and sometimes hit the detectors of the PET scanner (nearly) simultaneously. These observations are called coincidence events. If simplified, the coincidence events are assumed to originate from single positron-electron annihilation that occurred within the LOR that connects the two detectors.

However, practical electronic implementation of the PET scanner does not permit measurement of only those events with the energy of exactly 511 keV or exactly simultaneous detection; albeit the gamma-quanta travel at the speed of light a small timing difference occurs when the annihilation location is near the edge of the FOV, and the electronics does not permit exact timing of detection (within nanoseconds). Thus, the detector electronics are typically tuned to record events in an energy window of e.g. 350-650 KeV and timing window of 5-10 nanoseconds. Due to the broad energy and timing windows, unwanted erroneous observations have significant contribution in the measured data.

A coincidence event (true or random) is such that no more than two photons are detected within the time window (opened at the first detection and closed after 5-10 nanoseconds). A multiple event is such that more than two photons are detected within the time window, and these are usually ignored automatically. A random event is such that the two detected photons actually originate from different annihilations. Scattered events originate from the same annihilation but one of the gamma-

quanta has been disoriented (scattered) from the original path while travelling through the medium. The true, multiple, random and scattered events together are called prompt events without distinction.

The amount and distribution of the random and scattered events can be estimated from the measured data using established techniques (see [23] for review). Furthermore, absorption of the photons within the medium can be estimated using external photon transmitting source that travels around the medium and the non-absorbed energy is measured on the opposite side. The so-called attenuation correction compensates for the non-homogeneous contribution of the photon absorption within the medium.

2.2.2 Image signal-to-noise ratio

Statistical quality of the image data is commonly described by SNR that relates the true signal to the background noise. In imaging studies SNR is often defined as

$$SNR = \frac{\text{mean}(sig)}{\text{std}(sig)},$$

where mean(sig) depicts the mean and std(sig) the standard deviation (SD) of the true signal. The mean and SD are preferably determined from a high number of replicate measurements rather than more commonly employed voxel-by-voxel mean and SD within some predetermined region. SNR larger than some predefined threshold designates that the image contrast is adequate to distinguish an object from the background.

There is a connection between image SNR and so-called noise equivalent counts (NEC) derived from the PET acquisition statistics. NEC relates the square of the true counts (coincidences) to the total counts in a PET measurement and it can

be defined as [26]

$$NEC = \frac{((1-S)(P-R))^2}{P+R}, \tag{2.8}$$

where S is the scatter fraction, and P and R the prompt and random events, respectively. In Equation 2.8 the denominator $P + R = T + S + 2R$ reflecting the noise contribution from random events when they are calculated online during the PET measurement. Often, if the random events are estimated offline using a noise suppression algorithm the denominator is simply P. NEC is a commonly applied quantity in scanner performance evaluations and it has been shown to correlate well with the square of the reconstructed image SNR in a physical test object (phantom) measurement[26].

2.2.3 Image reconstruction

Analytical image reconstruction

Assuming the observations have been corrected for the unwanted events, summation of the true events results in quantities that approximate the line integrals through the radioactivity distribution. In the following theory is presented in the Cartesian plane (two-dimensional (2D)) for simplicity although PET is truly three-dimensional (3D). In the Cartesian plane line integral of a function can be expressed as

$$g \triangleq \int_L f(x,y)\, dl\,. \tag{2.9}$$

where L is the LOR connecting the two detectors, $f(x, y)$ is the unknown spatial radioactivity distribution (or emission density) and g is the observation.

If the LOR is given coordinates (s, θ), s being the perpendicular distance from the FOV origin and θ the angle to y-axis,

the line integral is expressed as [50]

$$\begin{aligned}\mathcal{R}_{(s,\theta)}\{\, f(x,y)\,\} &\triangleq g(s,\theta)\\ &= \int_{L_{(s,\theta)}} f(x,y)\, dl\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - s)\, dx\, dy.\end{aligned}$$

where $\delta()$ is the Dirac's delta-function.

Collection of all $g(s,\theta)$ through the unknown $f(x,y)$ is often called the Radon transform of object $f(x,y)$, for it was first introduced by Johann Radon (see [80] for English translation). It was also Johann Radon who first showed that the object $f(x,y)$ can be fully recovered from its projections, that is, the Radon transform can be inverted.

A simple proof of this analytical inversion algorithm can be formulated by considering the direct connection of the two-dimensional Radon transform of an object and its one-dimensional Fourier transform (central-slice theorem) [54]; if the two-dimensional Fourier-space can be filled with the one-dimensional Fourier transforms of the projections, the original object can be recovered by inverse two-dimensional Fourier transform (2DFT).

In X-ray CT-imaging the direct inverse fourier transform (DIFT) reconstruction is successful since the line integrals can be collected in a high enough number of angles, and with high enough X-ray energy the noise contribution in the measurements is tolerable [54]. In PET however (and in low-dose CT), the small number of projection angles, low sampling density and low SNR hamper the usage of DIFT significantly [65], although improved algorithms for the critical re-sampling step from polar to Cartesian coordinates have been recently developed. Instead of DIFT the analytical reconstruction is commonly achieved using the filtered backprojection (FBP) algorithm [54] that generates less artifacts due to limited projections and measurement

noise. Furthermore, the analytical FBP algorithm has been extended to 3D data, in particular the so-called three-dimensional reprojection FBP (3DRP) algorithm that exploits 2D-FBP in the estimation of missing data in the oblique segments prior to 3D-FBP is widely employed [57].

Iterative image reconstruction

The aim of the analytical image reconstruction algorithm is to recover the unknown spatial distribution of the object by mathematically solving the analytical representation of the measurements, in other words assuming deterministic data [79]. In PET however, this approach is suboptimal since the measurements represent stochastic rather than deterministic data.

A more complete model of the data collection can be crafted by taking into account the statistical nature of the collection process. Collection of the coincidence events in a given time interval can be regarded as a Poisson counting process [83], thus the measurements represent samples of Poisson random variables.

Let us first define a discrete model of the data collection. In the computer implementation a 3D object $f(x, y, z)$ is sampled into N (typically) cubicles (or voxels), let us denote them as $f = (f_1, f_2, \ldots, f_N)$. Now there are N voxels and the jth voxel has the value f_j. Also the number of LORs is finite, let us denote them as $g = (g_1, g_2, \ldots, g_M)$, where g_i is the measured value of coincidence events within LOR i.

In the discrete model a geometric correlation matrix (system matrix) with $N \times M$ elements can be calculated that determines the contribution of each voxel to the measurement of g. Let us denote the system matrix by $A = \{a_{11}, a_{12}, \ldots, a_{NM}\}$, with each a_{ij} representing the probability that an annihilation occurring at voxel j will be detected at LOR i. Now the detection process can be modelled by $g = Af$.

Let us denote the independent Poisson random variables describing the number of decays in voxel j and number of decays detected in LOR i by $F(j)$ and $G(i)$, respectively. Assuming that all decays are detected we have for the voxel j and unknown mean λ_j a probability that n decays are detected

$$P_{\lambda_j}(F(j) = n) = e^{-\lambda_j} \frac{\lambda_j^n}{n!}$$

The expected value of the random variable $G(i)$ is thus

$$\lambda_i = \sum_{j=1}^{N} a_{ij} \lambda_j, \quad \forall i$$

since the detection probabilities a_{ij} are independent.

Similarly as for the voxel data we get for LOR i and unknown mean λ_i a probability that the measured data of LOR i gets the value g_i

$$P_{\lambda_i}(G(i) = g_i) = e^{-\lambda_i} \frac{\lambda_i^{g_i}}{g_i!}.$$

Let us introduce a random variable x_{ij} for the number of emissions occurred in voxel j and detected in LOR i. It is a Poisson random variable with a mean $\lambda_{ij} = \lambda_j a_{ij}$. Now we have for the number of detections g_i in LOR i

$$g_i = \sum_{j=1}^{N} x_{ij}.$$

Likelihood function of the measured data is that of the multinomial distribution

$$l(\lambda) = l(\lambda_1, \lambda_2, \ldots \lambda_N) = \prod_{i=1}^{M} e^{-\lambda_i} \frac{\lambda_i^{g_i}}{g_i!}$$

and the log-likelihood function is

$$\begin{aligned} L(\lambda) = \ln l(\lambda) &= \sum_{i=1}^{M} \ln \left(e^{-\lambda_i} \frac{\lambda_i^{g_i}}{g_i!} \right) \\ &= \sum_{i=1}^{M} \left(-\lambda_i + g_i \ln (\lambda_i) - \ln (g_i!) \right) \\ &= \sum_{i=1}^{M} \left(- \sum_{j=1}^{N} a_{ij} \lambda_j + g_i \ln \sum_{j=1}^{N} a_{ij} \lambda_j - \ln (g_i!) \right). \end{aligned}$$

Thus the task is to find λ to maximize the log-likelihood function but without having the image data to calculate the log-likelihood function [72]. So instead the *expectation* of the log-likelihood function is maximized given the measurements g and the current estimate of λ.

This approach is called maximum likelihood expectation maximization (MLEM) and it can be employed to define a PET image reconstruction algorithm, where the expectation- and the maximization-steps are combined into one equation [72]. MLEM is an iterative algorithm to find the most likely image given the measurements. Let us denote the initial voxel value at voxel i as f_i^0, then for iterate f_j^k we have

$$f_j^k = f_j^{k-1} \frac{1}{\sum_{i=1}^{M} a_{ij}} \sum_{i=1}^{M} \frac{g_i a_{ij}}{\sum_{l=1}^{N} f_l^{k-1} a_{il}}. \tag{2.10}$$

The iterative MLEM algorithm has been shown to converge to the focal point of the log-likelihood function, without decreasing the likelihood at any iteration [83]. In practice the convergence can however be very slow [79], and typically the reconstruction is stopped at some prefixed number of iterations. This is due to the computational cost of the MLEM algorithm - each iteration lasting approximately the same time as FBP reconstruction - but also due to the ill-conditioned problem of finding the maximum likelihood (ML) estimate [85].

The late iterates of the MLEM are typically very noisy, and therefore the intermediate images are often visually more appealing. One way to reduce the voxel-wise noise is to post-smooth the image using for instance a three-dimensional gaussian smoothing function. A more general approach to limit the noise contribution is to apply spatially penalized-likelihood objective function. A good example of penalized-likelihood based regularization of the MLEM is the median root prior (MRP) algorithm introduced by Alenius and Ruotsalainen [4].

To speed up the convergence of the iteration in Eq. 2.10 a subsetized version was introduced in [46]. The subsets consist of a number of projections unique in each subset. Often the projections are chosen to define orthogonal basis [46], and the resulting algorithm is called ordered subsets expectation maximization (OSEM). In the OSEM algorithm the image is updated after completion of each subset, and the convergence is thus faster, while, the amount of computation of one full iteration is similar to that of MLEM.

Choice of the reconstruction method

In the light of NEC the choice of image reconstruction algorithm has the greatest impact on image accuracy versus SNR. In general, the analytical image reconstruction algorithms are regarded more accurate while the iterative algorithms are thought to yield better SNR but may require a higher NEC for accurate image reconstruction. Application of the iterative reconstruction algorithms is obstructed by the lack of explicit stopping rules for the iterative process, while the outcome of analytical reconstruction depends on the filter dimensions. Filter in the FBP reconstruction can be adjusted to yield a smooth appearance, while the iterative reconstruction can be stopped early and subsequently filtered using a large kernel size (5-10 mm FWHM) in the 3D Gaussian post-smoothing step. On the other hand,

if PET imaging is performed for research purposes to assess e.g. receptor binding the best achievable spatial and temporal resolution are often required to protect against partial volume effect (PVE) and to be able to distinguish rapid fluctuations in neurotransmission, respectively.

In [26] Dahlbom and colleagues showed that FBP reconstruction yielded inferior SNR^2 compared to OSEM reconstruction at the same NEC levels. However, in [26] the OSEM iterations were stopped very early although it was anticipated that later iterates approach SNR^2 of FBP. Nevertheless, NEC is a measure of count statistics that takes into account the variable randoms proportion at different count rates, and correlates with image SNR^2, thus it is recommended as a measure of count statistics.

In [15] Boellaard and colleagues employed physical phantoms to explore the relative image accuracy as a function of NEC as well as to study the acquired image SNR using various image reconstruction algorithms. Error in the image estimates (reconstruction bias) was calculated as bias(%) relative to the FBP reconstruction which is thought to be bias-free in reasonable count statistics. Among the iteratively reconstructed images they found biases of up to -15% in high contrast regions and up to 40% in the low contrast regions [15]. There were however large differences between the tested iterative methods with regard to the magnitude of bias(%) but all algorithms showed similar behaviour - there was a sustained steep decline in the image accuracy following the decline in NEC statistics.

The observed dual faceted bias in the high and low contrast regions indicates attenuated contrast recovery in the low count statistics as compared to high count statistics reconstruction. Thus suggesting a different degree of convergence of the iterative reconstruction at different count statistics. Boellaard and colleagues [15] did not experiment with different number of iterations but their experiments showed at least the risk of using

predetermined number of iterations in varying count statistics. In their experiments the image SNR of FBP was again inferior to that by iterative algorithms [15], but again at early iterations (c.f. [26]).

2.2.4 High resolution PET

The High Resolution Research Tomograph (HRRT) offers the current state-of-the-art human-size high-resolution PET imaging in the field (c.f. Table 2.1). The HRRT scanner differs from conventional scanners for its planar detector design. Typically the detector modules are small and they are arranged on a ring around the scanner FOV. In the HRRT, however, a different approach was chosen to allow quadrant sharing technique of the photomultiplier tubes on the one hand and to allow the depth of interaction (DOI) measurement on the other hand [108]. The HRRT scanner consists of eight planar detector heads arranged in an octagon around the scanner FOV, which results in large areas devoid of detectors at the seams of two adjacent heads. This feature renders the HRRT data discontinuous with gaps corresponding to the location of the joints between detector heads [70]. If the gaps are not properly treated the analytical reconstruction with missing data results in sizeable artefacts, practically precluding the direct usage of analytical reconstruction with the HRRT data [70].

2.2.5 Image reconstruction strategies for the HRRT

Dedicated image reconstruction strategies for the HRRT can be categorized into

a) gap-filling plus analytical reconstruction

b) iterative statistical reconstruction using a system matrix incorporating the missing data.

In the iterative reconstruction the system matrix can be easily modified to incorporate the inherent missing data in the HRRT measurements, while the analytical reconstruction generates sizeable artefacts in presence of gaps in the sinograms [29]. However, the early evaluations of analytical reconstruction suffered from poor gap-filling.

The recent advancements in the gap-filling strategies as well as the HRRT-specific implementation of full 3D analytical reconstruction using the 3DRP [57] algorithm have helped to alleviate the reconstruction artefacts considerably. The current more auspicious gap-filling methods have been extensively studied by Tuna and colleagues [92 – 95] and by van Velden and colleagues [99 – 101].

The proposed gap-filling methods can be categorized into

a) model-based methods that employ forward projection of the iteratively reconstructed image [28];

b) constrained Fourier space (CFS) methods [55]

c) discrete cosine transform (DCT) domain methods [94, 95]

d) transradial bicubic interpolation method [92, 93].

The CFS, the DCT-domain and the bicubic transradial interpolation gap-filling methods have all performed well in phantom studies, but the latter method outperforms the others in computation times; the bicubic transradial interpolation works in nearly linear time while the others can be as slow as full iterative reconstruction. In contrast, the gap-filling based on forward projected data would allow consistent estimation of the data within the gaps but it fails to model the noise in the

missing parts; thus generating noise dependent artefacts in the analytical reconstruction.

The most recently proposed method employing a transradial bicubic interpolation approach has performed well in phantom studies in comparison to DCT domain method but with much lesser computational cost. In short, the transradial bicubic interpolation approach can be described as follows: the transradial slice in the sinogram domain is defined as the slice in the direction of the radial samples, as it happens that the missing data in the transradial slice lie in parallel lines in the sinogram domain they can be omitted, yielding a shrunk, gap-free sinogram (see Figure 3 in [92] for illustration). The shrunk sinogram is resized back to the original dimensions using bicubic interpolation [56], and rows corresponding to the gaps are inserted to the original sinogram. Thus the transradial bicubic interpolation method can provide fast and robust estimation of the gaps.

Iterative statistical reconstruction using the MLEM algorithm and its variants allow usage of non-complete projections directly. Therefore - and for the more complete model of the measurement - the iterative full 3D reconstruction has been preferred for the HRRT image reconstruction [70]. However, algorithm based on the Equation 2.10 with pre-corrected measurement data resulted in considerable bias as was shown in a phantom experiment [24]. The bias was attributable to the non-negativity constraints in Equation 2.10 and the many negative observations truncated to zero occurring in the pre-corrected measurements. Albeit the HRRT is regarded sensitive relative to conventional scanners [28], the area covered by a single LOR in the HRRT is a small fraction of that in a conventional scanner, thereby rendering the number of detections per LOR in the HRRT very low (c.f. discussions in Chapter 2.2.2).

To avoid the zero-truncation bias a HRRT-specific modification to the MLEM equation (Eq. 2.10) was presented by

Comtat and colleagues [24]. The algorithm of Comtat and colleagues [24] is called ordinary Poisson ordered subsets expectation maximization (OP-OSEM) for it aims to maintain the Poisson distribution characteristics in measured data, which are normally violated in the pre-correction process. The modified equation is expressed as

$$f_j^k = f_j^{k-1} \frac{1}{\sum_{i=1}^{M} \frac{a_{ij}}{N_i A_i}} \sum_{i=1}^{M} \frac{p_i a_{ij}}{\sum_{l=1}^{N} f_l^{k-1} a_{il} + (d_i N_i + s_i) A_i}, \qquad (2.11)$$

where p_i represent the prompt coincidences, d_i the randoms and multiples estimated from the delayed coincidence window data [18], N_i the scanner normalization factor, and A_i the attenuation correction factor, all for LOR i. Formulation in Equation 2.11 removes the need for zero-truncation, thus the bias should be avoided, but it was also noted that the additive term in the denominator of 2.11 slows down the convergence [24].

The OP-OSEM algorithm was shown quantitative in a uniform cylinder phantom experiment which resulted in significant bias with the conventional OSEM reconstruction [24], and it was therefore recommended for dynamic brain studies on the HRRT. However, in an experiment with more realistic anthropomorphic phantom data the OP-OSEM reconstruction was found to suffer from remnant bias at very low count statistics, but still realistic in dynamic human brain studies [98].

Flexibility of the iterative framework has been exploited to accommodate the signal blurring function, or PSF, in Equation 2.11 [25]. Let us denote the signal blurring kernel in the image space as h, then the update equation with resolution modeling

(RM) is expressed as

$$f_j^k = f_j^{k-1} \frac{1}{\sum_{i=1}^{M} \frac{a_{ij}}{N_i A_i * h}} \sum_{i=1}^{M} \frac{p_i a_{ij}}{\sum_{l=1}^{N} f_l^{k-1} a_{il} * h + (d_i N_i + s_i) A_i}. \tag{2.12}$$

The algorithm depicted in Equation 2.12 is called resolution modeling OP-OSEM (RM-OP-OSEM).

The convolution kernel h for the HRRT has been estimated using pseudo point-source measurements within the FOV of three HRRT scanners [25]. In the work of Comtat and colleageus [25] isotropic convolution kernel was formed on the basis of the three sets of measurements.

The RM-OP-OSEM showed reduced PVE without noise-amplification in a human experiment [88]. It has been also found that the RM-OP-OSEM can significantly reduce the low count statistics bias as compared to OP-OSEM [105]. On the other hand the PSF-based resolution modelling has been shown to suffer from edge artefacts [91].

Other modifications to the MLEM equation have been suggested that might be more robust with regard to the low count statistics, such as the AB-MLEM [19] and the NEG-MLEM [77] algorithms for sinogram-based reconstruction. Furthermore, the list-mode-based reconstruction algorithm (MOLAR; [20]) using the conventional OSEM algorithm have been shown less vulnerable to low count statistics bias [51]. It can be however regarded that the list-mode-based reconstruction using the conventional OSEM algorithm does not provide theoretical advantage with regard to the low count statistics bias, while the optimization of the additional parameters introduced in AB-MLEM and NEG-MLEM algorithms may depend on the imaging conditions and thus become intractable relative to the expected advantage of using these methods.

2.3 Region of interest analysis

A commonplace practice to improve the SNR in brain PET studies is to pool the signal from functionally similar image elements. Aggregation of the signals is achieved through region of interest (ROI) analysis, consisting of the ROI-definition and averaging of the signal cross the image elements within the ROI.

ROI-analysis can be applied either on parametric images or on the PET frames to represent regional pharmacokinetics of the tracer. The outcome of the latter approach is a TAC that represents the tracer pharmacokinetics at improved SNR, which can then be applied for more robust estimation of the physiological parameters e.g. BP_{ND}. Topographical organization of the brain allows the aggregation of adjacent image elements, but inaccuracies in the ROI-definition can easily result in amplified PVE.

In the present work the focus was on the ROI-delineation strategies for dopamine (DA) receptor type 2/3 PET-imaging, although optimization of ROI-delineation strategies is of high importance in various nt-PET studies targeting different receptors. The choice is motivated by the facts that DA PET-imaging, in particular using ^{11}C-raclopride, constitutes a widespread methodology in human neurotransmission research, and its ROI-analysis has been particularly demanding due to vast functional repertoire of the striatum, the hub of dopaminergic transmission.

2.3.1 Implications from the functional organization of the striatum

An illustrative schematic representation of the striatal connections can be found in [41] Fig. 12 by Haber and colleagues. Anatomically striatum is divided into caudate nucleus (CAU),

putamen (PUT) and nucleus accumbens (NAcc), while the functional division has ventro-medial to dorso-lateral and rostral to caudal gradients not obeying the anatomical division [41]. The limbic striatum (LSTR) which includes the NAcc and the ventral parts of the rostral CAU and PUT receives its direct frontal input from the orbital and medial prefrontal cortex (OMPFC) [41], which is thought to have an important role in the development of reward-guided behaviours [41].

Dopamine has an important role in the neuromodulation of the striatal signalling. The striatum has the highest concentration of DA receptors in the human brain, and it mediates most of the dopamine-related signalling. The role of e.g. LSTR has been widely studied in different pathological states involving the reward system using ^{11}C-raclopride and other dopamine $D_{2/3}$ receptor specific tracers, mostly in drug abuse reinforcement [103] but recently also in connection with pathological gambling [53]. In contrast, the associative striatum (ASTR) receives its direct cortical inputs mostly from the dorsolateral prefrontal cortex (DLPFC) which is thought to be involved in cognitive functions such as working memory [39], while the sensorimotor striatum (SMST) receives its direct cortical inputs mostly from motor and pre-motor cortical regions.

Differentiated response to pharmacological as well as cognitive and behavioral activations have been reported in the functional striatal subregions as measured using ^{11}C-raclopride-PET [32, 66]. It has been regarded that the functional division of the striatum would be more relevant in PET imaging studies investigating alterations in DA levels than the anatomical division. Furthermore, introduction of the new high-resolution PET scanners has opened new possibilities in the striatal investigations highlighting the importance of accurate striatal subdivision.

For instance, in our recent study we showed significant rostro-

caudal gradients in the striatal BP as measured using ^{11}C-raclopride and the HRRT [2]. In the caudate nucleus and limbic striatum there was a declining trend, while in the putamen there was an inclining trend of BP_{ND} in rostro-caudal axis. Differences between the rostral and caudal BP_{ND} were in the order of 10-20%.

In a recent study Tziortzi and colleagues [97] showed increased sensitivity to DA level alteration using functional subdivision as compared to anatomical subdivision as measured with ^{11}C-raclopride and d-amphetamine challenge. Tziortzi and colleagues employed a MRI technique called probabilistic tractography [12], that can be used to explore the connectivity based (CB) functional subdivision of the striatum.

The process of CB functional subdivision of the striatum employed in their study is illustrated in [97] Fig. 1. In the first phase they investigated the projections from the striatum to the frontal, parietal, temporal and occipital cortices. In probabilistic tractography each striatal voxel was designated probabilities at which its connections would reach a given cortical region. Then, the striatum was segmented according to these probabilities by assigning each voxel to the lobe with highest probability.

The temporal lobe innervates a very small territory in the ventro-medial striatum, while the frontal lobe was assigned to approximately 2/3 of the striatal voxels (see Figure 1 in [97]). In the second phase the frontal lobe was further subdivided into functional regions of limbic, executive, rostral-motor and caudal-motor cortex, according to literature not based on individual functional-MRI measurements. Tractography analysis was rerun for the striatal voxels assigned to frontal lobe to find striatal territories associated with the aforementioned frontal lobe subregions. The CB functional subdivision was in good concordance with the previous studies in humans and primates (c.f. Figure 12 in [41]). The individual subdivisions were in

addition analysed in Montreal Neurological Institute (MNI) normalized coordinates in orded to define probabilistic atlases at different probability thresholds and voxel sizes.

2.3.2 ROI delineation strategies in the striatum

Methods for ROI delineation can be categorized into manual and automated. Manual ROI delineation is typically achieved by using an anatomical reference image (in the brain studies mostly T1 weighted MRI images), a dedicated ROI delineation software and expertise of a trained researcher. In ^{11}C-raclopride studies the manual ROI delineation in the striatum is typically achieved according to the guidelines described by Mawlawi and colleagues yielding the functional ROIs of limbic striatum (LSTR), associative striatum (ASTR) and sensorimotor striatum (SMST) [66, 67].

To date the manual ROI drawing has been regarded as the gold standard, for the ability to accurately delineate the gray matter boundaries. However, the manual interaction introduces a subjective component to the analysis that can produce biased results between subject groups and between operators [44]. Moreover, the ever-increasing spatial resolution of the PET images requires more manual labour due to higher sampling density - in a typical high-resolution ^{11}C-raclopride-PET study the ROIs are drawn on over thirty coronal slices.

Automated ROI delineation methods can provide operator-independent and reproducible segmentations that do not require manual labour of expensive experts [30]. Furthermore, automated methods might enable more accurate functional subdivision of the striatum as compared to the manual method. Automated methods can be categorized into atlas-based methods, and those that segment the imaging data directly, as depicted

in Figure 2.4.

Atlas-based methods employ a mapping from the atlas defining space to the individual space that is subsequently applied to individualize the atlas-mask image. The mapping is typically calculated using non-linear matching between individual anatomical MRI image and an anatomical template image, but can be also achieved between individual PET image and a PET template. Direct segmentation can be obtained from the anatomical or functional imaging data. Method depicted in Figure 2.4 employs the known gradients in ^{11}C-raclopride binding [2] to find clusters of functionally homogeneous voxels from the PET image [90].

Currently, there is a consensus that the functional subdivision should be favoured over the anatomical division in the ROI analysis of the ^{11}C-raclopride studies. However, the manual method based on anatomical landmarks is still regarded as the gold standard for ROI delineation. The rules for manual delineation described by Mawlawi and colleagues [66, 67] and later by Tziortzi and colleagues [96] have served well, but there is still room for subjective bias, especially in the LSTR, that might confound the analysis.

Recently, there have been advances in the definition of the striatal subdivision [97] as well as in the PET imaging through the introduction of probabilistic tractography and high-resolution PET scanners, respectively. Currently, there exists automated methods that might provide more robust functional ROI delineation than manual ROI drawing. At current the feasibility of the atlas-individualization approach using the novel CB-atlases is however unknown, thus, comparison with the current ROI-methodologies is warranted in order to assess its feasibility.

In addition to the striatal ROIs the pharmacokinetic analysis of ^{11}C-raclopride studies using reference tissue methods requires a reference region ROI. The cerebellar cortex (CERC)

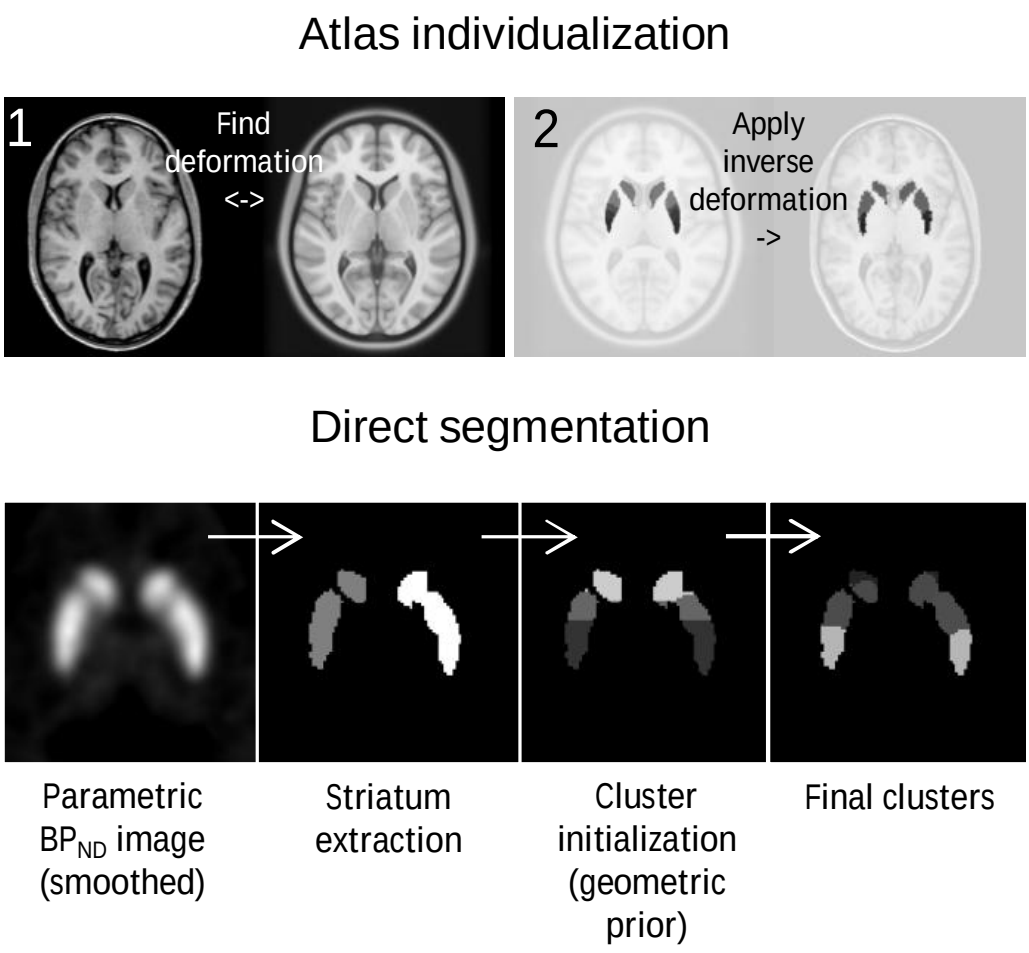


Figure 2.4: Illustration of methods for automated ROI generation that apply atlas individualization or direct segmentation of the ^{11}C-raclopride PET BP_{ND} image.

is commonly used as a reference region as it is almost devoid of $D_{2/3}$ receptors, and its usage has been validated before [61]. Currently, manual drawing is often applied in the reference region ROI delineation but automated methods based on atlas individualization exist. In a recent study Schain and colleagues compared manual delineation, to that of two atlas individualization for the extraction of reference region ROI [82]. Their

finding was that manual delineation might be superior to that of automated methods, but the behaviour is somewhat ligand dependent. For fully automated analysis of ^{11}C-raclopride studies it would be highly beneficial to obtain the reference region ROI also automatically.

2.4 Summary

Dynamic PET offers a versatile tool to assess receptor binding. Receptor binding assays can be used to probe direct occupation by exogenous drugs or indirect occupation by endogenous neurotransmitter. In particular the rapid fluctuations in endogenous neurotransmitter release due to cognitive/behavioural stimulation constitute an endeavour that is not easily achieved, but holds great potential with regard to better understanding the pathophysiology of various disorders and also the brain function in health.

The current state-of-the-art PET methodology employs single-injection measurements and analysis that is claimed to characterize the temporal profile of the neurotransmitter release. The setup has been solely employed with ^{11}C-raclopride but the theory applies to any tracer that has amenable features.

Temporal characterization of the neurotransmitter response may permit new kind of analysis on PET data that has not been available before. In cognitive/behavioural stimulations the achievable length of neurotransmitter activation is often short due to learning and habituation, thus implicating need for more accurate temporal modelling per se, but also some brain disorders involving neurotransmission may be largely of temporal nature rather than about magnitude.

The lp-ntPET model contains four explicit parameters (R_1, k_2, k_{2a}, γ) and three implicit parameters of the gamma-variate functions. The concern is whether the statistical quality of the

dynamic PET data allow robust estimation of all the parameters. Furthermore, the temporal sampling of the PET data must accommodate with the expected temporal profile of the neurotransmitter fluctuations, increasing the statistical variation as the frame length decreases. In addition, PVE due to poor spatial resolution of PET or due to head motion or inaccurate ROI-delineation may mask the subtle localized variations in the PET signal.

A major concern, however, is the absolute quantitative accuracy of the dynamic PET data. It is apparent from the lp-ntPET formulation that the risk of type I error in the statistical hypothesis testing is considerable. Considering the re-analysis of Egerton and colleagues [32] of the famous battlefield-game experiment of Koepp and colleagues [58] it is anticipated that rigorous error recognition and compensation is required in order to implement a cognitive/behavioural stimulation experiment to study neurotransmitter release.

At current, HRRT provides the best achievable spatial resolution in human brain PET studies but it is associated with considerable risk of quantification bias due to reconstruction inaccuracies. The so-called low count statistics bias in the iterative reconstruction is correlated with the measurement SNR and its impact should be carefully addressed in particular in lp-ntPET analysis. It is also unclear at which circumstances the HRRT quantification can be considered accurate.

ROI-based analysis of PET imaging data can in turn provide significant improvement in measurement SNR, but erroneous ROI-delineation may introduce additional bias. The most recent automated methods for striatal ROI-delineation show potential to improve the outcome but their feasibility has not yet been fully addressed.

3 Objectives

The purpose of the present work was to investigate the reconstruction strategies for high-resolution PET and the feasibility of the novel ROI-delineation strategies in DA receptor type 2/3 studies using ^{11}C-raclopride. The particular aim was to find and develop optimal methodologies to allow more accurate temporal characterization in nt-PET studies. The work is divided into three parts with the following objectives:

a) To investigate the quantitative accuracy of OP-OSEM algorithm for high-resolution PET image reconstruction, and further, to set the minimum NEC-statistics yielding a tolerable error, and translate the minimum NEC-statistics into temporal resolution. (Publications I and II)

b) To investigate the quantitative accuracy of RM-OP-OSEM and 3DRP with transradial bicubic interpolation gap-filling algorithms for high-resolution PET image reconstruction. (Publications III and IV)

c) To evaluate the feasibility of novel automated ROI-delineation methods in the human striatum in comparison to current manual ROI-delineation. (Publication V)

4 Experiments

The PET experiments presented in Publications I-V (referred to as studies I-V) were implemented using antropomorphic brain phantoms as well as real human test-retest measurements. Real high-resolution PET experiments presented in this work were all acquired using the HRRT-scanner.

Phantom experiments

Three different anthropomorphic brain phantoms were scanned for the reconstruction experiments:

a) the Hoffman phantom [43] (study I);

b) the Striatum phantom [60] (study I); and

c) the Iida phantom [47] (study III).

All the phantom scans were made in list-mode format, allowing retrospective framing of the data, and transmission scans were performed for μ-map calculation. All sinograms were generated from the list-mode data using span 9 and maximum ring difference of 67, while the reconstructions were made using $256 \times 256 \times 207$ image matrix with voxel-size of 1.21875 mm $\times$ 1.21875 mm $\times$ 1.21875 mm.

The Hoffman 3D brain phantom provides accurate anatomical simulation of the radioactive distribution in the human brain [43]. The phantom consist of 95 thin plastic slices that are

carved to simulate a 4:1 ratio of radioactivity between grey and white matter. The slices are packed in a plastic cylinder and filled with the radioactive solution. Thus the Hoffman phantom does not provide a realistic head contour nor attenuation effect of the skull, but it provides realistic contrast between the grey and white matter, and it is widely applied in the performance evaluation of PET scanners and reconstruction algorithms. In study I the Hoffman phantom was filled with ^{18}F-FDG in water solution, having a starting activity of 35 MBq. The relatively low starting activity resulted in a reasonable scanner dead-time (range 6.5%-5.8%) and randoms-to-trues ratio (range 0.147-0.132), during the 70 minutes scanning time.

The Striatum phantom provides realistic simulation of the anatomy of the striatum [60]. The phantom consists of two main parts, the head and the brain. The head part provides a realistic head contour and attenuation from the skull, and it can be opened through a calvarial cut to place the brain part inside. The brain part has five compartments that can be filled separately; left and right nucleus caudate, left and right putamen, and the remainder of the brain. In study I the striatum to background ratio was approximately 5:1, while the total amount of activity in the ^{18}F-FDG solution was 21 MBq at scan start. The scan duration was 57 minutes (corresponding to a typical ^{11}C-raclopride bolus scan).

The Iida phantom provides accurate anatomical simulation of the radioactivity distribution in the human brain, that is based on digital model of a real human MRI scan [47]. The digital model was realized using a 3D-printing technique. 3D-printing technique allowed generation of separate compartments for the radioactivity, skull and air cavities within the head region. The skull compartment was filled with K_2HPO_4 solution to simulate the attenuation effect of the skull, while in study III the radioactivity compartment was filled with ^{18}F-FDG solution

with approximately 69 MBq of activity at the scan start. The scan duration was 20 minutes.

Purpose of study I was to investigate the linearity of the radioactivity measurement at varying count statistics using the OP-OSEM reconstruction. For this, the state-of-the-art correction algorithms were employed for scatter, randoms and attenuation correction, while the varying count statistics were simulated using increasing number of NEC in overlapping frames (c.f. Chapter 2.2.2).

The list-mode data sub-sampling resulted in NEC range of 6×10^4 to 3×10^8 in the Hoffman phantom experiment and 1×10^5 to 3×10^8 NEC range in the Striatum phantom experiment.

Reconstructions at varying count statistics were made using the speed-optimized version of the OP-OSEM algorithm [45], with 16 subsets, and 11 full iterations, based on the initial convergence experiment with high count statistics data and literature [24].

Evaluation of the reconstruction accuracy was made within anatomically determined ROIs using remainder relative to the internal reference as a figure-of-merit (FOM):

$$REM_k = \frac{\text{mean}(ROI_k, \hat{F})/T_{\hat{F}}}{\text{mean}(ROI_k, F)/T_F}, \tag{4.1}$$

where $\hat{F}$ is any estimate image, F is the reference (high statistics) image and T_F is the total trues in frame F. Furthermore, regional contrast factors were calculated as

$$Q_k(F) = \frac{\text{mean}(ROI_k, F)}{\text{mean}(\partial ROI_k, F)}\,, \tag{4.2}$$

where $\text{mean}(R, F)$ is the mean pixel value inside region R and ∂ROI_k is a two to three pixels wide edge of the region ROI_k.

Threshold of $|REM_k - 1| < 0.025$ was used to determine the minimum NEC to obtain reasonably low bias. The possible

correlation between regional contrast and bias was explored using Pearson product moment correlation coefficient between the miminum NEC and Q_k over the regions.

In study III the full Iida phantom acquisition of 20 minutes (2.3×10^8 NEC) was subdivided into 28 samples of prefixed NEC (8.1×10^6). Image reconstructions were obtained using OP-OSEM, RM-OP-OSEM and 3DRP with bicubic transradial interpolation gap-filling algorithms as described in Chapter 2.2.4. Input to the iterative reconstruction was separate prompts, randoms and scatter sinograms, whereas, input to the analytical reconstruction was a precorrected, gap-filled trues sinogram. Analytical 3DRP reconstruction was made with Hamming filter with a transaxial and axial cuf-off frequency of 1 cycle per voxel. For iterative methods the number of iterations were 1-50 when 16 subsets were employed, and 1-250 without subsets. All image data were post-smoothed using a 2.5 mm FWHM 3D Gaussian kernel size.

Evaluation of the reconstructions were made within anatomically determined ROIs using either internal reference or the analytical reconstruction as a reference. The voxel values were scaled with the total counts in each image prior to ROI evaluation, and the bias was expressed as $100\% \times (ROI_{cur} - ROI_{ref})/ROI_{ref}$. Furthermore, the replicate design allowed determination of the true SNR through

$$SNR = \frac{\text{mean}(AC)}{\text{std}(AC)}, \tag{4.3}$$

where mean(AC) is the activity concentration average and std(AC) the standard deviation over the replicates.

Human experiments

The human data were collected for evaluation of the short-term repeatability and reliability of ^{11}C-Raclopride measurement

using the HRRT. The dataset consisted of two ^{11}C-Raclopride scans of each subject during the same day at least 2.5 hours apart. Seven healthy male subjects free of any somatic or psychiatric illness were recruited for the study. All were right handed and non-smokers. To exclude any structural brain abnormalities and to obtain anatomic reference, all subjects underwent 1.5T MRI. The age, height, and weight of the subjects were 24.5±3.5 years, 185.5±12.5cm, and 74±14kg, respectively (mean±s.d.).

Each PET scan was preceded by a transmission scan with a moving ^{137}Cs point source for attenuation correction calculation. PET scanning and motion tracking (MT) were started at the time of injection, and continued for 55 minutes. PET data were collected in event-by-event list-mode format, and the MT data was collected using an external infra-red position sensor, the Polaris Vicra system (NDI Medical, Waterloo, Canada).

Each subject was accommodated with an individually shaped thermoplastic mask to minimize the head motion during PET scanning. With regard to the head motion the data were of high quality, as indicated by small within and between frame motion, c.f. publication II. In addition, the injected doses and masses did not show any statistically significant differences between the scans.

Purpose of study II was to compare the image reconstructions obtained using the OP-OSEM and conventional framing to those obtained using RM-OP-OSEM algorithm and those obtained with fixed NEC statistics framing.

The OP-OSEM reconstructions were made using 16 subsets and 8 iterations, while the RM-OP-OSEM reconstruction were made using 16 subsets and 10 iterations.

The conventional framing from the literature was based on the kinetic profile of the ^{11}C-Raclopride TAC, that is, there was more dense framing in the initial part of the scan to depict the peak in radiotracer uptake, while the decaying part had more

sparse framing to accommodate for decline of the radioactivity. The conventional framing was 2×30, 9×60, 3×120, 3×180, and 6×300 seconds, resulting in altogether 23 frames.

The NEC-based framing was introduced in study II for retrospective sampling of the emission list-mode data according to a predefined NEC-threshold. In the conventional framing the time points were fixed while the NEC was let to vary, while in the NEC-based framing the count statistics were constant but the time points may vary. It was conceived that variability in the framing sequence may introduce bias in the SRTM-based kinetic analysis, whence, it was deemed necessary to validate the NEC-based framing through the test-retest performance evaluation.

The NEC-threshold was set at 7×10^6 on the basis of experiments in study I, yielding 10 to 17 frames per scan, depending on the injected dose and subject weight. Furthermore, parts of each scan were sampled into short duration frames yielding 5×10^5 to 2×10^6 NECs in 268 frames. The purpose of the low statistics framing was to explore the low statistics bias in real human data.

Reconstructions were evaluated using manually delineated anatomical ROIs as described in publication II. Primary evaluation was made with regard to the test-retest performance regionally. The test-retest performance was quantified using test-retest variability (TRV) for intra-subject variability and intraclass-correlation coefficient (ICC) for reliability, calculated as:

$$TRV = \frac{2|BP_t - BP_{rt}|}{BP_t + BP_{rt}} \tag{4.4}$$

where $BP_{t/rt}$ stand for the test and retest estimate of the measurement outcome within subject;

$$ICC = \frac{MS_B - MS_W}{MS_B + (k-1)MS_W} \tag{4.5}$$

where MS_W indicates the within, and MS_B the between subjects sum of squares of the difference in the outcome measure and $k = 2$ [84]. Moreover, the range of $\Delta BP = 100\% \times (BP_{rt} - BP_t)/BP_t$ and the arithmetic mean of the outcome measures in test and retest ± SD were reported.

In the secondary analysis the reconstructions were evaluated according to the area-under the curve (AUC) of the TAC data, relative to the AUC of an internal reference. The relative difference in the AUC was defined as

$$\Delta AUC(\%) = 100\% \frac{\int_{t_1}^{t_2} f(t)dt - \int_{t_1}^{t_2} f_0(t)dt}{\int_{t_1}^{t_2} f_0(t)dt}$$

where $f_0(t)$ is the TAC of the internal reference method, in this case the OP-OSEM reconstruction of the conventional framing.

Aim in study IV was to evaluate the transradial bicubic interpolation method and the analytical 3DRP reconstruction relative to the iterative reconstruction methods using real human data.

To investigate the impact of low count statistics in study IV, the conventional framing defined above was employed as a reference, and additional framing was obtained using short-duration frames by splitting each frame in the conventional framing into two, thus the short-duration framing included 46 time points.

The iterative reconstructions employed 16 subsets, with 12 full iterations for OP-OSEM and 16 full iterations for RM-OP-OSEM.

In study IV reconstructions were evaluated using the same manually delineated ROIs as in publication II. Primary evaluation was made with regard to the $\Delta AUC(\%)$, where the internal reference was obtained from the conventional framing using each of the reconstruction methods. In addition, the impact of framing to the $\mathrm{BP_{ND}}$ was investigated using

$$\Delta BP_{framing}(\%) = 100\% \frac{BP_{short} - BP_{long}}{BP_{long}} \tag{4.6}$$

The test-retest reliability was reported using ICC for long- and short-duration framings.

Purpose of study V was to evaluate the manual and automated methods for ROI-delineation in ^{11}C-raclopride studies. The human test-retest data firstly presented in study II was re-processed to maintain the best achievable spatial resolution, with regard to the image reconstruction as well as motion correction. Manual segmentation of the striatum was performed by two operators, one with more experience and other with less experience but instructed by the first; a T1-weighted MRI fused to PET sum image was used as a reference. The anatomical landmarks described by Mawlawi and colleagues [66, 67] were applied in the ROI delineation.

The individualizing mappings required in the atlas individualization were obtained using the FSL (version 5.0.8, FMRIB, Oxford, UK) and SPM (version 8, Wellcome Trust Centre for Neuroimaging, London, UK) software packages. The atlas of Fischl and colleagues was individualized within the FreeSurfer software (version 5.3.0, Martinos Center for Biomedical Imaging, Charlestown, Massachusetts, USA). The direct PET image clustering algorithm previously described in e.g. [36] was revised for this work through inclusion of Markov random field (MRF)-based resolution modelling in the initial striatum extraction step and inclusion of geometrically defined initialization of the clusters.

Primary analysis of the method performance was made on the basis of the test-retest characteristics. Test-retest characteristics of the ROI methods were assessed using SRTM kinetic analysis yielding the BP_{ND}, where the CERC TAC was used as a reference. To limit the number of comparisons, one of the CERC delineation methods was chosen as reference method. In addition to the ICC and TRV the standard error of measurement (SEM) was reported. SEM was calculated as $SEM = SD\sqrt{1 - ICC}$,

where SD was the standard deviation of all samples.

A secondary, more descriptive analysis was made relative to manual segmentation. That is, the manual delineations made by the first operator were used as a reference to which the other ROI-methods were compared. Comparisons were made at the voxel-level by measuring the similarity of the voxel-sets within the striatum (whole) (STR), LSTR, ASTR and SMST ROIs using the Jaccard coefficient:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \tag{4.7}$$

where A denotes the voxel-set of manual segmentation and B the voxel-set of the other method. Furthermore, agreement between the binding parameters was assessed using the limits of agreement (LOA) [14]. LOA is closely related to the Bland-Altman plot which can be used to assess the outcome of two methods that both contain methodological errors. The LOA was expressed as $\bar{d} \pm 1.96SD$, where $\bar{d}$ is the average of the differences between the two methods and SD is their standard deviation.

5 Results and discussion

Quantitative accuracy of OP-OSEM algorithm for high-resolution PET image reconstruction

Graphs in Figure 5.1 illustrate the relative regional accuracy (REM_k) as a function of NEC in Hoffman and Striatum phantom experiments (study I), in chosen brain regions. The graphs illustrate a two-way pattern; in the high activity ('hot') regions the activity was underestimated and in the low activity ('cold') regions it was overestimated. There was a nearly linear correlation between log(NEC) and REM_k, indicating a strong impact of the count statistics in the quantitative accuracy of the OP-OSEM reconstruction, and limit of $|REM_k - 1| < 0.025$ was subsequently used to find minimum NEC statistics to obtain an endurable error level of 2.5%.

In the Hoffman phantom experiment there was some variance in the minimum NEC requirements in the gray matter ROIs, including one outlier observation: the caudate nucleus. For the other gray matter ROIs the average requirement was 4.6×10^6 NECs. In the Striatum phantom experiment the average mimimum NEC requirement in the gray matter regions was somewhat higher (7.4×10^6) as compared to the Hoffman phantom experiment, while the cold background region showed rather low NEC requirement.

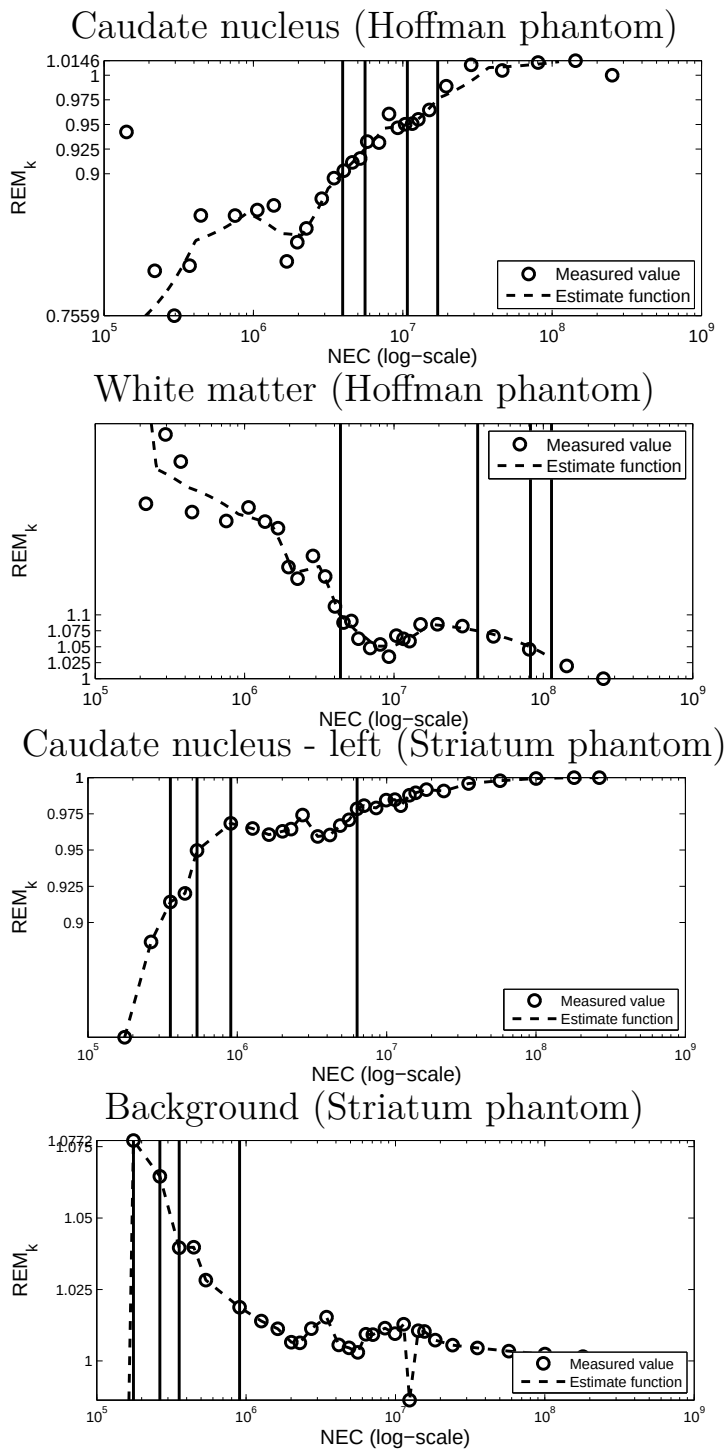


Figure 5.1: Phantom experiments (study I). Graphs represent the bias relative to the complete data (REM_k) as a function of the count statistics in chosen regions. Vertical lines correspond to minimum NEC statistics to obtain $|REM_k - 1| < 0.1, 0.075, 0.05, 0.025$.

To further investigate the relation between reconstruction bias and the image contrast the regional contrast factors $Q_k(F)$ were calculated and correlated with the minimum NEC requirements regionally. After removing the outlying observation of the caudate nucleus the contrast factors and the minimum NEC requirements in the gray matter regions showed a significant positive linear correlation in the Hoffman phantom experiment (Pearson's product moment correlation R=0.73, p=0.005), thus suggesting a strong dependence between the local image contrast and susceptibility to reconstruction bias.

The main findings in study I were that the previously recommended OP-OSEM algorithm for low count statistics HRRT reconstruction was not completely bias-free and that the low count statistics bias was related to the local image contrast. The bias seen in study I was of similar magnitude with that seen by van Velden and colleagues in [98].

In contrast to previous studies the NEC-based analysis presented in study I provided applicable threshold for retrospective framing of the HRRT PET data that can be effectively used to protect against the reconstruction bias. Furthermore, the contrast analysis implicated a problem in the convergence of the iterative reconstruction in low count statistics that may be reversible through modification of e.g. the number of iterations and/or subsets. These two implications were investigated further in studies II and III.

Firstly the impact of using fixed NEC-statistics to the outcome of pharmacokinetic modelling was explored using the ^{11}C-raclopride test-retest human dataset in study II. Safe NEC-statistics of 7×10^6 was estimated from the Striatum phantom experiment in study I that better resembles the activity distribution in ^{11}C-raclopride scans. The detrimental impact of fixed NEC-statistics framing to the temporal resolution in ^{11}C-raclopride scans was illustrated in study I, whence it was deemed

necessary to explore the effect of the proposed framing method to the outcome of pharmacokinetic modelling in real human data. In addition, the low count statistics bias was explored further using real human data in study II.

Figure 5.2 illustrates typical TAC-data obtained from a human scan using the tested reconstruction schemes. The fixed NEC high statistics framing yielded almost identical AUC in comparison with the conventional framing with the mean ΔAUC(%)<1.5% in all regions, implying a tolerable impact from degenerated temporal resolution. Whereas, the low statistics framing showed smaller values in all striatal regions, while the cerebellar values were similar, indicating a similar reconstruction bias in real human data as predicted from the phantom experiments.

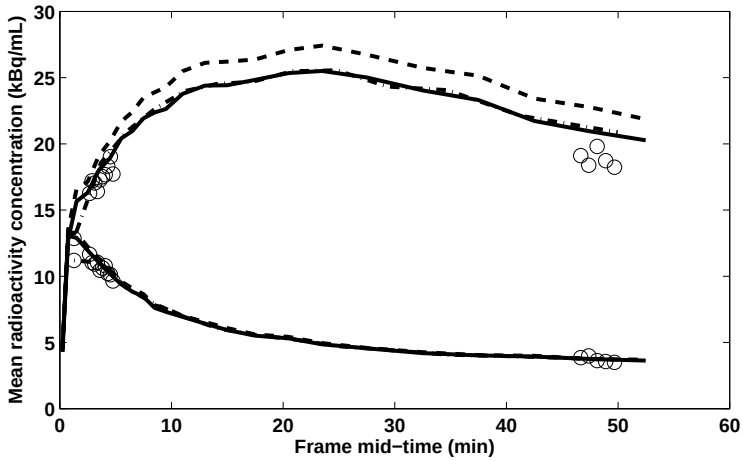


Figure 5.2: Human test-retest data (study II). A typical example of regional TAC data in the putamen and in the cerebellum estimated by the different framings and the OP-OSEM and RM-OP-OSEM reconstruction. Solid line is for original framing with OP-OSEM reconstruction, dash-dot line is for high statistics framing and OP-OSEM, dashed line is for RM-OP-OSEM reconstruction, and circles are for low statistics framing and OP-OSEM.

According to the analysis in study II there was a good agreement of the average BP_{ND} obtained from the OP-OSEM recon-

struction with conventional or NEC-based framings, and the TRV(%) and ICC were in the same rank.

Good agreement between the conventional and NEC-based framing was interpreted to provide cross-validation of the methods; the results suggested that the conventional framing was not associated with significant reconstruction bias with regard to BP_{ND} calculation and conversely the poorer temporal resolution due to fixed NEC-statistics framing did not yield significant reduction in the accuracy of the pharmacokinetic modelling. On the basis of these findings a novel, less dense framing scheme was proposed for the ^{11}C-raclopride HRRT studies that better accommodate the varying count statistics during the PET scanning without interfering with the pharmacokinetic modelling.

Quantitative accuracy of RM-OP-OSEM and 3DRP with transradial bicubic interpolation gap-filling algorithms for high-resolution PET image reconstruction

The solution of less dense framing was regarded applicable in resting state ^{11}C-raclopride PET, but in order to allow more temporal flexibility to accommodate for instance lp-ntPET modelling a bias-free reconstruction algorithm was required.

The contrast dependence of the reconstruction bias using OP-OSEM seen in study I suggested restricted convergence in low count statistics that may have been due to too early suspension of the iterations or due to divergence of the iterations in the presence of noise. These two implications were further investigated in studies III and IV.

Firstly, convergence of the OP-OSEM algorithm in high and low count statistics was further investigated in study III. Experiments in study III showed that the convergence in low statistics stopped much earlier as compared to high statistics in the OP-OSEM iteration and the activity concentrations remained at a lower level in high contrast regions as the iterations

proceeded (c.f. Figure 5.3 top). Furthermore, the behaviour was the same independent of the number of subsets, thus modifying the number of iterations or subsets did not help to alleviate the reconstruction bias. In basal ganglio ROI the bias to cerebellar ROI was approximately -3%, in medial frontal cortex ROI approximately -4% and in cingulate cortex ROI approximately -5%, thus implying a sizeable quantification non-linearity depending on the count statistics. The NEC-level employed in study III was not extraordinary low, indicating a risk of biased outcome in a neurotransmitter release assessment if OP-OSEM algorithm was applied.

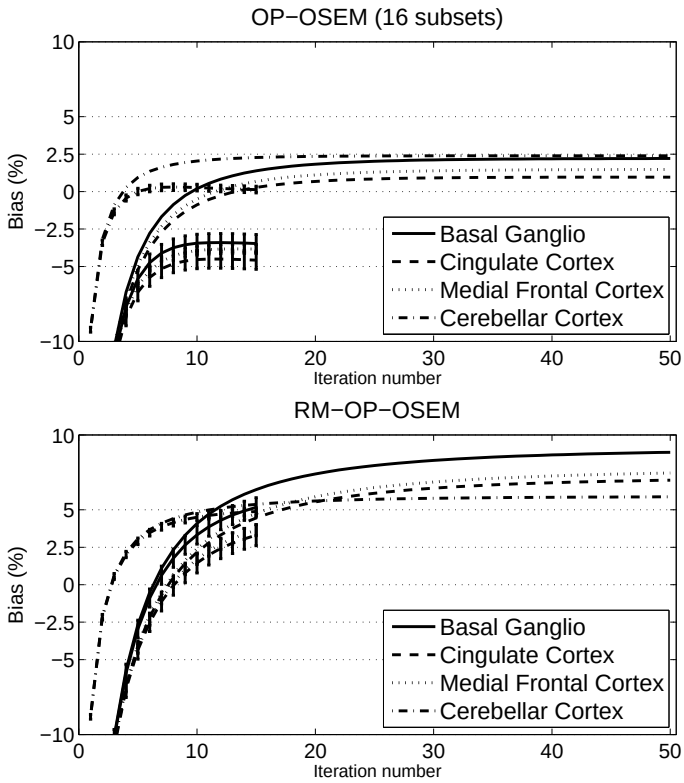


Figure 5.3: Bias in OP-OSEM and RM-OP-OSEM reconstruction with high and low statistics data. The graphs show regional difference relative to the high statistics 3DRP reconstruction as a function of iterations. The low statistics data are presented as mean±SD over the replicates.

A common shortcoming of generic MLEM-based iterative reconstruction is the lack of spatial connectivity between the neighbouring voxels, while in the living tissue the radioactivity concentration can be assumed fairly homogeneous in a millimeter-scale neighbourhood. In the case of the HRRT data this caveat has a stronger impact as compared to conventional scanners due to smaller volume of the LORs and image voxels, and thus less signal per measurement and image element. It is a viable supposition that the iterative reconstruction diverges in low count statistics due to poor SNR per voxel, but increasing the true covariation between voxels might help to reverse this unwanted behaviour.

Covariation between neighbouring voxels can be achieved by introduction of explicit or implicit spatial penalization. An example of explicit spatial penalization is the median root prior (MRP)-algorithm introduced by Alenius and Ruotsalainen [4], that might serve well for the current task, but was not included in the experiments due to lack of HRRT-specific implementation. In contrast, the HRRT-specific PSF-kernel implemented in the RM-OP-OSEM algorithm can be regarded as an implicit spatial penalization. The RM-OP-OSEM algorithm was primarily developed to alleviate the PVE in the HRRT measurement, but while doing so it also increases the voxel-wise covariation.

The benefit of using the RM-OP-OSEM algorithm as compared to OP-OSEM algorithm in low count statistics have been earlier shown by Walker and colleagues [105], but some bias was still visible. Thus, it was regarded that further investigation of the RM-OP-OSEM algorithm was required with regard to its performance in low count statistics as well as its noise characteristics. Feasibility of the RM-OP-OSEM algorithm in human studies was earlier shown by Sureau and colleagues [88] and Varrone and colleagues [102] using two HRRT scanners, but its

impact on the repeatability and reliability of the ^{11}C-raclopride assessment had not been reported.

In study II the test-retest characteristics of RM-OP-OSEM were investigated using the conventional framing. In study II we observed increased image contrast with RM-OP-OSEM that was associated with higher BP_{ND} estimates in the striatum, but not necessarily improved test-retest repeatability or reliability. In contrast, the test-retest characteristics of RM-OP-OSEM algorithm were mainly inferior to those obtained using OP-OSEM.

On the other hand, in study IV the same human data was employed to investigate the impact of low count statistics and it was found that the bias in RM-OP-OSEM was approximately one-third of that in the OP-OSEM reconstruction, indicating a benefit of using RM, similar to that seen by Walker and colleagues [105].

The characteristics of the RM-OP-OSEM algorithm were further investigated in a phantom experiment in study III. The phantom experiment showed non-linearities in the RM-OP-OSEM reconstruction with regard to the count statistics, but in slightly smaller magnitude as compared to the OP-OSEM, in good concordance with the earlier findings by others [88, 105].

A possibly more intriguing facet of the findings in study III was the non-negligible change of the iterates between a commonly suggested 15 iterations [88] and 50 iterations, with a non-zero slope of the tangent of the bias curve (c.f. Figure 5.3 bottom).

It is commonly acknowledged that the RM-OP-OSEM requires more iterations to converge [88], but our findings suggest that convergence might not be reached even after 50 iterations (with 16 subsets). On the other hand there was a regime close to 15 iterations where the high and low count statistics data showed smallest difference between each other (c.f. Figure 5.3 bottom), but the regime might be dependent on the phantom design and the tested NEC-levels. The NEC-level employed in

study III was not extraordinary low.

Over-correction in thin high contrast territories with RM-OP-OSEM (see Figure 2 column C in Publication III for edge artefact) have been observed by other researchers too. Stute and Comtat [87] have suggested employment of high number of iterations followed by spatial smoothing with a 3D Gaussian kernel size matching with the scanner intrinsic resolution to solve the edge artefact without loosing the high frequency information. In their experiments they employed the method for simulated data with known PSF, while in study III we employed true HRRT data and estimated PSF.

In our experiments post-smoothing using a 2.5 mm FWHM 3D Gaussian kernel did not change the behaviour of RM-OP-OSEM within 50 iterations, although the magnitude of bias decreased from approximately 11% to approximately 8% in the basal ganglia ROI (data not shown). Thus, albeit the RM-OP-OSEM algorithm might provide relative alleviation to the low count statistics bias issue, it did not improve the test-retest characteristics of ^{11}C-Raclopride assessment (study II) and its convergence can not be guaranteed, in contrary, there was a sizeable edge artefact effect in a phantom study (study III).

Thus, despite the relative advantage of RM-OP-OSEM over OP-OSEM in low count statistics its usage in lp-ntPET studies can not be recommended without caution.

In contrast, the analytical reconstruction implemented using transradial bicubic interpolation method for gap-filling and 3DRP algorithm showed virtually no bias in our experiments.

In study III the low count statistics replicates yielded identical outcome in average relative to the high count statistics reference and the experiments in study IV with half duration frames showed zero bias in 3DRP reconstruction, while the iterative methods showed biases of up to 3% in BP_{ND} estimates. Thus within the tested NEC-statistics analytical reconstruction was

bias-free while the iterative methods showed significant over- and underestimations of the activity concentrations, indicating that with regard to reconstruction bias usage of analytical algorithm is recommended over the iterative algorithms, and should thus be recommended in lp-ntPET studies.

Signal-to-noise ratio considerations

Quantitative accuracy is an essential prerequisite in pharmacokinetic modelling but the temporal resolution may be expendable on behalf of enhanced SNR. In particular, in the lp-ntPET analysis the SNR plays a central role in order to robustly detect the subtle changes in the PET signal. On the other hand, without prior knowledge about the time course of the fluctuations due to e.g. cognitive stimulation the best achievable temporal resolution should usually be pursued.

It is commonly thought that the iterative reconstruction algorithms would yield significantly better SNR as compared to analytical reconstruction due to more accurate modelling of the statistical variation in PET data.

For the HRRT van Velden and colleagues have employed homogeneous cylinder phantom and Hoffman phantom in a replicate design to assess the true variation in the ROI mean, and found that the 3DRP reconstruction was associated with approximately two-fold higher noise level than the OP-OSEM reconstruction [100]. However, the high noise level might be at least partly attributed to the CFS gap-filling method that has been shown more sensitive to noise than the DCT-domain and bicubic transradial interpolation gap-filling methods (c.f. phantom experiment in study IV).

We tested the SNR behaviour of iterative and analytical reconstructions in study III using the Iida phantom in a replicate design. In our experiments the analytical reconstruction with 3DRP was preceded by bicubic transradial interpolation gap-filling instead of CFS gap-filling.

Results from the SNR experiment are shown in Figure 5.4. Graphs in Figure 5.4 indicate that the SNR decreases while the iterations proceed, although the initial peak in SNR should be ignored, as it mostly reflects the uniformity of the initial guess. It can be seen that the iterative algorithms do not necessarily produce better SNR as compared to the analytical algorithm if convergence of the iterative reconstruction is required. There were also differences between the iterative methods; RM-OP-OSEM yielded systematically lowest SNR while the OP-OSEM reconstruction yielded slightly higher SNR with 1 subset than with 16 subsets. Visualization of the voxel-by-voxel SNR can be found in Figure 2 last row in Publication III.

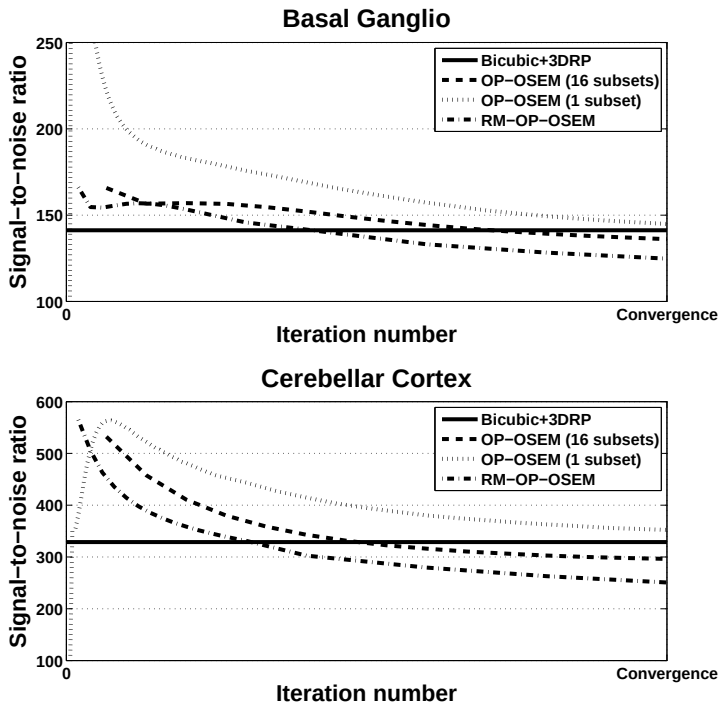


Figure 5.4: Iida phantom (study III). Graphs represent the SNR behaviour of various HRRT reconstruction schemes as calculated over the count-corrected samples. See legend for the applied reconstruction methods. Iteration number was scaled between 0 and convergence, to allow simultaneous presentation of the different methods.

Thus, according to the phantom experiments in study III the 3DRP reconstruction with transradial bicubic interpolation based gap-filling can provide comparable SNR with iterative reconstruction at a NEC-level that may introduce bias in the iterative reconstruction. Furthermore, experiments in study IV with half duration frames showed zero bias in 3DRP reconstruction, while the iterative methods showed biases up to 3% in BP_{ND} estimates. While, the regional ICCs of the 3DRP reconstruction were slightly superior to those of the iterative reconstruction algorithms.

Based on the current experiments usage of 3DRP and transradial bicubic interpolation for gap-filling can be recommended in nt-PET studies with high requirements for temporal resolution. In the current work employment of a small apodization filter size in 3DRP reconstruction resulted in similar reconstructed spatial resolutions (in particular after matched post-smoothing using 2.5 mm 3D Gaussian filter kernel) between analytical and iterative reconstructions, with only slight underestimation as compared to OP-OSEM outcome (c.f. Fig. 5.3).

Feasibility of novel automated ROI-delineation methods in the human striatum in comparison to current manual ROI-delineation

Experiments in study V using various ROI-delineation methods in the human striatum and cerebellum showed feasibility of all the tested methods but also important differences. The manual ROI method was associated with considerable inter-operator variability while the automated methods suffered from increased PVE of varying degree. It was however deemed that some of the automated methods were associated with lesser methodological variation as compared to manual method, suggesting a benefit in e.g. group level comparisons if automated methods were used. Furthermore, flexibility with regard to image normalization

procedure in the atlas-based approach was shown.

High inter-operator variability in the manual ROI-delineation was indicated at the voxel-by-voxel level by low Jaccard coefficient (0.46-0.61). Furthermore, there was a statistically significant difference between the outcome measures (BP_{ND}) as indicated by LOA range in ASTR and SMST ROIs, implicating that the ROI-delineations from many operators can not be employed without a risk of exaggerated variability and/or systematic bias in individual BP_{ND} estimates. Moreover, the manual ROIs were associated with systematically higher CoV(%) of both the ROI volume and BP_{ND} as compared to some automated methods suggesting a possibly exaggerated variability between subjects (or attenuated variability using automated methods).

Among the atlas-based methods the structural atlases were lacking the subdivision between the anterior and posterior parts of caudate and putamen, thus lacking specificity between the associative and sensorimotor striatum. Considering the disparate functionalities of ASTR and SMST the definition in structural atlases was deemed inadequate for high-resolution ^{11}C-raclopride PET studies.

In contrast, the novel CB atlas contained 7 functional subregions in both striata instead of three conventionally described in manual ROI-delineation. For direct comparison with manual delineation a combination strategy was employed to yield 3 subregions while the original subregions were analysed independently.

With regard to the test-retest characteristics the combined CB-atlas ROIs showed best performance. In particular, the ROIs generated using PET-based normalization showed smallest average TRV(%) (4.47%) and average SEM (0.15), somewhat superior to those obtained using MRI-based normalization. These differences may be attributed to the limitations of MRI-based

normalization specifically in the striatum, and others have also suggested PET-based normalization in the striatum [59].

In turn, the direct PET image clustering algorithm showed poor performance in the presence of noise despite good overall performance. The algorithm failed in a single scan with extraordinarily low SNR, thus in case the image SNR can be guaranteed through for instance averaging of many BP_{ND} images the direct PET clustering method may provide a viable alternative.

Albeit the unwanted methodological variation can not be fully dissected from the true biological variation it was deemed likely that the larger between subject variability in manual ROI-delineation was more due to random differences in the manual ROIs between subjects than due to increased PVE associated with the automated methods.

At least according to statistical power calculations that takes into account the mean±SD of BP_{ND} showed benefit of using CB-atlas with PET-based normalization over the manual ROI-delineation. A power calculation software G*Power (version 3.1.9.2, Universität Kiel, Germany) was used to calculate the minimum sample sizes to detect a 10% change in BP_{ND} in a one-tailed paired t-test (with $\alpha = 0.05$ and power=0.95). The results showed that in the LSTR the minimum sample size using CB-atlas was 10 while using manual ROI-delineation it was as much as 25. Thus, showing a remarkable improvement in the sensitivity by using the automated method.

Furthermore, the CB subdivision has been shown to provide more functionally relevant ROI-delineation than the structure based delineation using e.g. manual drawing [97]. Smallest of the 7 subregions from CB-atlas approach however showed markedly poorer test-retest characteristics and their independent usage in ROI-analysis may be critisized. Nevertheless, the most important ROIs may be better protected from nuisance

components if the 7 subregions are segmented independently.

The reference region ROI-delineations in the cerebellum were evaluated using the test-retest characteristics of the resulting BP_{ND} estimates, but also of cerebellar distribution volume (DV) derived with the help of arterial plasma input. DV provided an independent and quantitative assessment of the ROI-method accuracy. The DV estimates from all the tested methods were within 5% from each other showing good agreement, while the TRV(%) and SEM were slightly superior using the automated methods in comparison to manual drawings.

ROI volumes from automated methods were markedly larger as compared to manual drawings suggesting a benefit with regard to SNR in the reference region TAC when using automated methods. Improvement in SNR may have contributed to the improved test-retest characteristics of DV using the automated methods. Moreover, the test-retest characteristics of BP_{ND} were slightly superior using automated methods further establishing a benefit from improved SNR.

In conclusion, the automated methods based on atlas individualizing approach using either PET- or MRI-based normalization can provide operator-independent and replicable ROI-delineations with improved test-retest characteristics and sensitivity to detect group-level changes in BP_{ND}. Thus usage of fully automated ROI-generation can be recommended in the analysis of high-resolution ^{11}C-raclopride data.

6 Implications

In this work high-resolution PET image reconstruction and ROI-delineation strategies were evaluated with regard to requirements in PET neuroreceptor binding assay. The particular demands in the implementation of cognitive/behavioral activation studies have set the basis for this work.

Depiction and subsequent pharmacokinetic modelling of rapid fluctuations in PET signal due to e.g. cognitive stimulation require high temporal sampling. On the other hand, stimulation of the neurotransmitter system may induce the relevant alterations within small functionally specialized brain regions that are beyond the spatial resolution of conventional PET scanners. Thus the best achievable temporal and spatial resolution are simultaneously required to be able to detect the most subtle alterations in the PET signal.

Poisson characteristics of the PET signal sets serious limitations with regard to spatiotemporal resolution of PET. In particular, the current state-of-the-art high-resolution human PET imaging provided by the HRRT-scanner has been associated with significant limitations with regard to temporal resolution.

In this work it was repeatedly shown that the best temporal resolution achievable using the commonly recommended iterative reconstruction method was relatively poor, and the performance was not significantly improved through modifications of the algorithm. In turn, a recently introduced gap-filling strategy based on transradial bicubic interpolation in tandem

with 3D analytical reconstruction showed markedly improved performance with regard to quantitative accuracy and similar performance with regard to noise contribution, suggesting a true benefit from using this novel approach in neurotransmission PET studies.

Due to the long time-span of this work the initial solution using the iterative reconstruction algorithm and a predefined noise equivalent counts (NEC)-threshold to guarantee quantitative accuracy have been employed in the analysis of a cognitive stimulation PET study at Turku PET Centre. Despite the poor temporal resolution of the approach a significant stimulation effect was detected from the PET signal using lp-ntPET type modelling and the findings related to training effect were published in the prestige Science-journal in 2011 [7]. In this study the late onset of the activation task relative to the ^{11}C-raclopride injection (55 minutes), however, resulted in temporal resolution of approximately 9 minutes.

According to literature temporal resolution of 9 minutes is clearly suboptimal in case of cognitive stimulation that is associated with learning and habituation effects. Thus it can be regarded that the assessments achieved in [7] may have been somewhat attenuated although statistically significant changes were found. A viable supposition is that the improved temporal resolution as provided by the analytical reconstruction algorithm provide significant advantage in the measurement of the rapid fluctuations in the PET signal in the future high-resolution PET studies.

Nevertheless, higher temporal resolution is always associated with deteriorated image SNR in PET. The commonly employed approach of ROI-analysis to improve SNR was optimized in this work for high-resolution DA PET studies. It was found that automated method based on atlas individualization can provide improved repeatability as compared to currently employed

manual ROI-delineation. In particular, in cognitive activation studies where the alteration in PET signal can be spatially very narrow the higher accuracy of atlas-based approach may prove beneficial, and even provide an alternative to voxel-by-voxel statistical analysis such as Statistical Parametric Mapping (SPM).

Feasibility of neurotransmission PET was improved through the achievements in this work. To date, PET offers the best approach for human in vivo neurotransmission assessment and demands for higher accuracy are ever-increasing. Neurotransmission PET holds great potential in the investigation of the pathophysiology of various brain disorders that is not currently fully exploited. Important obstacle on the way of more widespread usage of neurotransmission PET is its methodological complexity, and all improvements in its accessibility have a great impact in the investigation of human neurotransmission. It is wished that the observations made in this work would spark new more widespread interest in the application of neurotransmission PET in e.g. cognitive/behavioral stimulation studies, but also in pharmacological stimulations where the temporal drug response is of interest.

Bibliography

[1] Alakurtti, K., Aalto, S., Johansson, J. J., Någren, K., Tuokkola, T., Oikonen, V., Laine, M., and Rinne, J. O., "Reproducibility of striatal and thalamic dopamine d2 receptor binding using [11c]raclopride with high-resolution positron emission tomography." *J Cereb Blood Flow Metab*, vol. 31, no. 1, pp. 155–165, Jan 2011. [Online]. Available: http://dx.doi.org/10.1038/jcbfm.2010.64

[2] Alakurtti, K., Johansson, J. J., Tuokkola, T., Någren, K., and Rinne, J. O., "Rostrocaudal gradients of dopamine d2/3 receptor binding in striatal subregions measured with [(11)c]raclopride and high-resolution positron emission tomography." *Neuroimage*, vol. 82, pp. 252–259, Nov 2013. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2013.05.091

[3] Alakurtti, K., Johansson, J. J., Joutsa, J., Laine, M., Bäckman, L., Nyberg, L., and Rinne, J. O., "Long-term test-retest reliability of striatal and extrastriatal dopamine d2/3 receptor binding: study with [(11)c]raclopride and high-resolution pet." *J Cereb Blood Flow Metab*, vol. 35, no. 7, pp. 1199–1205, Jul 2015. [Online]. Available: http://dx.doi.org/10.1038/jcbfm.2015.53

[4] Alenius, S. and Ruotsalainen, U., "Bayesian image reconstruction for emission tomography based on median root

prior." *Eur J Nucl Med*, vol. 24, no. 3, pp. 258–265, Mar 1997.

[5] Alpert, N. M., Badgaiyan, R. D., Livni, E., and Fischman, A. J., "A novel method for noninvasive detection of neuromodulatory changes in specific neurotransmitter systems." *Neuroimage*, vol. 19, no. 3, pp. 1049–1060, Jul 2003.

[6] Azevedo, F. A. C., Carvalho, L. R. B., Grinberg, L. T., Farfel, J. M., Ferretti, R. E. L., Leite, R. E. P., Jacob Filho, W., Lent, R., and Herculano-Houzel, S., "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain." *J Comp Neurol*, vol. 513, no. 5, pp. 532–541, Apr 2009. [Online]. Available: http://dx.doi.org/10.1002/cne.21974

[7] Bäckman, L., Nyberg, L., Soveri, A., Johansson, J., Andersson, M., Dahlin, E., Neely, A. S., Virta, J., Laine, M., and Rinne, J. O., "Effects of working-memory training on striatal dopamine release." *Science*, vol. 333, no. 6043, p. 718, Aug 2011. [Online]. Available: http://dx.doi.org/10.1126/science.1204978

[8] Badgaiyan, R. D., Fischman, A. J., and Alpert, N. M., "Striatal dopamine release during unrewarded motor task in human volunteers." *Neuroreport*, vol. 14, no. 11, pp. 1421–1424, Aug 2003. [Online]. Available: http://dx.doi.org/10.1097/01.wnr.0000080021.91618.ee

[9] ———, "Striatal dopamine release in sequential learning." *Neuroimage*, vol. 38, no. 3, pp. 549–556, Nov 2007. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2007.07.052

[10] ———, "Explicit motor memory activates the striatal dopamine system." *Neuroreport*, vol. 19, no. 4,

pp. 409–412, Mar 2008. [Online]. Available: http://dx.doi.org/10.1097/WNR.0b013e3282f6435f

[11] Bataille, F., Comtat, C., Jan, S., and Trebossen, R., "Monte Carlo simulation for the ecat hrrt using gate," in *Nuclear Science Symposium Conference Record, 2004 IEEE*, vol. 4, 2004, pp. 2570–2574. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1462778

[12] Behrens, T. E. J., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A. M., Boulby, P. A., Barker, G. J., Sillery, E. L., Sheehan, K., Ciccarelli, O., Thompson, A. J., Brady, J. M., and Matthews, P. M., "Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging." *Nat Neurosci*, vol. 6, no. 7, pp. 750–757, Jul 2003. [Online]. Available: http://dx.doi.org/10.1038/nn1075

[13] Bettinardi, V., Presotto, L., Rapisarda, E., Picchio, M., Gianolli, L., and Gilardi, M. C., "Physical performance of the new hybrid pet-ct discovery-690." *Med Phys*, vol. 38, no. 10, pp. 5394–5411, Oct 2011. [Online]. Available: http://dx.doi.org/10.1118/1.3635220

[14] Bland, J. M. and Altman, D. G., "Statistical methods for assessing agreement between two methods of clinical measurement." *Lancet*, vol. 1, no. 8476, pp. 307–310, Feb 1986.

[15] Boellaard, R., Lubberink, M., de Jong, H., Kropholler, M., and Lammertsma, A., "Application of various iterative reconstruction for quantitative 3D dynamic brain PET studies," in *Nuclear Science Symposium Conference Record, 2004 IEEE*, vol. 4, 2004, pp. 2553–2556. [Online].

Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1462774

[16] Boellaard, R., "Standards for pet image acquisition and quantitative data analysis," *Journal of nuclear medicine*, vol. 50, no. Suppl 1, pp. 11S–20S, 2009.

[17] Brix, G., Zaers, J., Adam, L. E., Bellemann, M. E., Ostertag, H., Trojan, H., Haberkorn, U., Doll, J., Oberdorfer, F., and Lorenz, W. J., "Performance evaluation of a whole-body pet scanner using the nema protocol. national electrical manufacturers association." *J Nucl Med*, vol. 38, no. 10, pp. 1614–1623, Oct 1997.

[18] Byars, L., Sibomana, M., Burbar, Z., Jones, J., Panin, V., Barker, W., Liow, J.-S., Carson, R., and Michel, C., "Variance reduction on randoms from coincidence histograms for the hrrt," in *Nuclear Science Symposium Conference Record, 2005 IEEE*, vol. 5, 2005, pp. 2622–2626. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1596876

[19] Byrne, C., "Iterative algorithms for deblurring and deconvolution with constraints," *Inverse Problems*, vol. 14, no. 6, p. 1455, 1998.

[20] Carson, R., Barker, W., Liow, J.-S., and Johnson, C., "Design of a motion-compensation osem list-mode algorithm for resolution-recovery reconstruction for the hrrt," in *Nuclear Science Symposium Conference Record, 2003 IEEE*, vol. 5, 2003, pp. 3281–3285. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1352597

[21] Carson, R. E., Channing, M. A., Blasberg, R. G., Dunn, B. B., Cohen, R. M., Rice, K. C., and Herscovitch, P.,

"Comparison of bolus and infusion methods for receptor quantitation: application to [18f] cyclofoxy and positron emission tomography," *Journal of Cerebral Blood Flow & Metabolism*, vol. 13, no. 1, pp. 24–42, 1993.

[22] Carson, R. E., Breier, A., de Bartolomeis, A., Saunders, R. C., Su, T. P., Schmall, B., Der, M. G., Pickar, D., and Eckelman, W. C., "Quantification of amphetamine-induced changes in [11c] raclopride binding with continuous infusion," *Journal of Cerebral Blood Flow & Metabolism*, vol. 17, no. 4, pp. 437–447, 1997.

[23] Cherry, S. and Dahlbom, M., "Pet: Physics, instrumentation, and scanners," in *PET*, Phelps, M., Ed. Springer New York, 2006, pp. 1–117. [Online]. Available: http://dx.doi.org/10.1007/0-387-34946-4_1

[24] Comtat, C., Bataille, F., Michel, C., Jones, J., Sibomana, M., Janeiro, L., and Trebossen, R., "Osem-3D reconstruction strategies for the ecat hrrt," in *Nuclear Science Symposium Conference Record, 2004 IEEE*, vol. 6, 2004, pp. 3492–3496. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1466639

[25] Comtat, C., Sureau, F., Sibomana, M., Hong, I., Sjoholm, N., and Trebossen, R., "Image based resolution modeling for the hrrt osem reconstructions software," in *Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE*, 2008, pp. 4120–4123. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4774188

[26] Dahlbom, M., Schiepers, C., and Czernin, J., "Comparison of noise equivalent count rates and image noise," *IEEE Transactions on Nuclear Science*, vol. 52, no. 5, pp. 1386–1390, 2005. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1546425

[27] Daube-Witherspoon, M. E., Karp, J. S., Casey, M. E., DiFilippo, F. P., Hines, H., Muehllehner, G., Simcic, V., Stearns, C. W., Adam, L.-E., Kohlmyer, S. *et al.*, "Pet performance measurements using the nema nu 2-2001 standard," *Journal of Nuclear Medicine*, vol. 43, no. 10, pp. 1398–1409, 2002.

[28] de Jong, H. W. A. M., van Velden, F. H. P., Kloet, R. W., Buijs, F. L., Boellaard, R., and Lammertsma, A. A., "Performance evaluation of the ecat hrrt: an lso-lyso double layer high resolution, high sensitivity scanner." *Phys Med Biol*, vol. 52, no. 5, pp. 1505–1526, Mar 2007. [Online]. Available: http://dx.doi.org/10.1088/0031-9155/52/5/019

[29] de Jong, H., Boellaard, R., Knoess, C., Lenox, M., Michel, C., Casey, M., and Lammertsma, A., "Correction methods for missing data in sinograms of the hrrt PET scanner," *IEEE Transactions on Nuclear Science*, vol. 50, no. 5, pp. 1452–1456, 2003. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1236948

[30] del Campo, N., Tait, R. J., Acosta-Cabronero, J., Hong, Y. T., Izquierdo-Garcia, D., Smith, R., Aigbirhio, F. I., Sahakian, B. J., Müller, U., Robbins, T. W., and Fryer, T. D., "Quantification of receptor-ligand binding potential in sub-striatal domains using probabilistic and template regions of interest." *Neuroimage*, vol. 55, no. 1, pp. 101–112, Mar 2011. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2010.11.071

[31] Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., and Ramnani, N., "A probabilistic mr atlas of the human cerebellum." *Neuroimage*, vol. 46,

no. 1, pp. 39–46, May 2009. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2009.01.045

[32] Egerton, A., Mehta, M. A., Montgomery, A. J., Lappin, J. M., Howes, O. D., Reeves, S. J., Cunningham, V. J., and Grasby, P. M., "The dopaminergic basis of human behaviors: A review of molecular imaging studies." *Neurosci Biobehav Rev*, vol. 33, no. 7, pp. 1109–1132, Jul 2009. [Online]. Available: http://dx.doi.org/10.1016/j.neubiorev.2009.05.005

[33] Endres, C. J. and Carson, R. E., "Assessment of dynamic neurotransmitter changes with bolus or infusion delivery of neuroreceptor ligands." *J Cereb Blood Flow Metab*, vol. 18, no. 11, pp. 1196–1210, Nov 1998. [Online]. Available: http://dx.doi.org/10.1097/00004647-199811000-00006

[34] Farde, L., Hall, H., Ehrin, E., and Sedvall, G., "Quantitative analysis of d2 dopamine receptor binding in the living human brain by pet." *Science*, vol. 231, no. 4735, pp. 258–261, Jan 1986.

[35] Farde, L., Eriksson, L., Blomquist, G., and Halldin, C., "Kinetic analysis of central [11c]raclopride binding to d2-dopamine receptors studied by pet–a comparison to the equilibrium analysis." *J Cereb Blood Flow Metab*, vol. 9, no. 5, pp. 696–708, Oct 1989. [Online]. Available: http://dx.doi.org/10.1038/jcbfm.1989.98

[36] Farinha, R. J. P. C., Ruotsalainen, U., Hirvonen, J., Tuominen, L., Hietala, J., Fonseca, J. M., and Tohka, J., "Segmentation of striatal brain structures from high resolution pet images." *Int J Biomed Imaging*, vol. 2009, p. 156234, 2009. [Online]. Available: http://dx.doi.org/10.1155/2009/156234

[37] Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. M., "Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain." *Neuron*, vol. 33, no. 3, pp. 341–355, Jan 2002.

[38] Fisher, R. E., Morris, E. D., Alpert, N. M., and Fischman, A. J., "In vivo imaging of neuromodulatory synaptic transmission using pet: a review of relevant neurophysiology," *Human Brain Mapping*, vol. 3, no. 1, pp. 24–34, 1995.

[39] Goldman-Rakic, P. S., "Working memory dysfunction in schizophrenia." *The Journal of neuropsychiatry and clinical neurosciences*, 1994.

[40] Gunn, R. N., Lammertsma, A. A., Hume, S. P., and Cunningham, V. J., "Parametric imaging of ligand-receptor binding in pet using a simplified reference region model." *Neuroimage*, vol. 6, no. 4, pp. 279–287, Nov 1997. [Online]. Available: http://dx.doi.org/10.1006/nimg.1997.0303

[41] Haber, S. N., Fudge, J. L., and McFarland, N. R., "Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum," *The Journal of neuroscience*, vol. 20, no. 6, pp. 2369–2382, 2000.

[42] Hirvonen, J., Aalto, S., Lumme, V., Någren, K., Kajander, J., Vilkman, H., Hagelberg, N., Oikonen, V., and Hietala, J., "Measurement of striatal and thalamic dopamine d2 receptor binding with 11c-raclopride." *Nucl Med Commun*, vol. 24, no. 12,

pp. 1207–1214, Dec 2003. [Online]. Available: http://dx.doi.org/10.1097/01.mnm.0000104642.79626.e8

[43] Hoffman, E., Cutler, P., Digby, W., and Mazziotta, J., "3-D phantom to simulate cerebral blood flow and metabolic images for PET," *IEEE Transactions on Nuclear Science*, vol. 37, no. 2, pp. 616–620, 1990. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=106686

[44] Hofheinz, F., Pötzsch, C., Oehme, L., Beuthien-Baumann, B., Steinbach, J., Kotzerke, J., and van den Hoff, J., "Automatic volume delineation in oncological pet. evaluation of a dedicated software tool and comparison with manual delineation in clinical data sets." *Nuklearmedizin*, vol. 51, no. 1, pp. 9–16, 2012. [Online]. Available: http://dx.doi.org/10.3413/Nukmed-0419-11-07

[45] Hong, I., Chung, S., Kim, H., Kim, Y., Son, Y., and Cho, Z., "Ultra fast symmetry and SIMD-based projection-backprojection (ssp) algorithm for 3-D PET image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 26, no. 6, pp. 789–803, 2007. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4214878

[46] Hudson, H. M. and Larkin, R. S., "Accelerated image reconstruction using ordered subsets of projection data." *IEEE Trans Med Imaging*, vol. 13, no. 4, pp. 601–609, 1994. [Online]. Available: http://dx.doi.org/10.1109/42.363108

[47] Iida, H., Hori, Y., Ishida, K., Imabayashi, E., Matsuda, H., Takahashi, M., Maruno, H., Yamamoto, A., Koshino, K., Enmi, J., Iguchi, S., Moriguchi, T., Kawashima,

H., and Zeniya, T., "Three-dimensional brain phantom containing bone and grey matter structures with a realistic head contour." *Ann Nucl Med*, vol. 27, no. 1, pp. 25–36, Jan 2013. [Online]. Available: http://dx.doi.org/10.1007/s12149-012-0655-7

[48] Innis, R. B., Cunningham, V. J., Delforge, J., Fujita, M., Gjedde, A., Gunn, R. N., Holden, J., Houle, S., Huang, S.-C., Ichise, M., Iida, H., Ito, H., Kimura, Y., Koeppe, R. A., Knudsen, G. M., Knuuti, J., Lammertsma, A. A., Laruelle, M., Logan, J., Maguire, R. P., Mintun, M. A., Morris, E. D., Parsey, R., Price, J. C., Slifstein, M., Sossi, V., Suhara, T., Votaw, J. R., Wong, D. F., and Carson, R. E., "Consensus nomenclature for in vivo imaging of reversibly binding radioligands." *J Cereb Blood Flow Metab*, vol. 27, no. 9, pp. 1533–1539, Sep 2007. [Online]. Available: http://dx.doi.org/10.1038/sj.jcbfm.9600493

[49] Ito, H., Hietala, J., Blomqvist, G., Halldin, C., and Farde, L., "Comparison of the transient equilibrium and continuous infusion method for quantitative pet analysis of [11c] raclopride binding," *Journal of Cerebral Blood Flow & Metabolism*, vol. 18, no. 9, pp. 941–950, 1998.

[50] Jain, A. K., *Fundamentals of digital image processing*. Prentice-Hall, Inc., 1989.

[51] Jian, Y., Planeta, B., and Carson, R. E., "Evaluation of bias and variance in low-count osem list mode reconstruction." *Phys Med Biol*, vol. 60, no. 1, pp. 15–29, Jan 2015. [Online]. Available: http://dx.doi.org/10.1088/0031-9155/60/1/15

[52] Jones, T., "Positron emission tomography and measurements of regional tissue function in man," *British medical bulletin*, vol. 36, no. 3, pp. 231–236, 1980.

[53] Joutsa, J., Johansson, J., Niemelä, S., Ollikainen, A., Hirvonen, M. M., Piepponen, P., Arponen, E., Alho, H., Voon, V., Rinne, J. O., Hietala, J., and Kaasinen, V., "Mesolimbic dopamine release is linked to symptom severity in pathological gambling." *Neuroimage*, vol. 60, no. 4, pp. 1992–1999, May 2012. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2012.02.006

[54] Kak, A. C. and Slaney, M., *Principles of computerized tomographic imaging*. Society for Industrial and Applied Mathematics, 2001.

[55] Karp, J. S., Muehllehner, G., and Lewitt, R. M., "Constrained fourier space method for compensation of missing data in emission computed tomography." *IEEE Trans Med Imaging*, vol. 7, no. 1, pp. 21–25, 1988. [Online]. Available: http://dx.doi.org/10.1109/42.3925

[56] Keys, R., "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, 1981. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1163711

[57] Kinahan, P. and Rogers, J., "Analytic 3D image reconstruction using all detected events," *IEEE Transactions on Nuclear Science*, vol. 36, no. 1, pp. 964–968, 1989. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=34585

[58] Koepp, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., Brooks, D. J., Bench, C. J., and Grasby, P. M., "Evidence for striatal dopamine release during a video game." *Nature*, vol. 393, no. 6682, pp. 266–268, May 1998. [Online]. Available: http://dx.doi.org/10.1038/30498

[59] Kuhn, F. P., Warnock, G. I., Burger, C., Ledermann, K., Martin-Soelch, C., and Buck, A., "Comparison of pet template-based and mri-based image processing in the quantitative analysis of c11-raclopride pet," *EJNMMI Res*, vol. 4, no. 1, p. 7, 2014.

[60] Kuikka, J. T., Yang, J., Karhu, J., Laitinen, T., Tupala, E., Hallikainen, T., and Tiihonen, J., "Imaging the structure of the striatum: a fractal approach to spect image interpretation." *Physiol Meas*, vol. 19, no. 3, pp. 367–374, Aug 1998.

[61] Lammertsma, A. A. and Hume, S. P., "Simplified reference tissue model for pet receptor studies." *Neuroimage*, vol. 4, no. 3 Pt 1, pp. 153–158, Dec 1996. [Online]. Available: http://dx.doi.org/10.1006/nimg.1996.0066

[62] Laruelle, M., "Imaging dopamine transmission in schizophrenia. a review and meta-analysis." *Q J Nucl Med*, vol. 42, no. 3, pp. 211–221, Sep 1998.

[63] ———, "Imaging synaptic neurotransmission with in vivo binding competition techniques: a critical review." *J Cereb Blood Flow Metab*, vol. 20, no. 3, pp. 423–451, Mar 2000. [Online]. Available: http://dx.doi.org/10.1097/00004647-200003000-00001

[64] Madsen, M. T., "A simplified formulation of the gamma variate function," *Physics in Medicine and Biology*, vol. 37, no. 7, p. 1597, 1992.

[65] Magnusson, M., Danielsson, P.-E., and Edholm, P., "Artefacts and remedies in direct Fourier tomographic reconstruction," in *Nuclear Science Symposium and Medical Imaging Conference, 1992., Conference Record of the 1992 IEEE*, 1992, pp. 1138–1140. [Online].

Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=301056

[66] Martinez, D., Slifstein, M., Broft, A., Mawlawi, O., Hwang, D.-R., Huang, Y., Cooper, T., Kegeles, L., Zarahn, E., Abi-Dargham, A., Haber, S. N., and Laruelle, M., "Imaging human mesolimbic dopamine transmission with positron emission tomography. part ii: amphetamine-induced dopamine release in the functional subdivisions of the striatum." *J Cereb Blood Flow Metab*, vol. 23, no. 3, pp. 285–300, Mar 2003.

[67] Mawlawi, O., Martinez, D., Slifstein, M., Broft, A., Chatterjee, R., Hwang, D. R., Huang, Y., Simpson, N., Ngo, K., Van Heertum, R., and Laruelle, M., "Imaging human mesolimbic dopamine transmission with positron emission tomography: I. accuracy and precision of d(2) receptor parameter measurements in ventral striatum." *J Cereb Blood Flow Metab*, vol. 21, no. 9, pp. 1034–1057, Sep 2001. [Online]. Available: http://dx.doi.org/10.1097/00004647-200109000-00002

[68] Mazière, B. and Mazière, M., "Where have we got to with neuroreceptor mapping of the human brain?" *Eur J Nucl Med*, vol. 16, no. 11, pp. 817–835, 1990.

[69] Mazziotta, J. C., Phelps, M. E., Plummer, D., and Kuhl, D. E., "Quantitation in positron emission computed tomography: 5. physical–anatomical effects." *J Comput Assist Tomogr*, vol. 5, no. 5, pp. 734–743, Oct 1981.

[70] Michel, C., Schmand, M., Liu, X., Sibomana, M., Vollmar, S., Knoss, C., Lercher, M., Watson, C., Newport, D., Casey, M., Defrise, M., Wienhard, K., and Heiss, W., "Reconstruction strategies for the hrrt," in *Nuclear*

Science Symposium Conference Record, 2000 IEEE, vol. 2, 2000. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=950103

[71] Mintun, M. A., Raichle, M. E., Kilbourn, M. R., Wooten, G. F., and Welch, M. J., "A quantitative model for the in vivo assessment of drug binding sites with positron emission tomography." *Ann Neurol*, vol. 15, no. 3, pp. 217–227, Mar 1984. [Online]. Available: http://dx.doi.org/10.1002/ana.410150302

[72] Moon, T., "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, 1996. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=543975

[73] Morris, E. D., Fisher, R. E., Alpert, N. M., Rauch, S. L., and Fischman, A. J., "In vivo imaging of neuromodulation using positron emission tomography: optimal ligand characteristics and task length for detection of activation," *Human Brain Mapping*, vol. 3, no. 1, pp. 35–55, 1995.

[74] Morris, E. D., Kim, S. J., Sullivan, J. M., Wang, S., Normandin, M. D., Constantinescu, C. C., and Cosgrove, K. P., "Creating dynamic images of short-lived dopamine fluctuations with lp-ntpet: dopamine movies of cigarette smoking." *J Vis Exp*, no. 78, 2013. [Online]. Available: http://dx.doi.org/10.3791/50358

[75] Myers, R., Cunningham, V., and Bailey, D., *Quantification of brain function using PET*. Academic Press, 1996.

[76] Normandin, M. D., Schiffer, W. K., and Morris, E. D., "A linear model for estimation of neurotransmitter response profiles from dynamic pet data." *Neuroimage*, vol. 59,

no. 3, pp. 2689–2699, Feb 2012. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2011.07.002

[77] Nuyts, J., Stroobants, S., Dupont, P., Vleugels, S., Flamen, P., and Mortelmans, L., "Reducing loss of image quality because of the attenuation artifact in uncorrected pet whole-body images." *J Nucl Med*, vol. 43, no. 8, pp. 1054–1062, Aug 2002.

[78] Nyberg, S., Farde, L., and Halldin, C., "Test-retest reliability of central [11c]raclopride binding at high d2 receptor occupancy. a pet study in haloperidol-treated patients." *Psychiatry Res*, vol. 67, no. 3, pp. 163–171, Oct 1996.

[79] Ollinger, J. and Fessler, J., "Positron-emission tomography," *IEEE Signal Processing Magazine*, vol. 14, no. 1, pp. 43–55, 1997. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=560323

[80] Radon, J., "On the determination of functions from their integral values along certain manifolds," *Medical Imaging, IEEE Transactions on*, vol. 5, no. 4, pp. 170–176, 1986.

[81] Rinne, J. O., Laihinen, A., Rinne, U. K., Någren, K., Bergman, J., and Ruotsalainen, U., "Pet study on striatal dopamine d2 receptor changes during the progression of early parkinson's disease." *Mov Disord*, vol. 8, no. 2, pp. 134–138, Apr 1993. [Online]. Available: http://dx.doi.org/10.1002/mds.870080203

[82] Schain, M., Varnäs, K., Cselényi, Z., Halldin, C., Farde, L., and Varrone, A., "Evaluation of two automated methods for pet region of interest analysis." *Neuroinformatics*, vol. 12, no. 4, pp. 551–562, Oct 2014. [Online]. Available: http://dx.doi.org/10.1007/s12021-014-9233-6

[83] Shepp, L. and Vardi, Y., "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging*, vol. 1, no. 2, pp. 113–122, 1982. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4307558

[84] Shrout, P. E. and Fleiss, J. L., "Intraclass correlations: uses in assessing rater reliability." *Psychological bulletin*, vol. 86, no. 2, p. 420, 1979.

[85] Snyder, D. L., Miller, M., Thomas Jr, L. J., Politte, D. G. *et al.*, "Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography," *Medical Imaging, IEEE Transactions on*, vol. 6, no. 3, pp. 228–238, 1987.

[86] Sossi, V., de Jong, H. W. A. M., Barker, W. C., Bloomfield, P., Burbar, Z., Camborde, M.-L., Comtat, C., Eriksson, L. A., Houle, S., Keator, D., Knob, C., Krais, R., Lammertsma, A. A., Rahmim, A., Sibomana, M., Teras, M., Thompson, C. J., Trebossen, R., Votaw, J., Walker, M., Wienhard, K., and Wong, D. F., "The second generation hrrt - a multi-centre scanner performance investigation," in *Nuclear Science Symposium Conference Record, 2005 IEEE*, vol. 4, Oct. 2005, pp. 2195–2199.

[87] Stute, S. and Comtat, C., "Practical considerations for image-based psf and blobs reconstruction in pet." *Phys Med Biol*, vol. 58, no. 11, pp. 3849–3870, Jun 2013. [Online]. Available: http://dx.doi.org/10.1088/0031-9155/58/11/3849

[88] Sureau, F. C., Reader, A. J., Comtat, C., Leroy, C., Ribeiro, M.-J., Buvat, I., and Trébossen, R., "Impact of image-space resolution modeling for studies with the high-resolution research tomograph." *J Nucl Med*, vol. 49,

no. 6, pp. 1000–1008, Jun 2008. [Online]. Available: http://dx.doi.org/10.2967/jnumed.107.045351

[89] Teräs, M., Tolvanen, T., Johansson, J. J., Williams, J. J., and Knuuti, J., "Performance of the new generation of whole-body pet/ct scanners: Discovery ste and discovery vct." *Eur J Nucl Med Mol Imaging*, vol. 34, no. 10, pp. 1683–1692, Oct 2007. [Online]. Available: http://dx.doi.org/10.1007/s00259-007-0493-3

[90] Tohka, J., Wallius, E., Hirvonen, J., Hietala, J., and Ruotsalainen, U., "Automatic extraction of caudate and putamen in [c] raclopride PET using deformable surface models and normalized cuts," *IEEE Transactions on Nuclear Science*, vol. 53, no. 1, pp. 220–227, 2006. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1610975

[91] Tong, S., Alessio, A. M., Thielemans, K., Stearns, C., Ross, S., and Kinahan, P. E., "Properties and mitigation of edge artifacts in psf-based PET reconstruction," *IEEE Transactions on Nuclear Science*, vol. 58, no. 5, pp. 2264–2275, Oct. 2011.

[92] Tuna, U., Peltonen, S., and Ruotsalainen, U., "Interpolation for the gap-filling of the hrrt PET sinograms by using the slices in the direction of the radial samples," in *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, 2009, pp. 3273–3279. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5401727

[93] Tuna, U., Johansson, J., and Ruotsalainen, U., "Comparison of 3D-rp and 3D-oposem reconstructions of the ecat hrrt PET data," in *Nuclear Science Symposium*

Conference Record (NSS/MIC), 2010 IEEE, 2010, pp. 3511–3515. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5874460

[94] Tuna, U., Peltonen, S., and Ruotsalainen, U., "Data estimation for the ecat hrrt sinograms by utilizing the dct domain," in *Nuclear Science Symposium Conference Record, 2008. NSS'08. IEEE*. IEEE, 2008, pp. 5076–5080.

[95] ———, "Gap-filling for the high-resolution pet sinograms with a dedicated dct-domain filter." *IEEE Trans Med Imaging*, vol. 29, no. 3, pp. 830–839, Mar 2010. [Online]. Available: http://dx.doi.org/10.1109/TMI.2009.2037957

[96] Tziortzi, A. C., Searle, G. E., Tzimopoulou, S., Salinas, C., Beaver, J. D., Jenkinson, M., Laruelle, M., Rabiner, E. A., and Gunn, R. N., "Imaging dopamine receptors in humans with [11c]-(+)-phno: dissection of d3 signal and anatomy." *Neuroimage*, vol. 54, no. 1, pp. 264–277, Jan 2011. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2010.06.044

[97] Tziortzi, A. C., Haber, S. N., Searle, G. E., Tsoumpas, C., Long, C. J., Shotbolt, P., Douaud, G., Jbabdi, S., Behrens, T. E. J., Rabiner, E. A., Jenkinson, M., and Gunn, R. N., "Connectivity-based functional analysis of dopamine release in the striatum using diffusion-weighted mri and positron emission tomography." *Cereb Cortex*, vol. 24, no. 5, pp. 1165–1177, May 2014. [Online]. Available: http://dx.doi.org/10.1093/cercor/bhs397

[98] van Velden, F., Kloet, R., de Jong, H., Lammertsma, A., and Boellaard, R., "Quantitative experimental comparison of hrrt versus hr+ PET brain studies,"

in *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol. 5, 2006, pp. 3097–3099. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4179686

[99] van Velden, F., Kloet, R., van Berckel, B. N., Molthoff, C. F., Lammertsma, A., and Boellaard, R., "Gap filling strategies for 3-D-fbp reconstructions of high-resolution research tomograph scans," *IEEE Transactions on Medical Imaging*, vol. 27, no. 7, pp. 934–942, 2008. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4483776

[100] van Velden, F. H. P., Kloet, R. W., van Berckel, B. N. M., Wolfensberger, S. P. A., Lammertsma, A. A., and Boellaard, R., "Comparison of 3d-op-osem and 3d-fbp reconstruction algorithms for high-resolution research tomograph studies: effects of randoms estimation methods." *Phys Med Biol*, vol. 53, no. 12, pp. 3217–3230, Jun 2008. [Online]. Available: http://dx.doi.org/10.1088/0031-9155/53/12/010

[101] van Velden, F. H. P., Kloet, R. W., van Berckel, B. N. M., Lammertsma, A. A., and Boellaard, R., "Accuracy of 3-dimensional reconstruction algorithms for the high-resolution research tomograph." *J Nucl Med*, vol. 50, no. 1, pp. 72–80, Jan 2009. [Online]. Available: http://dx.doi.org/10.2967/jnumed.108.052985

[102] Varrone, A., Sjöholm, N., Eriksson, L., Gulyás, B., Halldin, C., and Farde, L., "Advancement in pet quantification using 3d-op-osem point spread function reconstruction with the hrrt." *Eur J Nucl Med Mol Imaging*, vol. 36, no. 10, pp. 1639–1650, Oct 2009. [Online]. Available: http://dx.doi.org/10.1007/s00259-009-1156-3

[103] Volkow, N. D., Fowler, J. S., Wang, G.-J., and Swanson, J. M., "Dopamine in drug abuse and addiction: results from imaging studies and treatment implications." *Mol Psychiatry*, vol. 9, no. 6, pp. 557–569, Jun 2004. [Online]. Available: http://dx.doi.org/10.1038/sj.mp.4001507

[104] Volkow, N. D. and Swanson, J. M., "Variables that affect the clinical use and abuse of methylphenidate in the treatment of adhd." *Am J Psychiatry*, vol. 160, no. 11, pp. 1909–1918, Nov 2003. [Online]. Available: http://dx.doi.org/10.1176/appi.ajp.160.11.1909

[105] Walker, M. D., Asselin, M.-C., Julyan, P. J., Feldmann, M., Talbot, P. S., Jones, T., and Matthews, J. C., "Bias in iterative reconstruction of low-statistics pet data: benefits of a resolution model." *Phys Med Biol*, vol. 56, no. 4, pp. 931–949, Feb 2011. [Online]. Available: http://dx.doi.org/10.1088/0031-9155/56/4/004

[106] Watabe, H., Endres, C. J., Breier, A., Schmall, B., Eckelman, W. C., and Carson, R. E., "Measurement of dopamine release with continuous infusion of [11c]raclopride: optimization and signal-to-noise considerations." *J Nucl Med*, vol. 41, no. 3, pp. 522–530, Mar 2000.

[107] Weir, J. P., "Quantifying test-retest reliability using the intraclass correlation coefficient and the sem." *J Strength Cond Res*, vol. 19, no. 1, pp. 231–240, Feb 2005. [Online]. Available: http://dx.doi.org/10.1519/15184.1

[108] Wienhard, K., Schmand, M., Casey, M., Baker, K., Bao, J., Eriksson, L., Jones, W., Knoess, C., Lenox, M., Lercher, M., Luk, P., Michel, C., Reed, J., Richerzhagen, N., Treffert, J., Vollmar, S., Young, J.,

Heiss, W., and Nutt, R., "The ecat hrrt: performance and first clinical application of the new high resolution research tomograph," *IEEE Transactions on Nuclear Science*, vol. 49, no. 1, pp. 104–110, 2002. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=998689

[109] Zaidi, H., Ojha, N., Morich, M., Griesmer, J., Hu, Z., Maniawski, P., Ratib, O., Izquierdo-Garcia, D., Fayad, Z. A., and Shao, L., "Design and performance evaluation of a whole-body ingenuity tf pet-mri system." *Phys Med Biol*, vol. 56, no. 10, pp. 3091–3106, May 2011. [Online]. Available: http://dx.doi.org/10.1088/0031-9155/56/10/013

[110] Zhou, Y., Chen, M.-K., Endres, C. J., Ye, W., Brašić, J. R., Alexander, M., Crabb, A. H., Guilarte, T. R., and Wong, D. F., "An extended simplified reference tissue model for the quantification of dynamic pet with amphetamine challenge," *Neuroimage*, vol. 33, no. 2, pp. 550–563, 2006.

Publications

Publication I

Johansson, J., Oikonen, V., Teräs, M., "Quantitative brain imaging using the new, fast iterative histogram-mode reconstruction for the HRRT PET scanner," *Nuclear Science Symposium Conference Record.*, 2007, NSS '07. IEEE.

Quantitative Brain Imaging Using the New, Fast Iterative Histogram-Mode Reconstruction for the HRRT PET Scanner

Jarkko Johansson, *Student Member, IEEE,* Vesa Oikonen, and Mika Teräs, *Member, IEEE*

***Abstract*—Specialized methods to overcome limitations in low count statistics HRRT brain imaging have been developed and implemented recently. Combination of ordinary Poisson (OP) OSEM-3D reconstruction with advanced randoms calculation has been shown to decrease quantification bias in low count statistics, but a residual bias is however seen.**

In the present study the extent and origin of the quantification bias in the histogram-mode iterative reconstruction (OP-OSEM-3D) scheme was further investigated. Aim of the study was to assess the minimum emission measurement signal strength, to express the minimum signal strength in noise-equivalent-counts (NEC) and to identify the origin of the residual bias.

To assess the minimum count statistics anthropomorphic brain phantoms were scanned in the scanners dynamic range. Dynamic range was simulated by histogramming the coincidence data into overlapping variable length frames. True distribution was estimated from a high statistics frame and ROI analysis was applied to observe biases in the quantification.

ROI analysis showed significant quantification errors in low count statistics. Moreover, regional sensitivity to bias was shown to correlate with image features. Our conclusion is that the residual quantification bias is originated in the iterative reconstruction, and that it needs to be taken into consideration in dynamic HRRT imaging when OP-OSEM-3D is used.

***Index Terms*—positron emission tomography, image reconstruction, dynamic imaging, high resolution tomograph.**

I. Introduction

THE ECAT High Resolution Research Tomograph (HRRT) is a dedicated human brain PET scanner with nearly isotropic and shift invariant reconstructed resolution of 2.5 mm (FWHM) [1]. High spatial resolution is obtained with high number (119808) of small-size ($2.1 \times 2.1 \times 10$ mm^3) LSO/LYSO crystals. Due to high number of small crystals HRRT sinograms are substantially larger and typically have less counts per bin than conventional PET sinograms. Moreover, HRRT planar detector head design with gaps introduces zero bins in the sinogram data. These missing data hinders usage of the analytical reconstruction methods with the HRRT.

Alternative to analytical reconstruction is the statistical iterative OSEM reconstruction, where missing data is easily incorporated into the system model [2]. It has, however, been reported that OSEM reconstruction is a source of quantification biases in limited count statistics in general [3], and especially in the case of the HRRT [4].

To overcome low count statistics limitations in the HRRT, specialized methods have been developed and implemented. Firstly, implementation of ordinary Poisson (OP) weighting scheme in the OSEM-3D reconstruction [5] has been shown to decrease the bias due to negatives truncation at a price of slower convergence in comparison to ANW-OSEM-3D scheme. Secondly, usage of variance reduction on random (VRR) coincidences algorithm [6] in combination with OP-OSEM-3D has been shown to further reduce the bias in low statistics [4]. Thirdly, speed optimized version of the OSEM-3D algorithm [7] increases feasibility of continued reconstructions.

According to a previous study the above described, optimized histogram-mode reconstruction still suffers from residual bias and further refinement is requested. However, specific refinement strategies or practicable limit for count statistics was not presented.

Therefore, in the present study the extent and origin of the quantification bias in the histogram-mode iterative reconstruction (OP-OSEM-3D) scheme was further investigated. Aim of the study was to assess the minimum emission signal strength, to express the minimum signal strength in noise-equivalent-counts (NEC) and to identify the origin of the bias.

J. Johansson, V. Oikonen and M. Teräs are with the Turku PET Centre, Turku University Hospital.

Address correspondence to Jarkko Johansson, Turku PET Centre, Kiinamyllynkatu 4-8, 20520, Turku, Finland.

Address e-mail to jarkko.johansson@tyks.fi.

II. Materials and Methods

A. Positron emission tomograph

All emission and transmission data was acquired using the HRRT PET scanner. This 3D only scanner has approximately 4.5×10^9 lines-of-response (LOR) are detected, in the field-of-view (FOV) of 252 mm in axial and 312 mm in transaxial directions. Emission acquisitions were made in list-mode using 400-650 keV energy window.

B. Phantom measurements

Hoffman phantom was filled with ^{18}F-FDG in water solution so that the activity concentration at start was 35 MBq and emission data was acquired for 70 mins. Countrate was such that dead time and randoms ratio as assessed from the head-curve were negligible (DT $= 6.5\% - 5.8\%$ and RTR $= 0.147 - 0.132$).

Striatal phantom was filled with 21 MBq of ^{18}F-FDG in water solution and emission data was acquired for 57 mins.

Striatal structures contained approximately five times higher activity concentration compared to the background.

C. Dynamic human measurements

Twelve ^{11}C-raclopride scans with healthy volunteers were made within a scanner validation project, and their head-curve data was exploited in this study. Head-curve data were integrated with respect to time to yield sufficient and equal count statistics in each time frame. Sufficient count statistics was estimated from the phantom studies. The head of the subjects were always aligned to the center of the FOV, and injected dose per kilogram varied from 3.3 to 7.3 MBq/kg, with the mean 5.1 MBq/kg.

D. Histogramming and reconstructions

In phantom measurements transmission data were acquired with activity in the FOV, and histogrammed using shifted-mock technique to estimate emission contamination [8]. MAP-TR reconstructions were made to create μ-images using water phantom priors for Hoffman phantom and human brain priors for Striatal phantom and human scans. Attenuation correction factors were obtained by forward projecting the μ-images in 3D.

Phantom measurements were histogrammed into 31 overlapping frames always adding more statistics to the previous sinogram, so that the last frame spanned the whole scan duration. Histogramming was made into separate prompts and randoms files, with the axial compression of span 9.

Framing yielded NEC range of 6×10^4 to 3×10^8 counts in the Hoffman phantom measurement, and 1×10^5 to 3×10^8 counts in the Striatal phantom measurement. NEC was calculated as [9]

$$NEC = \frac{((1-S)(P-R))^2}{P+R}\ , \tag{1}$$

with S the scatter fraction, and P and R the prompts and randoms respectively. In phantom measurements total prompts (P) and randoms (R) were extracted from sinogram counts and scatter fraction (S) was reported in the scatter calculation. In human measurements prompts and randoms were extracted from the head-curve and scatter fraction was set to 0.41.

Scatter estimates were calculated using the single scatter simulation method from the 3D attenuation, emission and normalization [10] sinograms. Variations in the scatter fraction calculation were negligible in both phantom measurements (within $\pm 5\%$ from the high statistics frame).

Additional set of randoms sinograms were created using measured delayed coincidence histograms with the new VRR algorithm [6].

OP-OSEM-3D reconstructions were made with the above described data using the new accelerated version of the CTI/Siemens software [11] with 16 subsets and eleven full iterations. Reconstructed image volume contains $256 \times 256 \times 207$ voxels, with isotropic voxel dimensions of $1.22 \times 1.22 \times 1.22$ mm^3.

E. Regional analysis

Twentyseven separate regions-of-interest (ROI) in the brain and one region to cover the whole FOV were delineated on the high statistics image of the Hoffman phantom. Bilateral regions were summed in the analysis to yield less noise in a total of fifteen ROIs.

Striatal phantom ROI set consists of five regions, including caudatus and putamen in both hemispheres and one region in the background.

Systematic errors in the ROI mean values were estimated as

$$REM_k = \frac{\text{mean}(ROI_k, \hat{F})/T_{\hat{F}}}{\text{mean}(ROI_k, F)/T_F}, \tag{2}$$

where $\hat{F}$ is any estimate image, F is the reference (high statistics) image and T_F is the total trues in frame F.

REM_k-values were analyzed as function of NEC for each ROI k, and minimum NECs were extracted from the graphs to obtain desired error level ($REM_k \in [0.975, 1.025]$). In this context the minimum NEC was interpreted as regional sensitivity to bias.

In the Hoffman phantom study the possible correlations between regional sensitivity to bias and reference image contrast factors were statistically evaluated. Regional contrast factors were calculated as

$$Q_k(F) = \frac{\text{mean}(ROI_k, F)}{\text{mean}(\partial ROI_k, F)}\ , \tag{3}$$

where $\text{mean}(R, F)$ is the mean pixel value inside region R and ∂ROI_k is a two to three pixels wide edge of the region ROI_k. Pearson's product-moment coefficient and corresponding p-values were used to measure the signifigance of the correlation.

F. Convergence analysis

Optimal number of iterations in the OP-OSEM-3D reconstruction was approximated using a high count statistics (NEC $> 10^8$) measurement of the Hoffman phantom and REM_k-values. Iterations were stopped after 32 full iterations and intermediate estimates were compared to this reference. Accuracy was considered sufficient at intermediate estimate if it deviates from the reference less than $\pm 2.0\%$ in all regions, that is $REM_k \in [0.98, 1.02]\ \forall\, k$.

G. Image noise analysis

In the Hoffman phantom study the image noise was estimated regionally as

$$SNR_k = \frac{\text{mean}(ROI_k, \hat{F})}{\text{std}(ROI_k, \hat{F})}\ , \tag{4}$$

where $\text{std}(ROI_k, \hat{F})$ is the standard deviation of the pixel values of image estimate $\hat{F}$ inside ROI_k. Linear fittings between the image derived SNR_k^2 and sinogram NEC were made (cf. [12]), and the slopes were examined to assess the noise properties of the reconstruction methods.

III. Results and Discussion

A. Convergence and noise properties of the OP-OSEM-3D iterates

Depending on the brain region sufficient accuracy was reached after 3-11 iterations. Rate of convergence was highest in the cerebellum and slowest in the medial frontal cortex and caudate nucleus. Ratio image of the second and last iterate in the Figure 1 visualizes spatially varying rate of convergence.

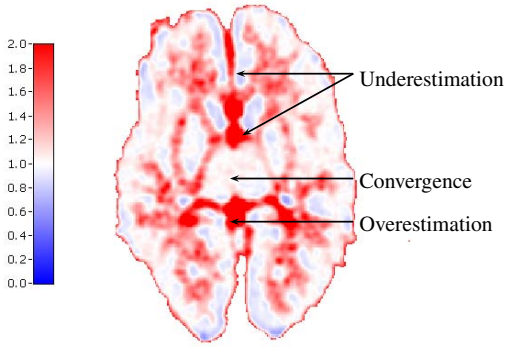


Fig. 1. Ratio image of the second and last iterate in the OP-OSEM-3D Hoffman phantom reconstruction.

Noise properties of the OP-OSEM-3D iterates were assessed by evaluating the NEC versus SNR_k^2 (Eq. 4) responses. Graphs in the Figure 2A-C show that later iterate has significantly smaller slope than eleventh iterate, and it is also seen that noise variance reduction on randoms has little or no effect to the slope.

According to above convergence and noise analysis OP-OSEM-3D requires eleven iterations for sufficient accuracy and that noise amplification limits number of iterations from above.

B. Hoffman phantom study

Accuracy of the OP-OSEM-3D in the presence of measurement noise was assessed in the Hoffman phantom study. Assessment was made regionally using the REM_k-values in the Equation (2).

In the whole FOV and in the white brain matter REM_k-graphs (Figure 3A-B) showed systematic overestimation with decreasing NEC, whereas in the grey brain matter (example in Figure 3C) systematic underestimation was seen.

Minimum NEC values to obtain maximum $\pm 2.5\%$ error in the REM_k-values were extracted and tabulated in Table I. Region dependency in the sensitivity to bias was examined by evaluating the possible correlation between minimum NECs and contrast factors Q_k in the Table I. A weak correlation was seen with $r^2 \approx 0.3$ and $p \approx 0.053$. Correlation was substantially higher after removing one outlier (caudatus).

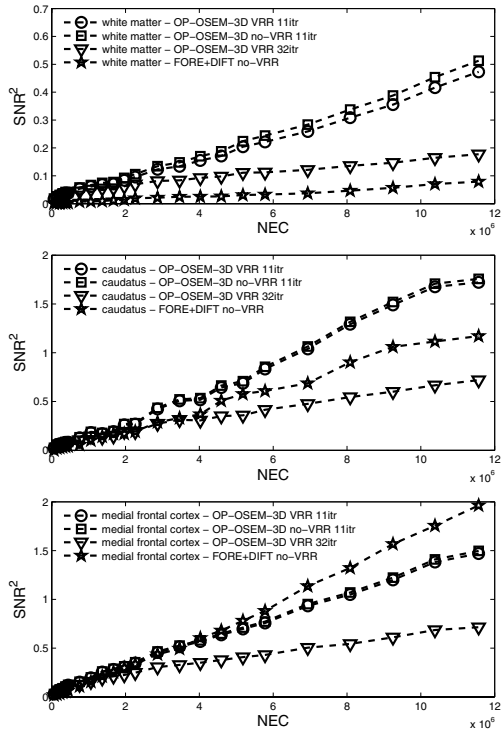


Fig. 2. SNR^2 versus NEC. Analytical FORE+DIFT reconstructions added for scaling. A: white brain matter. B: caudate nucleus. C: medial frontal cortex.

TABLE I
Minimum NEC values to obtain maximum $\pm 2.5\%$ error regionally - Hoffman phantom study.

Region name	Q_k	Minimum NEC
anterior cingulate gyrus	1.9	7.9×10^6
angular gyrus	1.6	7.7×10^6
caudatus	1.6	1.7×10^7
cerebellum	1.2	3.2×10^5
dorsolateral prefrontal cortex	1.6	6.3×10^6
inferior parietal cortex	1.7	2.4×10^6
insular gyrus	1.3	2.5×10^6
lentiform nucleus	1.5	5.8×10^6
medial frontal cortex	1.7	8.4×10^6
posterior cingulate cortex	2.0	8.2×10^6
pons	1.4	4.8×10^6
supramarginal gyrus	1.7	2.2×10^6
lateral thalamus	1.0	2.3×10^6
medial thalamus	1.1	6.9×10^5
white matter	1.0	1.1×10^8
whole field-of-view		3.9×10^7
Mean (all)		1.4×10^7
Mean (gray matter)		5.5×10^6

C. Striatal phantom study

REM_k-graphs indicated systematic underestimation in the hot regions and overestimation in the cold regions (data not shown here). NEC values to obtain maximum $\pm 2.5\%$ error in the REM_k-values were extracted and tabulated in the Table II.

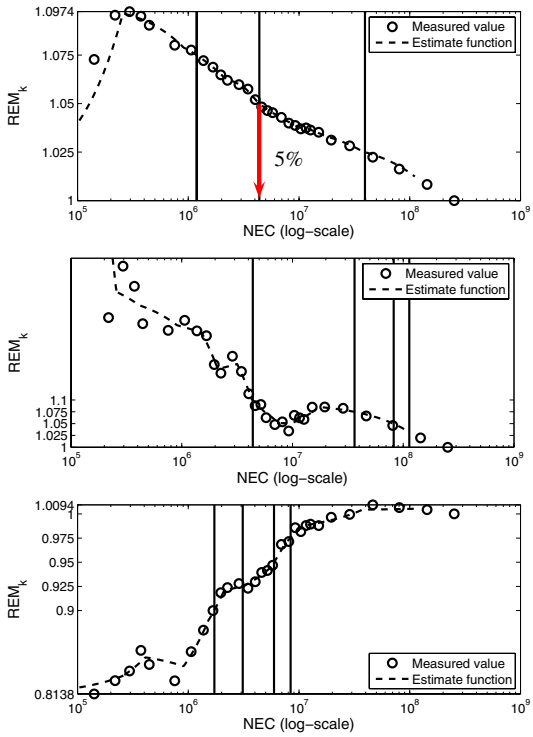


Fig. 3. REM_k-values as function of the NEC. A: the whole FOV. B: white brain matter. C: gray brain matter.

TABLE II
MINIMUM NEC VALUES TO OBTAIN MAXIMUM $\pm 2.5\%$ ERROR REGIONALLY - STRIATAL PHANTOM STUDY.

Region name	Minimum NEC
background	7.35×10^5
caudatus dx	1.30×10^7
caudatus sin	6.43×10^6
putamen dx	6.27×10^6
putamen sin	3.87×10^6
Mean (all)	6.06×10^6
Mean (gray matter)	7.39×10^6

D. Human measurements

According to our Hoffman and Striatal phantom measurements, count statistics in human raclopride-studies should exceed approximately 7.5×10^6 NEC in each frame. To assess temporal resolution head-curve data from twelve human scans were integrated with respect to time to yield 7.5×10^6 NEC in each frame. Shortest frame durations to yield sufficient count statistics at peak count rate are plotted against the injected dose in the Figure 4. According to the graph minimum frame duration at peak count rate would be approximately 130 sec. in a typical raclopride study. In the later parts of the scans the minimum frame durations corresponded well to the decay of the isotope (data not shown).

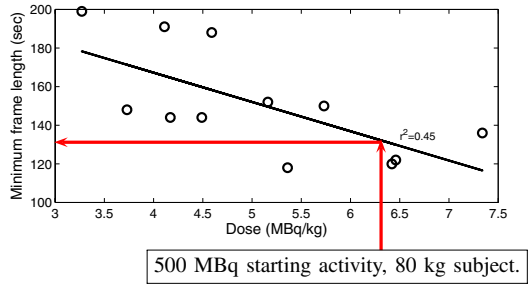


Fig. 4. Injected doses versus minimum frame durations to yield 7.5×10^6 NEC at peak count rate.

IV. CONCLUSION

Count statistics limitations in ROI quantification of OP-OSEM-3D reconstructed HRRT images was analysed in this study. Practicable guidelines for quantified dynamic imaging was given in the form of minimum noise-equivalent-counts and the limit was applied to human raclopride scans to assess the temporal resolution.

Rationale of a global measure in human scans is questionable however. For instance the positioning of the subject and distribution of the radiopharmaceutical have a great impact to a global measure. However, when a reasonably reproducible positioning of a known object is available, which is the case in the HRRT brain scans, application of a global measure is justified.

It has been seen in previous studies that OSEM reconstruction suffers from contrast degradation in the presence of noise [6], [4]. This in conjuction with our observation of region dependence in the sensitivity to bias suggest that the bias is originated mainly in the iterative procedure (OP-OSEM-3D).

For study protocols where higher temporal resolution is required there are several options. In the case of the HRRT there is a trade-off between the temporal resolution and the spatial resolution. In principle better temporal resolution is expected when spatial resolution is lowered, for instance by summing up adjacent lines-of-response.

In the case that spatial resolution can not be compromised other strategies must be applied. For the ROI analysis fully three dimensional analytical reconstruction (3D-FBP) has been shown to provide robust estimates in varying noise levels [13]. In the pixel-by-pixel analysis, however, the analytical method is likely suboptimal because of poor noise properties.

As the bias appears to be contrast dependent inclusion of a contrast enhancing mechanism, such as scanner PSF model [14] or a bayesian prior [15] in the iterative reconstruction possibly reduces bias. In our future work we are planning to further investigate different OP-OSEM-3D variants and their impact to the relialibility of the clinical parameters.

ACKNOWLEDGMENT

The authors would like to thank Jussi Hirvonen and Mika Hirvonen for ROI delineation, and Juha Rinne for providing the human data.

This work was financially supported by the Turku University Foundation.

REFERENCES

[1] H. W. A. M. de Jong, F. H. P. van Velden, R. W. Kloet, F. L. Buijs, R. Boellaard, and A. A. Lammertsma, "Performance evaluation of the ecat hrrt: an lso-lyso double layer high resolution, high sensitivity scanner." *Phys Med Biol*, vol. 52, no. 5, pp. 1505–1526, Mar 2007. [Online]. Available: http://dx.doi.org/10.1088/0031-9155/52/5/019

[2] C. Michel, M. Schmand, X. Liu, M. Sibomana, S. Vollmar, C. Knoss, M. Lercher, C. Watson, D. Newport, M. Casey, M. Defrise, K. Wienhard, and W. Heiss, "Reconstruction strategies for the hrrt," in *Nuclear Science Symposium Conference Record, 2000 IEEE*, vol. 2, 15-20 Oct. 2000, pp. 15/207–15/212vol.2.

[3] R. Boellaard, M. Lubberink, H. de Jong, M. Kropholler, and A. Lammertsma, "Application of various iterative reconstruction for quantitative 3d dynamic brain pet studies," in *Nuclear Science Symposium Conference Record, 2004 IEEE*, vol. 4, 16-22 Oct. 2004, pp. 2553–2556Vol.4.

[4] F. H. P. van Velden, R. W. Kloet, H. W. A. M. de Jong, A. A. Lammertsma, and R. Boellaard, "Quantitative experimental comparison of hrrt versus hr+ pet brain studies," in *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol. 5, Oct. 2006, pp. 3097–3099.

[5] C. Comtat, F. Bataille, C. Michel, J. Jones, M. Sibomana, L. Janeiro, and R. Trebossen, "Osem-3d reconstruction strategies for the ecat hrrt," in *Nuclear Science Symposium Conference Record, 2004 IEEE*, vol. 6, 16-22 Oct. 2004, pp. 3492–3496Vol.6.

[6] L. Byars, M. Sibomana, Z. Burbar, J. Jones, V. Panin, W. Barker, J.-S. Liow, R. Carson, and C. Michel, "Variance reduction on randoms from coincidence histograms for the hrrt," in *Nuclear Science Symposium Conference Record, 2005 IEEE*, vol. 5, 23-29 Oct. 2005, pp. 2622–2626.

[7] I. Hong, S. Chung, Y. Son, H. Kim, Y. Kim, and Z. Cho, "Some new developments in fast reconstruction with osem3d using symmetric simd in rotational coordinate," 2006, in HRRT Users Meeting 2006.

[8] M. Sibomana, L. Byars, V. Panin, M. Lenox, F. Kehren, J. Rist, Z. Burbar, C. Michel, C. Knoess, and H. deJong, "Simultaneous measurement of transmission and emission contamination using a collimated 137cs point source for the hrrt," in *IEEE Nuclear Science Symposium Conference Record*, 2004.

[9] S. Strother, M. Casey, and E. Hoffman, "Measuring pet scanner sensitivity: relating countrates to image signal-to-noise ratios using noise equivalents counts," *Nuclear Science, IEEE Transactions on*, vol. 37, no. 2, pp. 783–788, April 1990.

[10] C. Watson, "New, faster, image-based scatter correction for 3d pet," *IEEE Transactions on Nuclear Science*, vol. 47, pp. 1588–1594, 2000.

[11] I. K. Hong, S. T. Chung, H. K. Kim, Y. B. Kim, Y. D. Son, and Z. H. Cho, "Ultra fast symmetry and simd-based projection-backprojection (ssp) algorithm for 3-d pet image reconstruction." *IEEE Trans Med Imaging*, vol. 26, no. 6, pp. 789–803, Jun 2007.

[12] M. Dahlbom, C. Schiepers, and J. Czernin, "Comparison of noise equivalent count rates and image noise," in *Nuclear Science Symposium Conference Record, 2004 IEEE*, vol. 6, 16-22 Oct. 2004, pp. 3638–3642Vol.6.

[13] F. H. P. van Velden, R. W. Kloet, B. N. M. Berckel, A. A. Lammertsma, and R. Boellaard, "Accuracy of various 3d-osem versus 3d-fbp reconstruction of hrrt pet studies," in *Nuclear Science Symposium Conference Record, 2007. IEEE*, 2007.

[14] F. Sureau, C. Comtat, A. Reader, C. Leroy, M. Santiago-Ribeiro, I. Buvat, and R. Trebossen, "Improved clinical parametric imaging using list-mode reconstruction via resolution modeling," in *Nuclear Science Symposium Conference Record, 2006. IEEE*, vol. 6, Oct. 29 2006-Nov. 1 2006, pp. 3507–3510.

[15] S. Alenius, U. Ruotsalainen, and J. Astola, "Using local median as the location of the prior distribution in iterative emission tomography image reconstruction," in *Nuclear Science Symposium, 1997. IEEE*, vol. 2, 9-15 Nov. 1997, pp. 1726–1730vol.2.

Publication II

Alakurtti, K., Aalto, S., Johansson, J., Någren, K., Tuokkola, T., Oikonen, V., Laine, M., Rinne, J. O., "Reproducibility of striatal and thalamic dopamine D2 receptor binding using [11C]raclopride with high-resolution positron emission tomography.," *J Cereb Blood Flow Metab.*, 2011, 31, 1, 155-165.

Journal of Cerebral Blood Flow & Metabolism (2011) 31, 155–165

www.jcbfm.com

Reproducibility of striatal and thalamic dopamine D2 receptor binding using [^{11}C]raclopride with high-resolution positron emission tomography

Kati Alakurtti[1], Sargo Aalto[1], Jarkko J Johansson[1], Kjell Någren[1], Terhi Tuokkola[1], Vesa Oikonen[1], Matti Laine[2] and Juha O Rinne[1]

[1]*Turku PET Centre, University of Turku and Turku University Hospital, Turku, Finland;* [2]*Department of Psychology and Logopedics, Åbo Akademi University, Turku, Finland*

Positron emission tomography (PET) imaging of small striatal brain structures such as the ventral striatum (VST) has been hampered by low spatial resolution causing partial-volume effects. The high-resolution research tomograph (HRRT) is a brain-dedicated PET scanner that has considerably better spatial resolution than its predecessors. However, its superior spatial resolution is associated with a lower signal-to-noise ratio. We evaluated the test–retest reliability of the striatal and thalamic dopamine D_2 receptor binding using the HRRT scanner. Seven healthy male volunteers underwent two [^{11}C]raclopride PET scans with a 2.5-hour interval. Dopamine D_2 receptor availability was quantified as binding potential (BP_{ND}) using the simplified reference tissue model. To evaluate the reproducibility of repeated BP_{ND} estimations, absolute variability (VAR) and intraclass correlation coefficients (ICCs) were calculated. VAR values indicated fairly good reproducibility and were 3.6% to 4.5% for the caudate nucleus and putamen and 4.5% to 6.4% for the lateral and medial part of the thalamus. In the VST, the VAR value was 5.8% when the definition was made in the coronal plane. However, the ICC values were only moderate, in the range of 0.34 to 0.66, for all regions except the putamen (0.87). Experimental signal processing methods improved neither ICC nor VAR values significantly.
Journal of Cerebral Blood Flow & Metabolism (2011) **31,** 155–165; doi:10.1038/jcbfm.2010.64; published online 5 May 2010

Keywords: dopamine receptors; PET; raclopride; reproducibility; test–retest; ventral striatum

Introduction

In quantification of striatal D_2 receptor binding *in vivo* in humans using positron emission tomography (PET), [^{11}C]raclopride is a well-validated tracer. Previous studies with different test–retest time intervals and quantification methods indicate that the region of interest (ROI) measurement of striatal [^{11}C]raclopride binding is highly reliable (Volkow *et al*, 1993; Mawlawi *et al*, 2001; Hirvonen *et al*, 2003). However, PET measurements have bias because of the partial-volume effect (PVE) caused by the low spatial resolution, especially when imaging the small, D_2 receptor-rich striatal brain structures such as the caudate nucleus and the ventral striatum (VST) (Sureau *et al*, 2008). Brain imaging has recently benefited from the development of scanners such as the high-resolution research tomograph (HRRT), which has a spatial resolution in the range of about 2.5 to ~3 mm in the 10-cm field of view that covers most of the brain (de Jong *et al*, 2007). The ability to reliably measure dopamine D_2 receptor binding in striatal subregions is important, because these regions are differentially involved in cognitive and motor performance both in healthy volunteers and in patients with Parkinson's disease (DeLong, 1983; Evarts and Wise, 1984; Volkow *et al*, 1993; Bäckman *et al*, 2000; Kaasinen and Rinne, 2002). In addition, the VST has recently attracted attention because of its role in human behavior and its disorders such as addiction (Floresco *et al*, 2008; Belin and Everitt, 2008). However, applying a high-resolution scanner is associated with a lower signal-to-noise ratio, which may impair the reliability of the measurement (Sureau *et al*, 2008; Johansson *et al*, 2007). Evaluation of the test–retest variability of PET measurement is particularly important in studies in which the same individuals are scanned repeatedly before and after intervention or when small brain

Correspondence: K Alakurtti, Turku PET Centre, Kiinanmyllynkatu 4-8, Turku FIN-20520, Finland.
E-mail: khalak@utu.fi
This study was financially supported by the Academy of Finland (project #111879) and by clinical grants of the Turku University Hospital (EVO).
Received 2 December 2009; revised 5 March 2010; accepted 6 April 2010; published online 5 May 2010

structures, such as the VST, are analyzed. Knowledge about the test–retest variability also allows to estimate the number of subjects needed to detect a change in tracer binding, for instance after experimental intervention.

In this study, we evaluated the test–retest reliability of the striatal and thalamic D_2 binding by administering a single bolus of [11C]raclopride twice on the same day in seven healthy males and collecting data using the HRRT scanner. A novel approach in data sampling was implemented to account for potential bias in statistical image reconstruction of the HRRT data, and a recently introduced resolution-modeling image reconstruction algorithm (Comtat *et al*, 2008) was applied for evaluation purposes.

Materials and methods

The study protocol was approved by the Ethics Committee of the Hospital District of Southwestern Finland. The study subjects were given written information about all the relevant issues involved in this study. A written consent was obtained from each subject. This study was performed according to the ethical guidelines given by the Declaration of Helsinki.

Subjects

Seven healthy male subjects free of any somatic or psychiatric illness volunteered to take part in the study. All were right handed and nonsmokers. To exclude any structural brain abnormalities and to obtain anatomic reference, all subjects underwent 1.5 T magnetic resonance imaging. The age, height, and weight of the subjects were 24.5 ± 3.5 years, 185.5 ± 12.5 cm, and 74 ± 14 kg, respectively (mean ± s.d.). Each subject underwent two [11C]raclopride PET scans during the same day between 10.00 a.m. and 6.00 p.m., at least 2.5 hours apart.

Preparation of [11C]raclopride

The precursor *O*-desmethylraclopride free base and raclopride were obtained from ABX advanced biochemical compounds (Radeberg, Germany). [11C]methane was produced at the Accelerator Laboratory of Åbo Akademi with a 103-cm isochronous Efremov cyclotron using the $^{14}N(p,\alpha)^{11}C$ reaction. [11C]methyl iodide with high specific radioactivity was prepared from [11C]methane (Larsen *et al*, 1997; Någren *et al*, 2003). [11C]raclopride was prepared from [11C]methyl triflate according to a published procedure (Langer *et al*, 1999). The radiochemical purity and the specific radioactivity of the product were determined using high-performance liquid chromatography and ultraviolet detection at 214 nm. The volume of the final product solution was calculated by weighing the sterile product vessel before and after sterile filtration and by dividing by the density of the sterile solvent.

Positron Emission Tomography Imaging

Positron emission tomography experiments were performed using a brain-dedicated high-resolution PET scanner, the ECAT HRRT (Siemens Medical Solutions, Knoxville, TN, USA). The HRRT is a dual-layer, crystal-detector scanner allowing a depth-of-interaction measurement for the coincident photons. It is characterized by an isotropic 2.5-mm intrinsic spatial resolution. Spatial resolution in the reconstructed images varies in the radial and tangential directions from about 2.5 to ~3 mm and in the axial directions from 2.5 to 3.5 mm in the 10-cm field of view covering most of the brain (de Jong *et al*, 2007).

The left antecubital vein was cannulated and a single [11C]raclopride bolus injection was administered intravenously and flushed with saline. The injected dose, specific radioactivity, and the injected mass for the test and retest scans were 413 ± 91 MBq, 222 ± 154 MBq/nmol, and $2.47 \pm 4.68\ \mu$g versus 393 ± 99 MBq, 257 ± 142 MBq/nmol, and $0.68 \pm 0.35\ \mu$g, respectively, (mean ± s.d.), see Table 1. There were no statistically significant differences in the injected doses (paired *t*-test, $P = 0.56$), in the specific radioactivities (paired *t*-test, $P = 0.15$), or in the injected masses (paired *t*-test, $P = 0.33$) between the test and the retest scans. However, one outlier was detected in the boxplot analysis of the injected masses. In the outlier case, the injected mass was 13.03 μg, which is considerably higher than the average of others. Although the injected mass was higher than the others, it was considered tolerable with regard to pharmacological effects and thus this subject was not excluded from the analysis. Radiochemical purity was over 99.5% in all cases. Before each emission scan, a transmission scan was performed using ^{137}Cs point source. Emission scan was started at injection time and data were collected for 55 minutes in list mode format.

An individually shaped thermoplastic mask was used with each subject to minimize the head movement. To monitor the head movement, an external position detector (Polaris Vicra, Northern Digital, Waterloo, Ontario, Canada) was used. The position information was converted into motion information by subtracting the average reference position (transmission scan position) coordinates from the position data. A representative motion parameter was obtained by taking the norm of the translation vector. The external motion data were used to estimate three features of motion, which hamper the quality of PET imaging. First, translocation from the transmission scan (TX) position may cause erroneous attenuation correction. To measure distance from the TX position, the mean of the norm vector was calculated within each frame. Second, within-frame motion, which may cause signal blurring, was measured as the amplitude of the norm vector within each frame. Finally, frame-to-frame motion, which may cause sudden changes in the time–activity curve, was measured as the mean of the norm from the previous frame position (transmission scan position for the first frame). Distance from the TX position, the within-frame motion, and the frame-to-frame motion for the test and retest scans were 1.11 ± 0.58, 0.46 ± 0.47, 0.21 ± 0.25 mm versus 1.09 ± 0.56, 0.47 ± 0.46, and 0.21 ± 0.19 mm, respectively (mean ± s.d.),

Table 1 [11C]raclopride injections and their test–retest characteristics

[11C]raclopride injection	*Scan 1*		*Scan 2*		*Both scans*	*VAR (%)*	*(Scan 2−Scan 1)/ Scan 1*	*t-Test sign.*	*ICC*
	M ± s.d.	*COV (%)*	*M ± s.d.*	*COV (%)*	*M ± s.d.*	*M ± s.d.*	*Range*		
Injected dose (MBq)	413 ± 91	22.03	393 ± 99	25.26	403 ± 92	19.87 ± 11.62	−0.32–0.25	0.56	0.60
Specific radioactivity (MBq/nmol)	222 ± 154	69.40	257 ± 142	55.32	240 ± 144	42.89 ± 53.98	−0.24–1.56	0.15	0.91
Injected mass (μg)	2.47 ± 4.68[a]	189.60	0.68 ± 0.35	51.86	1.57 ± 3.32	43.41 ± 56.44	−1.65–0.24	0.33	0.11

COV, coefficient of variation; ICC, intraclass correlation coefficient; M ± s.d., mean ± standard deviation; VAR, absolute variability.
[a]One outlier (case 7, injected mass 13.03 μg).

and there were no systematic differences in motion between the test and retest scans. Thus, the amount of motion was negligible and in average tolerable compared with scanner resolution. Frame-to-frame coregistration was deemed unnecessary based on external motion recordings.

Quantification of [11C]raclopride Binding

All images were corrected for attenuation, scattering, and random events, scanner dead time, and detector normalization, and finally calibrated to kBq/mL within decay correction. Prompt and random events, attenuation, and scatter correction and normalization factors were processed in 3D sinogram mode, with an axial compression of span 9 and a maximum ring difference of 67, resulting in 2,209 sinograms in 16 segments. Tissue attenuation maps were reconstructed using the maximum a posteriori for transmission data (MAP-TR) algorithm with the standard human brain priors for air, bone, noise, water, or soft tissue (Nuyts *et al*, 1999). The 3D attenuation correction sinograms were obtained by forward projecting the tissue attenuation maps. Scattered events were estimated using the single scatter simulation algorithm (Watson, 2000), and randoms were estimated from the block singles using a variance reduction algorithm (Byars *et al*, 2005).

The [11C]raclopride uptake was measured continuously for 55 minutes after the injection. Histogramming was made in the following sequence of frames: 2×30, 9×60, 3×120, 3×180, and 6×300 seconds, resulting in altogether 23 frames. Image estimates with isotropic voxel dimensions of $1.22\,\text{mm} \times 1.22\,\text{mm} \times 1.22\,\text{mm}$ were generated using a speed-optimized version of ordinary poisson ordered subsets expectation maximization (OP-OSEM-3D) reconstruction with 16 subsets and 8 iterations, if not otherwise mentioned (Hong *et al*, 2007).

The [11C]raclopride images were processed with SPM2 in Matlab 7.0.4 as follows. First, to correct the head positioning between two PET scans, PET-to-PET realigment was made between the sum images. Second, for anatomic reference, T1-weighted magnetic resonance images were coregistered to a mean image of the realigned PET images. Regions of interest were drawn using Imadeus software, version 1.2 (Forima, Turku, Finland) on coregistered magnetic resonance images. Thus, an identical set of ROIs was applied for both [11C]raclopride scans of the same subject. The regions of interest were defined in a transaxial plane on the caudate nucleus, the putamen, the VST, the cerebellum, and the thalamus. We divided the thalamus into a medial and a lateral part because there is a gradient in D_2 receptor distribution between these sections (Kessler *et al*, 1993). In addition, VST was drawn in the coronal plane according to Mawlawi *et al* (2001) (Figure 1).

The mean activity concentration values in the ROIs were calculated and used to generate the regional time–activity curves. No partial-volume correction was made to the mean ROI values. These time–activity curves were used to calculate the regional [11C]raclopride binding potential (BP_{ND}) using the simplified reference tissue model with the cerebellum as reference tissue (Lammertsma and Hume, 1996). In this model, it is assumed that free plus nonspecific and specific binding equilibrium are reached so rapidly that they can be treated as a single compartment. Binding potential is defined according to the following equation: $BP_{ND} = f_{ND} B_{Avail}(1/K_D)$, where B_{Avail} is the total receptor concentration in tissue, $1/K_D$ is the apparent affinity, and f_{ND} is the fraction of free radiotracer in the nondisplaceable tissue compartment (Innis *et al*, 2007). In the model fitting, the data points were weighted with the estimate of the inverse variance for each sample.

Statistical Analyses

Test–retest variability was evaluated as follows:

$$\text{VAR} = \frac{2|T - \text{RT}|}{T + \text{RT}}$$

where T is the first (test) estimate and RT is the second (retest) estimate. We define VAR(%) = 100% VAR.

Test–retest consistency was measured with an intraclass correlation coefficient (ICC). Intraclass correlation coefficient can have a value between −1 and 1. Values closer to 1 indicate that most of the variance is due to between-subject variation rather than within-subject variation, which indicates good reliability, whereas values under 0 indicate greater within-subject than between-subject variation, indicating poor consistency. Values are expressed as mean ± s.d.

Paired two-tailed *t*-test was applied to the injected doses, specific radioactivities, injected masses, and BP_{ND} estimates between the test and retest scans and additionally

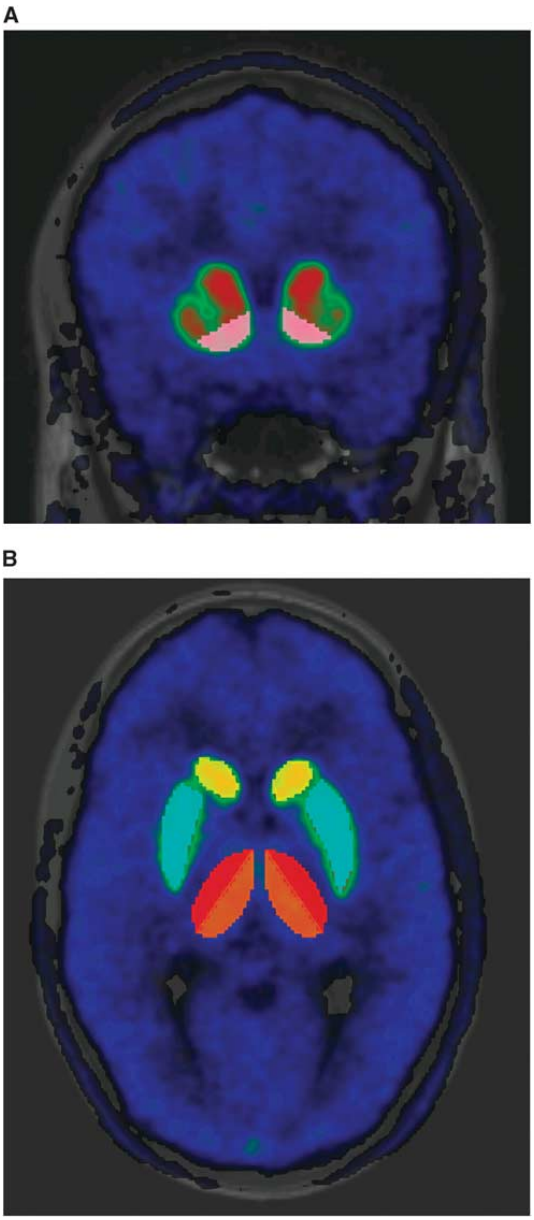


Figure 1 Activity distribution of [11C]raclopride in coronal view (**A**) and in axial plane (**B**). A mean image of positron emission tomography (PET)-to-PET coregistered summed PET images is overlaid on a T1-weighted magnetic resonance image of one subject at the level of the striatum. Regions of interest are defined on the caudatus, putamen, thalamus, and the ventral striatum.

for the BP_{ND} between the hemispheres. Moreover, the paired *t*-test was applied to the areas under the curve (AUC) (see below), BP_{ND} estimates, and VAR values between the reconstructions, so that the quantification method described above was always the reference.

Image Reconstruction With Fixed Counting Statistics

Conventionally, in the histogram mode image reconstruction, discrete time points are fixed in the sinogram space while counting statistics are let to vary. This procedure originated from old generation scanners that were not able to record all events as a stream in the acquisition disk. Modern scanners such as the HRRT record all events in a so-called list mode format, which allows us to determine framing after scanning, based on the number of acquired events for instance.

In PET, it is convenient to treat the acquired events as so-called noise-equivalent counts (NECs).

$$\mathrm{NEC} = \frac{T^2}{T + S + 2R}$$

where T denotes the number of trues, S the number of scattered, and R the number of random events. In conventional framing with fixed time points, the NEC per frame was $4.15 \pm 1.72 \times 10^6$ (mean ± s.d.) in the range of $\sim 10^6$ to $\sim 10^7$ counts, except for the first frame. A new set of sinograms with more homogeneous noise characteristics was created by fixing the minimum counts at 7×10^6 NEC per frame. NEC-based framing yielded 10 to 17 frames per scan, depending on the injected dose and the subject weight. Furthermore, parts of each scan were rebinned into frames of extremely low count statistics, yielding 5×10^5 to 2×10^6 NECs in 268 frames.

The rebinned sinogram data were reconstructed using the method described above. The head positioning between repeated PET scans was corrected using the coregistration parameters estimated from the original data, and thus the same ROIs were applied to yield new time–activity curves. Areas under the curve were calculated from both the original and the new sets of time–activity data and relative changes from original data were calculated as follows:

$$\Delta\mathrm{AUC} = \frac{\int_{t_1}^{t_2} f(t)\mathrm{d}t - \int_{t_1}^{t_2} f_0(t)\mathrm{d}t}{\int_{t_1}^{t_2} f_0(t)\mathrm{d}t}$$

where f_0 is the time–activity course estimated with the conventional framing. Integrals of the discrete time–activity course data were estimated using an in-house software with linear interpolation. We define $\Delta\mathrm{AUC}(\%) = 100\%$ ΔAUC. For the second full set of image data the BPs were reestimated as descibed above.

Image Reconstruction With Resolution Modeling

Additional image data were generated using a resolution model within the statistical reconstruction (Sureau *et al*, 2008). Resolution modeling exploits a measured point-spread function (PSF) of the scanner, which is assumed to be space invariant within the field of view. Implementation of the algorithm into OP-OSEM-3D has been described

earlier (Comtat *et al*, 2008). The same PSF parameters, iteration number (10), and number of subsets (16) were used as in an earlier paper (Comtat *et al*, 2008). The same PET-to-PET coregistration parameters and ROIs as with the original data were applied, and the resulting time–activity course data were evaluated using the ΔAUC defined above. The BPs were reestimated using the above described method.

Results

The mean BP_{ND} values of the test and retest scans ranged from 0.65 ± 0.04 to 4.41 ± 0.47 with the conventional framing using the OP-OSEM-3D reconstruction (Table 2). The BP_{ND} was highest in the putamen, followed by the caudate nucleus and the VST. The BP values were lowest in the thalamus. Thalamic BP_{ND} values were nearly identical for the medial and lateral sections (BP_{ND} values 0.65 ± 0.04 and 0.67 ± 0.04, respectively). No noticeable lateralization of the [11C]raclopride binding was observed (*t*-test, $P = 0.18$ to 0.98 for test, and $P = 0.08$ to 0.94 for retest). Moreover, the VAR values were nearly identical in both the left and the right hemisphere (*t*-test, $P = 0.30$ to 0.93). There was an increase in the mean of BP_{ND} values from the test to the retest scans in all regions except the VST. However, the increase was statistically significant only in the lateral (*t*-test, $P = 0.02$) and medial thalamus (*t*-test, $P = 0.01$). The mean left and right hemisphere values were used for the rest of the analyses.

Variability and Consistency of the Test and Retest Scans

The test–retest VAR and ICC values for the combined ROIs are presented in Table 2. Mean VAR (%) values were 3.7% to 4.7% for the caudate nucleus and the putamen, and 4.6% to 6.7% for the lateral and medial part of the thalamus, indicating fairly good reproducibility in these regions. The BP_{ND} values measured from the putamen yielded high ICC values both for the separate hemispheres and for the combined hemispheres (0.87 for the combined putamen). In the caudate nucleus, the hemispheric ICC values varied considerably and the ICC value for the combined ROI was only moderate (0.59). Thalamic BP_{ND} values resulted in moderate ICC values of 0.44 for the lateral part and 0.66 for the medial part. For the VST, the ROIs defined in the coronal plane yielded higher BP_{ND} values than the ROIs drawn in the axial plane (3.42 ± 0.23 versus 3.22 ± 0.32, respectively). In addition, the VAR (5.9%) and the ICC (0.50) values were superior for the coronally defined VST ROIs as compared with those drawn on the axial plane (9.7% and 0.34, respectively).

Image Reconstruction With Fixed Counting Statistics

Figure 2 shows a typical set of time–activity course data obtained with the different framings and the relative changes in the AUC (ΔAUC(%)) are presented in Table 3.

Regional ΔAUC(%) values have nonzero means for both the high and low statistics. High counting statistics yielded a positive mean in both the scans in the caudate nucleus (1.39 ± 1.14 and 1.65 ± 1.29),

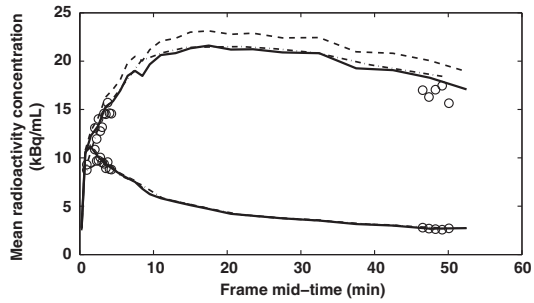


Figure 2 A typical example of regional time–activity course data in the putamen and the cerebellum estimated by the different framings and the point-spread function (PSF) reconstruction. Solid line is for original framing with non-PSF reconstruction, dashdot line is for high statistics framing, dashed line is for PSF reconstruction, and circles are for low statistics framing.

Table 2 Regional [11C]raclopride BP_{ND} values and their test–retest characteristics

Region	*Scan 1 (BP_{ND})*		*Scan 2 (BP_{ND})*		*Both scans (BP_{ND})*	*VAR (%)*	*100 × (Scan 2− Scan 1)/Scan 1*	*t-Test sign.*	*ICC*
	M ± s.d.	*COV (%)*	*M ± s.d.*	*COV (%)*	*M ± s.d.*	*M ± s.d.*	*Range*		
Putamen	4.33 ± 0.45	10.32	4.49 ± 0.51	11.36	4.41 ± 0.47	4.70 ± 3.91	−2.76–12.08	0.104	0.87
Caudate nucleus	3.81 ± 0.22	5.83	3.88 ± 0.23	6.05	3.85 ± 0.22	3.71 ± 4.72	−4.24–12.37	0.402	0.59
Ventral striatum (cor)	3.40 ± 0.28	8.34	3.44 ± 0.26	7.51	3.42 ± 0.26	5.93 ± 5.34	−8.84–15.37	0.731	0.50
Ventral striatum (ax)	3.27 ± 0.43	13.13	3.17 ± 0.36	11.41	3.22 ± 0.38	9.72 ± 9.67	−20.06–24.96	0.599	0.34
Medial thalamus	0.65 ± 0.03	5.03	0.68 ± 0.05	7.12	0.67 ± 0.04	4.59 ± 3.22	0.49–9.85	0.010	0.66
Lateral thalamus	0.63 ± 0.05	7.27	0.67 ± 0.04	5.67	0.65 ± 0.04	6.69 ± 4.82	−1.77–11.51	0.024	0.44

BP_{ND}, binding potential according to the simplified reference tissue model; COV, coefficient of variation; ICC, intraclass correlation coefficient; M ± s.d., mean ± standard deviation; VAR, absolute variability.

Table 3 Regional ΔAUC(%) values in the combined ROIs, for high statistics, low statistics initial part, low statistics end part, and PSF reconstruction

Region	*Scan 1 ΔAUC(%), M ± s.d.*				*Scan 2 ΔAUC(%), M ± s.d.*			
	High	*Low start*	*Low end*	*PSF*	*High*	*Low start*	*Low end*	*PSF*
Putamen	1.36 ± 1.04†	−2.52 ± 2.21	−6.21 ± 3.56‡	7.06 ± 1.02‡	1.48 ± 1.08‡	−2.80 ± 2.48	−6.43 ± 4.18‡	7.50 ± 0.92‡
Caudate nucleus	1.39 ± 1.14†	−3.85 ± 2.91‡	−6.66 ± 4.31‡	8.29 ± 1.17‡	1.65 ± 1.29‡	−3.13 ± 2.69	−5.90 ± 5.22‡	8.66 ± 0.86‡
Ventral striatum (cor)	0.88 ± 1.03	−1.43 ± 1.31‡	−5.04 ± 3.20‡	5.58 ± 0.53‡	1.12 ± 0.97†	−0.08 ± 1.35	−6.04 ± 3.93‡	6.16 ± 0.82‡
Ventral striatum (ax)	0.89 ± 0.99	−1.66 ± 1.91	−6.42 ± 4.54‡	6.45 ± 0.94‡	1.05 ± 1.08	−0.51 ± 1.23	−6.36 ± 4.62‡	6.94 ± 0.98‡
Medial thalamus	−0.05 ± 1.27	0.19 ± 1.26	3.27 ± 3.77	1.53 ± 2.11	0.00 ± 1.50	−0.52 ± 1.42	4.65 ± 4.14‡	2.19 ± 0.39‡
Lateral thalamus	−0.74 ± 1.31	1.99 ± 1.51†	5.70 ± 3.07‡	−1.41 ± 2.25	−0.93 ± 0.92	2.24 ± 1.47‡	4.88 ± 3.13‡	−0.77 ± 0.86†
Cerebellum	0.16 ± 0.53	−1.65 ± 0.74‡	0.09 ± 2.91	0.78 ± 1.70	0.17 ± 0.59	−1.53 ± 1.35†	1.63 ± 1.95†	1.04 ± 0.74‡

AUC, area under the curve; M ± s.d., mean ± standard deviation; PSF, point-spread function; ROI, region of interest.
†Statistically significant difference as compared with original framing with $P < 0.05$ in Student's *t*-test.
‡Statistically significant difference as compared with original framing with $P < 0.01$ in Student's *t*-test.

Table 4 Regional [11C]raclopride BP_{ND} values with high counting statistics reconstructions and their test–retest characteristics

Region	*Scan 1 (BP_{ND})*		*Scan 2 (BP_{ND})*		*Both scans (BP_{ND})*	*VAR (%)*	*100 × (Scan 2− Scan1)/Scan 1*	*t-Test sign.*	*ICC*
	M ± s.d.	*COV (%)*	*M ± s.d.*	*COV (%)*	*M ± s.d.*	*M ± s.d.*	*Range*		
Putamen	4.46 ± 0.49‡	10.89	4.62 ± 0.54‡	11.72	4.54 ± 0.50	5.03 ± 3.84	−2.93–12.27	0.138	0.87
Caudate nucleus	3.92 ± 0.24‡	6.16	3.99 ± 0.23‡	5.88	3.96 ± 0.23	4.00 ± 4.70	−3.95–12.37	0.430	0.57
Ventral striatum (cor)	3.47 ± 0.27‡	7.84	3.52 ± 0.28‡	7.86	3.49 ± 0.26	5.65 ± 5.73	−7.34–16.22	0.674	0.51
Ventral striatum (ax)	3.34 ± 0.42‡	12.45	3.25 ± 0.37‡	11.51	3.29 ± 0.38	9.53 ± 9.90	−19.37–25.45	0.639	0.31
Medial thalamus	0.65 ± 0.04	6.82	0.68 ± 0.06	8.59	0.67 ± 0.05	4.51 ± 3.10	1.94–9.32	0.011	0.78
Lateral thalamus	0.62 ± 0.05	7.85	0.65 ± 0.04†	5.69	0.64 ± 0.04	6.81 ± 3.63	−3.82–12.04	0.077	0.51

BP_{ND}, binding potential according to the simplified reference tissue model; COV, coefficient of variation; ICC, intraclass correlation coefficient; M ± s.d., mean ± standard deviation; VAR, absolute variability.
†Statistically significant difference as compared with original framing with $P < 0.05$ in Student's *t*-test.
‡Statistically significant difference as compared with original framing with $P < 0.01$ in Student's *t*-test.

in the putamen (1.36 ± 1.04 and 1.48 ± 1.08), in the VST (transaxial ROI, 0.89 ± 0.99 and 1.05 ± 1.08), and in the cerebellum (0.16 ± 0.53 and 0.17 ± 0.59). However, negative (or zero) means were observed in both the scans in the medial (−0.05 ± 1.27 and 0.00 ± 1.50) and in the lateral parts of the thalamus (−0.74 ± 1.31 and −0.93 ± 0.92). Thus, high counting statistics yielded larger AUC estimates as compared with the original in the regions with high radioactivity concentration (RC) and smaller AUC estimates in the regions with low RC. For the low statistics, the behavior of the mean ΔAUC(%) was inversed: the caudate nucleus and the putamen had negative means in both the scans in the beginning and at the end of the scan, whereas the thalamus had positive means in all cases except for the start of the second scan. In the cerebellum, the mean ΔAUC(%) was −1.65 ± 0.74 and −1.53 ± 1.35 in the beginning of the scan but 0.72 ± 3.02 and 1.49 ± 1.94 at the end of the scan. However small, the direction of the change in the AUC is clear—high counting statistics yielded larger estimates in the high RC regions and smaller estimates in the low RC regions as compared with low counting statistics. Although statistically significant in the paired *t*-test, the amount of change was minimal between the original and the high statistics framing, and between the first part of the low statistics framing and the original framing. As regards the end part with low statistics, the bias (ΔAUC(%)) was noticeably larger in both the scans, with mean values of −6.66 ± 4.31 and −5.90 ± 5.22 for the caudate nucleus, and 4.65 ± 4.14 and 3.96 ± 3.65 for the medial part of the thalamus.

BP_{ND} estimates and the VAR and ICC values for combined ROIs with the high statistics framing are presented in Table 4. Changes in the mean BP_{ND} were positive in the caudate nucleus (2.86%), in the putamen (2.95%), in the axially drawn VST (2.17%), and in the coronally drawn VST (2.05%), and negative or zero in the lateral (−1.59%) and medial parts of thalamus (0.00%) as compared with the conventional framing. In the VAR and ICC, the changes were more ambiguous, as high counting statistics yielded higher VAR in the caudate nucleus, the putamen, and in the lateral part of thalamus, but lower VAR in the medial part of thalamus and in the VST. None of the regions had significantly different mean VAR as compared with the original framing.

Table 5 Regional [11C]raclopride BP_{ND} values with PSF reconstruction and their test–retest characteristics

Region	*Scan 1 (BP_{ND})*		*Scan 2 (BP_{ND})*		*Both scans (BP_{ND})*	*VAR (%)*	*100 × (Scan 2− Scan 1)/Scan 1*	*t-Test sign.*	*ICC*
	M ± s.d.	*COV (%)*	*M ± s.d.*	*COV (%)*	*M ± s.d.*	*M ± s.d.*	*Range*		
Putamen	4.75 ± 0.49‡	10.41	4.93 ± 0.57‡	11.67	4.84 ± 0.52	5.76 ± 4.32	−3.90–14.08	0.158	0.82
Caudate nucleus	4.21 ± 0.24‡	5.65	4.29 ± 0.25‡	5.92	4.25 ± 0.24	4.78 ± 5.44	−4.68–15.26	0.497	0.35
Ventral striatum (cor)	3.65 ± 0.27‡	7.48	3.71 ± 0.27‡	7.37	3.68 ± 0.26	5.56 ± 5.57	−6.83–14.77	0.601	0.48
Ventral striatum (ax)	3.57 ± 0.45‡	12.7	3.46 ± 0.38‡	10.96	3.52 ± 0.41	9.17 ± 9.60	−20.18–23.31	0.587	0.33
Medial thalamus	0.66 ± 0.04	6.28	0.70 ± 0.06†	8.92	0.68 ± 0.06	6.07 ± 3.88	0.71–10.29	0.007	0.64
Lateral thalamus	0.59 ± 0.06‡	9.55	0.63 ± 0.05‡	7.54	0.61 ± 0.05	9.09 ± 3.37	−5.09–13.77	0.021	0.52

BP_{ND}, binding potential according to the simplified reference tissue model; COV, coefficient of variation; ICC, intraclass correlation coefficient; M ± s.d., mean ± standard deviation; PSF, point-spread function; VAR, absolute variability.
†Statistically significant difference as compared with original framing with $P < 0.05$ in Student's *t*-test.
‡Statistically significant difference as compared with original framing with $P < 0.01$ in Student's *t*-test.

The high counting statistics improved the ICC in the thalamus and in the coronally defined VST, whereas in the other regions, the change was negative or zero.

Image Reconstruction With Resolution Modeling

In Figure 2, a typical set of time–activity course data obtained with the PSF reconstruction along with the different framings using the non-PSF reconstruction is seen. Regional ΔAUC(%) values for PSF reconstruction are presented in Table 3.

Regional ΔAUC(%) values have positive means in both the scans except for the lateral thalamus (−1.41 ± 2.25 and −0.77 ± 0.86), with largest increases in the caudate nucleus (8.29 ± 1.17 and 8.66 ± 0.86) and the putamen (7.06 ± 1.02 and 7.50 ± 0.92), whereas the smallest changes were seen in the cerebellum (0.78 ± 1.70 and 1.04 ± 0.74).

The BP_{ND} estimates, the mean VAR, and the ICC values for combined ROIs with the PSF reconstruction are presented in Table 5. The PSF reconstruction yielded higher BP_{ND} estimates in all other regions expect the lateral thalamus when compared with the original framing using the non-PSF reconstruction. Increases were largest in the caudate nucleus and in the putamen. However, the increased BP_{ND} estimates were coupled with increased VAR and thereafter decreased ICC in many of the regions. In striatal regions, the largest increase in the VAR was seen in the caudate nucleus and in the putamen, whereas the VAR was somewhat smaller in the VST. The ICC values were decreased in all regions except the lateral thalamus, with largest decrease in the caudate nucleus.

Discussion

In this study, we set out to evaluate the test–retest reliability of the striatal, ventrostriatal, and thalamic D_2 binding using [11C]raclopride and the HRRT scanner. The reproducibility of striatal [11C]raclopride binding has been explored earlier in various studies (Volkow *et al*, 1993; Hietala *et al*, 1999; Mawlawi *et al*, 2001; Hirvonen *et al*, 2003), but not with a high-resolution tomograph. In this study, with the conventional framing and non-PSF reconstruction, the VAR values were 3.7% ± 4.7% in the caudate nucleus, 4.7% ± 3.9% in the putamen, and 5.9% ± 5.3% in the VST (coronal plane). These values are comparable to those reported in an earlier study in which striatal and ventrostriatal D_2 receptors were quantified (5.3% ± 5.2% for the caudate nucleus, 3.5% ± 3.9% for the putamen, and 8.6% ± 6.2% for the VST) (Mawlawi *et al*, 2001). However, the ICC values from our study were slighty lower for every region except for the putamen. This difference may, at least partly, be due to methodological differences such as the bolus plus infusion method used in the previous study (Mawlawi *et al*, 2001).

A previous study evaluating the reproducibility and validity of the thalamic D_2 receptor binding measurements in a test–retest setting found the VAR values to be 8.5% ± 6.1% for the medial part and 8.9% ± 5.3% for the lateral part of the thalamus (Hirvonen *et al*, 2003). In this study, we obtained slightly superior VAR values, with 4.5% ± 3.1% for the medial and 6.4% ± 4.5% for the lateral part of the thalamus. The earlier results (Hirvonen *et al*, 2003) for the caudate nucleus (4.0% ± 2.2%) and for the putamen (5.3% ± 3.4%) are well in line with this study. Nonetheless, except for the putamen (0.87), the ICC values were less than those reported earlier for the caudate nucleus (0.96), the putamen (0.84), the medial thalamus (0.84), and the lateral thalamus (0.86) (Hirvonen *et al*, 2003). The most likely explanation to the lower ICC values in this study is the small between-subject variability in our study group (see Table 2, coefficient of variation (%)) when compared with the earlier study (Hirvonen *et al*, 2003). Altogether, the VAR values in our study indicate good reproducibility and are well in line with the previous studies.

In this study, we observed an increase in the mean BP_{ND} values from the test to the retest scans in all regions except the VST, which might also impair

the ICC. However, the increase was statistically significant only in the lateral (*t*-test, $P = 0.02$) and medial (*t*-test, $P = 0.01$) thalamus. There were no statistically significant differences either in the specific radioactivities or in the injected masses between the test and retest scans. However, one case was an outlier in the analysis of the injected masses. Theoretically, low specific activity of [11C]raclopride would yield underestimation of the BP_{ND}. Exclusion of the low specific radioactivity case from the analysis decreased the difference between mean BP_{ND} in the test and retest scans, yielding slightly superior VAR and ICC values (data not shown).

This is the first study measuring the reproducibility of [11C]raclopride binding using the HRRT. There are some previous reports with other radioligands such as [18F]-FDG (fluorodeoxyglucose) showing that the better spatial resolution of the HRRT scanner decreased the PVE when determining metabolic rates in small brain nuclei (Heiss *et al*, 2004). Dopamine transporter binding was measured using the selective dopamine transporter radioligand [11C]-PE2I, which resulted in a significant increase (30%) in striatal BP values when using the HRRT versus a more conventional scanner (Leroy *et al*, 2007), and a recent study reported a good test–retest reliability of dopamine transporter binding using [11C]-PE2I with the HRRT (Hirvonen *et al*, 2008). Accordingly, we also discovered a significant benefit from the better spatial resolution of the HRRT scanner in noticeably higher BP_{ND} values compared with previous [11C]raclopride studies with more conventional scanners. For example, we obtained BP_{ND} values 20% higher for the caudate nucleus, the putamen, and for the thalamus than the BP_{ND} values reported using a lower-resolution scanner (Hirvonen *et al*, 2003).

The measured [11C]raclopride BPs in this study were high throughout the whole striatum, ranging from 3.22 (VST) to 4.41 (putamen), and moderate in the thalamus (0.65 to 0.67), using the conventional quantification methods. These results are in a good agreement with the previous *in vivo* studies (Mawlawi *et al*, 2001; Hirvonen *et al*, 2003) and also comparable with various *in vitro* studies (Hall *et al*, 1988, 1994). The BP_{ND} values found in the VST were 77% of that in the putamen, which is greater than the 63% reported earlier (Mawlawi *et al*, 2001). This could be explained by the increased signal because of the better spatial resolution of the HRRT scanner. Although BP_{ND} was higher in the VST, the VAR value ($5.93\% \pm 5.34\%$) in the VST (coronal) did not significantly differ from those observed in the larger caudate nucleus and putamen ($3.71\% \pm 4.72\%$ and $4.70\% \pm 3.91\%$, respectively). Furthermore, the BP_{ND} values were higher and the reproducibility was better when delineating VST ROI in the coronal plane than in the axial plane. This is probably because of a more accurate anatomic delineation of the VST in the coronal than in the axial plane.

Instrumentational and Signal Processing Considerations

The small size of the VST renders the signal highly vulnerable to head movement. Therefore, in our study we used individually shaped thermoplastic mask with each subject to minimize head movement. The external motion detector (Polaris Vicra) used in this study showed very little motion during the scans, despite a trend toward increased movement during the course of the scan. Even when being greatest, the movement was less than the spatial resolution of the scanner, thus no movement correction was applied as this potential confound was not obviously affecting the results.

It has been reported that the HRRT image reconstruction suffers from quantification bias in low count statistics, such as we saw in this study (Johansson *et al*, 2007). Despite the development of algorithms such as variance reduction on randoms (Byars *et al*, 2005) and ordinary poisson OSEM reconstruction (Comtat *et al*, 2004), a recent study still reported a noticeable bias in statistical reconstruction as compared with analytical reconstruction in low count statistics (van Velden *et al*, 2008). However, analytical image reconstruction is not optimal for the HRRT image reconstruction because of gaps in the sinogram data. In this study, we implemented a novel approach in PET data sampling to overcome the bias issue. Our aim was to collect a fixed number of counts in each frame, which is made possible by the list mode event recording. Noise-equivalent count was applied because it takes into account the varying randoms rates and scatter fractions. We extrapolated an NEC threshold per frame from a previous phantom study in which we studied the quantification bias in low statistics (Johansson *et al*, 2007). In the conventional framing, the threshold of 7×10^6 NECs was not reached during the first 600 seconds of the scan because of dense framing depicting the bolus peak. After that NECs per frame varied below and above the threshold, depending on the subject mass and injected dose. Absolute difference in the injected doses between the test and retest scans ranged between 3 and 113 MBq, with the mean injected dose of 403 MBq. Scans with lowest dosage may yield lower BP_{ND} as compared with their counterpart, which would impair the ICC and VAR values. Our NEC-based framing was aimed at removing this biasing effect if it was present in this study. First, to evaluate the effect of the NEC-based framing the bias was estimated directly from the time-activity course (TAC) data (Table 3). It was concluded that the NEC-based framing yielded higher contrast images as compared with the original framing, with statistically significant difference in the high RC regions. Second, the BP_{ND} estimates were recalculated from the NEC-based data and the ICC and VAR values were estimated (Table 4). As expected, the observed changes in the AUC were transferred into BP_{ND} so that in the high RC regions the BP_{ND} estimates were

higher as compared with the conventional framing. However, there was no improvement either in the ICC or in the VAR values, which could be expected because of relatively small within-subject differences in injected doses (Table 1). To study the bias further, we used NEC-based framing with thresholds of 5×10^5 to 2×10^6 NECs per frame. Parts of the data were histogrammed and reconstructed with the low count statistics, and bias was estimated directly from the TAC data (Table 3). Now, a statistically significant decrease in the image contrast was seen in the high and low RC regions when the NEC-based framing was compared with the original framing. The BP_{ND} estimates were not calculated for low statistics data because there were only a few samples from the beginning and end of each scan. Our observations add evidence to the claim that the current statistical reconstruction suffers from bias, which has significant magnitude, for instance, in follow-up studies if the injected doses vary substantially between scans. Thus, to cope with the bias the injected doses should be equalized, if not succeeded, the NEC-based framing should be used to maintain similar conditions for the statistical reconstruction. Ultimately, new bias-free reconstruction schemes must be explored to remove the count statistics limitation in the HRRT studies.

Detector design of the HRRT aims at removing the PVE from the PET measurement. Conventionally, the PVE is corrected from the collected data and a variety of algorithms have been proposed. An implementation of a scanner-specific PSF-based method was recently made for the HRRT and delivered to the HRRT users (Comtat *et al*, 2008). It has been reported that the PSF-based method is able to generate higher BP estimates as compared with the conventional reconstruction (Sureau *et al*, 2008; Comtat *et al*, 2008; Mourik *et al*, 2010; Varrone *et al*, 2009). Approximately, 25% higher BP_{ND} estimates in the caudate nucleus, 23% in the putamen, and 10% in the VST using [11C]-PE2I with listmode reconstruction and PSF modeling has been reported (Sureau *et al*, 2008). In a study including one human subject imaged with [11C]-raclopride $\sim$10% higher BP_{ND} estimates in the caudate nucleus and 6% higher in the putamen were obtained (Varrone *et al*, 2009). In a study using [11C]-flumazenil in five subjects no noticeably higher distribution volume estimates with PSF as compared with non-PSF reconstruction were obtained (Mourik *et al*, 2010). The authors concluded that the improvement obtained through the usage of PSF reconstruction is tracer dependent. In this study, we obtained $\sim$10% increase in the BP_{ND} estimates in the caudate nucleus and putamen and 8% increase in the VST, which are in good agreement with an earlier report (Varrone *et al*, 2009). Moreover, we obtained $\sim$ 7% lower BP_{ND} estimate in the lateral thalamus, and $<$2% increase in the medial thalamus, which would suggest that the PSF reconstruction alleviates the spill-in effect as well. In this study, our main objective was to evaluate the methods with regard to measurement reproducibility. Although PSF reconstruction provided BP_{ND} estimates that should be less contaminated by the PVE, its VAR values and, consequently, the ICC values were somewhat inferior to the non-PSF reconstruction. In the thalamus and VST, the VAR and ICC values were identical between the PSF and the non-PSF reconstructions.

As the line-of-response (LOR) of the HRRT is significantly thinner as compared with conventional scanners, the counting statistics rather than the spatial resolution becomes a limiting factor in the quantitative HRRT studies. In this report, we conclude that the statistical reconstruction of the HRRT is vulnerable to bias in low count statistics, but also, that the PSF reconstruction rather increases than decreases the VAR values. However, the PSF reconstruction was recently evaluated with regard to count statistics (Walker *et al*, 2009) and it was found to be less vulnerable to low count statistics bias. In this study, we did not evaluate the PSF reconstruction with regard to count statistics, but we observed that both the increment of count statistics and the inclusion of the resolution model in the reconstruction increase the contrast of the images. In addition, the preceding evaluation using list mode reconstruction and the resolution model (Planeta-Wilson *et al*, 2008) supports the usage of PSF reconstruction in the low statistics HRRT studies.

Conclusions

This study showed that the use of single bolus [11C]raclopride and the high-resolution research scanner HRRT results in good reproducibility of quantification of D_2 receptor binding. The use of the HRRT scanner increased the signal from the striatum and the thalamus and resulted in higher BP_{ND} values compared with previous studies using scanners with lower resolution. The good test–retest reproducibility of the quantification of the VST is encouraging and shows the feasibility of the use of the HRRT in studying the dopaminergic system in small brain regions. Our results show that [11C]raclopride BP has good reproducibility even when using a high-resolution scanner that allows analysis of small brain structures.

Acknowledgements

The authors thank the personnel of Turku PET Centre for skillful assistance during the scanning.

Conflict of interest

The authors declare no conflict of interest.

References

Bäckman L, Ginovart N, Dixon RA, Wahlin TB, Wahlin A, Halldin C, Farde L (2000) Age-related cognitive deficits

mediated by changes in the striatal dopamine system. *Am J Psychiatry* 157:635–7

Belin D, Everitt BJ (2008) Cocaine seeking habits depend upondopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron* 57:432–41

Byars LG, Sibomana M, Burbar Z, Jones J, Panin V, Barker WC, Liow J-S, Carson RE, Michel C (2005) Variance reduction on randoms from coincidence histograms for the HRRT. *IEEE Nucl Sci Conf R* 5:2622–6

Comtat C, Bataille F, Michel C, Jones JP, Sibomana M, Janeiro L, Trebossen R (2004) OSEM-3D reconstruction strategies for the ECAT HRRT. *IEEE Nucl Sci Conf R* 6:3492–6

Comtat C, Sureau FC, Sibomana M, Hong IK, Sjoholm N, Trebossen R (2008) Image based resolution modeling for the HRRT OSEM reconstructions software. *IEEE Nucl Sci Conf R* 2008:4120–3

de Jong HWAM, van Velden FHP, Kloet RW, Buijs FL, Boellaard R, Lammertsma AA (2007) Performance evaluation of the ECAT HRRT: an LSO-LYSO double layer high resolution, high sensitivity scanner. *Phys Med Biol* 52:1505–26

DeLong MR (1983) The neurophysiologic basis of abnormal movements in basal ganglia disorders. *Neurobehav Toxicol Teratol* 5:611–6

Evarts EV, Wise SP (1984) Basal ganglia outputs and motor control. *Ciba Found Symp* 107:83–102

Floresco SB, Onge JRS, Ghods-Sharifi S, Winstanley CA (2008) Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. *Cogn Affect Behav Neurosci* 8:375–89

Hall H, Farde L, Sedvall G (1988) Human dopamine receptor subtypes—*in vitro* binding analysis using 3H-SCH 23390 and 3H-raclopride. *J Neural Transm* 73:7–21

Hall H, Sedvall G, Magnusson O, Kopp J, Halldin C, Farde L (1994) Distribution of D1- and D2-dopamine receptors, and dopamine and its metabolites in the human brain. *Neuropsychopharmacol* 11:245–56

Heiss W-D, Habedank B, Klein JC, Herholz K, Wienhard K, Lenox M, Nutt R (2004) Metabolic rates in small brain nuclei determined by high-resolution PET. *J Nucl Med* 45:1811–5

Hietala J, Någren K, Lehikoinen P, Ruotsalainen U, Syvälahti E (1999) Measurement of striatal D2 dopamine receptor density and affinity with [11C]-raclopride *in vivo*: a test-retest analysis. *J Cereb Blood Flow Metab* 19:210–7

Hirvonen J, Aalto S, Lumme V, Någren K, Kajander J, Vilkman H, Hagelberg N, Oikonen V, Hietala J (2003) Measurement of striatal and thalamic dopamine D2 receptor binding with 11C-raclopride. *Nucl Med Commun* 24:1207–14

Hirvonen J, Johansson J, Teräs M, Oikonen V, Lumme V, Virsu P, Roivainen A, Någren K, Halldin C, Farde L, Hietala J (2008) Measurement of striatal and extrastriatal dopamine transporter binding with high-resolution PET and [11C]PE2I: quantitative modeling and test-retest reproducibility. *J Cereb Blood Flow Metab* 28: 1059–69

Hong IK, Chung ST, Kim HK, Kim YB, Son YD, Cho ZH (2007) Ultra fast symmetry and SIMD-based projection-backprojection (SSP) algorithm for 3-D PET image reconstruction. *IEEE Trans Med Imaging* 26:789–803

Innis RB, Cunningham VJ, Delforge J, Fujita M, Gjedde A, Gunn RN, Holden J, Houle S, Huang S-C, Ichise M, Iida H, Ito H, Kimura Y, Koeppe RA, Knudsen GM, Knuuti J, Lammertsma AA, Laruelle M, Logan J, Maguire RP, Mintun MA, Morris ED, Parsey R, Price JC, Slifstein M, Sossi V, Suhara T, Votaw JR, Wong DF, Carson RE (2007) Consensus nomenclature for *in vivo* imaging of reversibly binding radioligands. *J Cereb Blood Flow Metab* 27:1533–9

Johansson J, Oikonen V, Teras M (2007) Quantitative brain imaging using the new, fast iterative histogram-mode reconstruction for the HRRT PET scanner. *IEEE Nucl Sci Conf R* 5:3463–7

Kaasinen V, Rinne JO (2002) Functional imaging studies of dopamine system and cognition in normalaging and Parkinson's disease. *Neurosci Biobehav Rev* 26: 785–93

Kessler RM, Whetsell WO, Ansari MS, Votaw JR, de Paulis T, Clanton JA, Schmidt DE, Mason NS, Manning RG (1993) Identification of extrastriatal dopamine D2 receptors in postmortem human brain with [125I]epidepride. *Brain Res* 609:237–43

Lammertsma AA, Hume SP (1996) Simplified reference tissue model for PET receptor studies. *Neuroimage* 4:153–8

Langer O, Någren K, Dollé F, Lundkvist C, Sandell J, Swahn C-G, Vaufrey F, Crouzel C, Mazière B, Halldin C (1999) Precursor synthesis and radiolabelling of the dopamine D2 receptor ligand [11C]raclopride from [11C]methyl triflate. *J Labelled Comp Radiopharm* 42:1183–93

Larsen P, Ulin J, Dahlström K, Jensen M (1997) Synthesis of [11C]methyl iodide by iodination of [11C]methane. *Appl Radiat Isot* 48:153–7

Leroy C, Comtat C, Trébossen R, Syrota A, Martinot J-L, Ribeiro M-J (2007) Assessment of 11C-PE2I binding to the neuronal dopamine transporter in humans with the high-spatial-resolution PET scanner HRRT. *J Nucl Med* 48:538–46

Mawlawi O, Martinez D, Slifstein M, Broft A, Chatterjee R, Hwang DR, Huang Y, Simpson N, Ngo K, Heertum RV, Laruelle M (2001) Imaging human mesolimbic dopamine transmission with positron emission tomography: I. Accuracy and precision of D(2) receptor parameter measurements in ventral striatum. *J Cereb Blood Flow Metab* 21:1034–57

Mourik JEM, Lubberink M, van Velden FHP, Kloet RW, van Berckel BNM, Lammertsma AA, Boellaard R (2010) *In vivo* validation of reconstruction-based resolution recovery for human brain studies. *J Cereb Blood Flow Metab* 30:381–9

Någren K, Truong P, Helin S, Amir A, Halldin C (2003) Experience from two systems for recirculating production of [11C]methyl iodide from target produced [11C]methane. *J Labelled Comp Radiopharm* 46 (Suppl. 1):S76

Nuyts J, Dupont P, Stroobants S, Maes A, Mortelmans L, Suetens P (1999) Evaluation of maximum-likelihood based attenuation correction in positron emission tomography. *IEEE Trans Nucl Sci* 46:1136–41

Planeta-Wilson B, Yan J, Mulnix T, Carson RE (2008) Quantitative accuracy of HRRT list-mode econstructions: effect of low statistics. *IEEE Nucl Sci Conf R* 2008:5121–4

Sureau FC, Reader AJ, Comtat C, Leroy C, Ribeiro M-J, Buvat I, Trébossen R (2008) Impact of image-space resolution modeling for studies with the high-resolution research tomograph. *J Nucl Med* 49:1000–8

van Velden FHP, Kloet RW, van Berckel BNM, Wolfensberger SPA, Lammertsma AA, Boellaard R (2008) Comparison of 3D-OP-OSEM and 3D-FBP reconstruction algorithms for High-Resolution Research Tomograph

studies: effects of randoms estimation methods. *Phys Med Biol* 53:3217–30

Varrone A, Sjöholm N, Eriksson L, Gulyás B, Halldin C, Farde L (2009) Advancement in PET quantification using 3D-OP-OSEM point spread function reconstruction with the HRRT. *Eur J Nucl Med Mol Imaging* 10:1639–50

Volkow ND, Fowler JS, Wang GJ, Dewey SL, Schlyer D, MacGregor R, Logan J, Alexoff D, Shea C, Hitzemann R (1993) Reproducibility of repeated measures of carbon-11-raclopride binding in the human brain. *J Nucl Med* 34:609–13

Walker MD, Julyan PJ, Talbot PS, Jones T, Matthews JC (2009) Bias in iterative reconstruction of low-statistics PET data: benefits of a resolution model. *IEEE Nucl Sci Conf R* 2009:2857–63

Watson CC (2000) New, faster, image-based scatter correction for 3D PET. *IEEE Trans Nucl Sci* 47:1587–94

Publication III

Johansson, J., Teuho, J., Linden, J., Tuna, U., Tolvanen, T., Saunavaara, V., Teräs, M., "Image quantification in high-resolution PET assessed with a new anthropomorphic brain phantom," *Nuclear Science Symposium Conference Record.*, 2013, NSS '13. IEEE.

Image Quantification in High-resolution PET Assessed with a New Anthropomorphic Brain Phantom

Jarkko Johansson, Jarmo Teuho, Jani Lindén, Uygar Tuna, Tuula Tolvanen, Virva Saunavaara, Mika Teräs, *Member, IEEE*

***Abstract*– Choice of the PET scanner and image reconstruction parameters have significant impact in quantitative positron-emission tomography (PET). Hoffman phantom is probably the most widely used test object for assessing this impact in brain PET studies. In high-resolution PET, however, it's usability is questionable due to lesser partial-volume effect. Futhermore, Hoffman phantom is cylindrical and does not offer realistic attenuation effect for the skull. In the current work we used a novel brain phantom that was produced using a 3D-printer, and provides realistic head contour and skull attenuation effect. We scanned the phantom with latest generation whole-body PET/MR (Philips Ingenuity TF) and PET/CT (GE Discovery 690) scanners and in a brain dedicated high-resolution scanner (Siemens HRRT) to evaluate it's usability for intra- and inter-scanner comparisons with regard to PET brain imaging.**

In all scanners reconstruction algorithm choice and number of iterations had significant impact on anatomical gray matter ROI values. As compared to the HRRT, whole-body scanners showed 3% to 15% (Philips Ingenuity TF) and 0% to 5% (GE D690) negative biases in gray matter ROIs, when iterative reconstruction with high number of iterations but without resolution modeling was used. Whereas, low number of iterations in Philips Ingenuity yielded negative biases of 7% to 19%, but inclusion of resolution modeling in GE D690 yielded 19% to 7% higher values. In the HRRT count statistics related negative bias of up to 6% was seen, when iterative reconstruction without resolution modeling was used.

We conclude that the new three-dimensional brain phantom is suitable for assessing the impact of reconstruction parameters both within and between scanners. However, the lack of ground truth values hampers the interpretation of the results, and furthermore, the small differences we saw between whole-body and brain-dedicated scanners might be due to limited resolution of the 3D-printing.

I. Introduction

IMAGE reconstruction parameters, such as attenuation and scatter correction methods and reconstruction algorithm have significant impact on the quantification of PET images. This impact is typically assessed using mathematical or real test objects. Hoffman phantom is probably the most widely used test object in PET brain studies, due to it's advantageous features. Hoffman phantom consists of over fifty, three millimeter thick axial slices that cover a human brain. Each slice is carved open where there should be gray matter and to one-fourth of total thickness where there should be white matter, yielding a 4:1 ratio between gray and white matter activity concentrations. Due to axial partial-volume effect (PVE) the resulting PET images seem continuos. High-resolution PET scanners, such as Siemens HRRT however have better axial resolution than three millimeters, hence producing visible slices in images. In Hoffman phantom slices are packed in a cylinder which does not correspond to the real head contour, and furthermore, the skull attenuation effect is missing. In the current work we used a recently developed brain phantom that was build from a three-dimensional (3D) model of the brain gray matter, skull and trachea using a 3D printer [1]. The gray matter compartment was designed to be filled with positron-emitting isotope in water solution and the skull compartment with a bone-equivalent solution of K_2HPO_4, and trachea was left empty (air). Our aim was to evaluate the usability of this novel phantom in high resolution PET for intra- and inter-scanner comparisons. Hence, we scanned the phantom with latest generation whole-body PET/MR (Philips Ingenuity TF) and PET/CT (GE Discovery 690) scanners and in a brain dedicated high-resolution scanner (Siemens HRRT). Measurements were collected in listmode format allowing us to generate histograms with varying count statistics retrospectively. Various different reconstruction protocols were applied for each scanner and results were compared within and between scanners, using quantitative ROI-based analysis and visual assessment.

II. Materials and Methods

A. Phantom design and construction

Phantom was designed and constructed in professor Hidehiro Iida's lab in National Cerebral and Cardiovascular Center Research Institute, Japan. A full report of the phantom construction and initial experiments can be found in [1]. In short, the phantom design is based on T1-weighted MR-images of a 26-year-old Japanese male volunteer, that were segmented for different brain tissues. Skull, trachea and gray matter compartments were modeled using computer-aided-design (CAD), and the CAD model was realized using laser-modeling technique (3D-printing). The skull and gray matter

Manuscript received November 22, 2013. This work was supported in part by theTurun Yliopistosäätiö travel grant from Kosti Hämmärö foundation.

J. Johansson, J. Teuho, T. Tolvanen, V. Saunavaara and M. Teräs are with theTurku PET Centre, Turku University Central Hospital and University of Turku, Kiinamyllynkatu 4-8, 20540 Turku, Finland (e-mail: firsname.lastname@tyks.fi).

J. Lindén is with the Department of Mathematics, Faculty of Natural Sciences, University of Turku, 20540 Turku, Finland.

U. Tuna was with the M2oBSI research group, Department of Signal Processing, Tampere University of Technology, Tampere, Finland.

compartments were designed so that they can be entirely filled from one opening each. The skull compartment is to be filled with bone-equivalent solution of K_2HPO_4, and the gray matter compartment with radioisotope in water solution. White matter, CSF, scalp and other tissue types were not modeled separately, thus they were all made of photo-curable polymer used in laser-modeling. Fig. 1 shows a transaxial slice of an X-ray CT scan of the phantom.

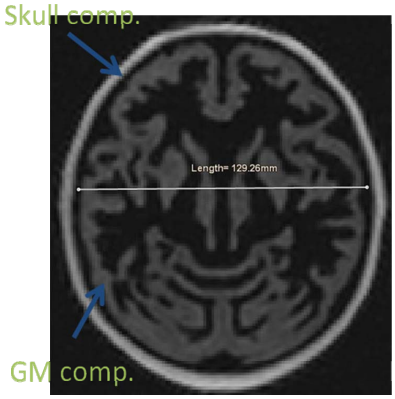


Fig. 1. Transaxial slice of an X-ray CT scan of the phantom at the level of basal ganglio. Skull compartment appears bright in the X-ray scan due to high density as compared to water or photocurable polymer.

B. *Phantom measurements*

The gray matter compartment was filled with ^{18}F-FDG in 550 mL water solution, and the skull compartment with K_2HPO_4 solution. The phantom was scanned for 15 to 20 minutes in Philips Ingenuity TF PET/MR (henceforth PhITF), in Siemens HRRT and finally in GE Discovery 690 PET/CT (henceforth GED690) with initial activities of 82, 69 and 48 MBq, respectively. All emission acquisitions were made in listmode format, and histograms from whole scan duration and from frames of 24 to 47 seconds were generated retrospectively. Short duration replicates were made to mimic dynamic scanning conditions, and the frame lengths were chosen according to noise-equivalent counts (NEC) (see [2] for formula) when available. Count statistics are summarizes in Table 1. In PhITF and GED690 the time-of-flight (ToF) information was collected. Attenuation correction measurements were made using MR (PhITF), CT (GED690) and ^{137}Cs point-source (HRRT).

TABLE I. COUNT STATISTICS

Scanner	Act (MBq)	Total NEC	NEC/frame	Frames
PhITF	81.7	2.1 x 10^8	changing	35 x 24 sec.
HRRT	69.4	2.3 x 10^8	8.1 x 10^6	41 – 44 sec.
GED690	47.8	6.1 x 10^8	1.2 x 10^7	42 – 47 sec.

C. *Image reconstructions*

PhITF images were reconstructed using MR-based attenuation correction that does not account for different tissue types, but incorporates the bed and coil effects to the attenuation map. ToF information was not applied in the image reconstruction. Reconstructions of high and low counting statistics data were made with iterative RAMLA algorithm using 2 to 30 iterations. The blob parameters were α=6.3716, blob-radius=2.8 and blob-increment=2.0375, based on the default brain protocol. Reconstructed image volume contained cubic 8 mm^3 voxels in 128 x 128 x 90 grid.

HRRT images were reconstructed using transmission based attenuation correction that was segmented into air, bone and other tissue, where other tissue class was close to the μ-value of water. Reconstructions of high and low count statistics data were made with iterative OSEM algorithm with (PSF-OP-OSEM) and without resolution modeling (OP-OSEM) using 1 to 50 iterations (1 to 15 in low statistics), and with analytical 3DRP algorithm after bi-cubic transradial gap-filling [3]. Reconstructed image volume contained cubic 1.8 mm^3 voxels in 256 x 256 x 207 grid.

GED690 images were reconstructed using CT-based attenuation correction. Reconstructions of high and low count statistics data were made with iterative OSEM algorithm with (PSF-OSEM) and without resolution modeling (OSEM) using 2 to 30 iterations, and with analytical 2D-FBP after Fourier-rebinning (FORE-2D-FBP). ToF information was not applied in the image reconstruction. Reconstructed image volume contained 1.4 x 1.4 x 3.3 = 6.5 mm^3 voxels in 256 x 256 x 47 grid.

D. *Image registrations and regions-of-interest delineation*

All images were registered and resliced to a representative HRRT image using rigid image registration algorithm in SPM8 (Wellcome institute, London, UK). ROIs were delineated on anatomically bound regions in Carimas 2.7 (Turku PET Centre, Turku, Finland) using a CT-image as reference. Cortical regions in ascending order of volume were orbitofrontal, temporal, cingulate, parieto-occipital, medial frontal, cerebellar and lateral frontal cortex, whereas basal ganglio was the only subcortical region. Volumes of cortical ROIs were 8 cm^3 to 97 cm^3, and the volume of basal ganglio ROI was 17 cm^3. To account for differences in activity concentrations between scanners and between high and low count statistics each image was divided by the sum of total counts in the image prior to voxel-level or ROI analysis

E. *Measures*

Intra-scanner comparisons included asymmetry indices (AI) between hemispheres, between anterior and posterior, and between superior and inferior regions, as in AI=100%(ROI_x/ROI_y-1), and convergence assessments of iterative reconstructions. Between scanner comparisons were made for varying count statistics using analytically reconstructed high count statistics HRRT image as reference. The intrinsic partial-volume effect (PVE) related to the imaging device and reconstruction method was assessed from native, non-post-smoothed data. Furthermore, post-smoothing using a 3D Gaussian kernel corresponding to the intrinsic resolution of each scanner was applied. Thus, for the HRRT we applied a kernel size of 2.5 mm x 2.5 mm x 2.5 mm

FWHM, and for the GED690 and PhITF a kernel size of 5 mm x 5 mm x 5mm FWHM was applied.

III. Results

A. Image quality

Fig. 2 shows representative images from the three scanners. Transaxial slices shown in Fig. 2 were extracted from 3D images reconstructed from the high count statistics data (rows 1 and 4), from the low count statistics data (rows 2 and 5), and standard deviation maps were generated from the low count statistics replicates (rows 3 and 6). Standard deviation images are not directly comparable, since different NEC levels were generated in different scanners. However, within scanner comparison between reconstruction methods is possible. Color scale in rows 1, 2, 4 and 5 was between 0 and 0.04, and in the rows 3 and 6 between 0 and 0.02. Rows 1 to 3 represent the native image resolution and rows 4 to 6 images that were post-smoothed with 3D Gaussian filter corresponding to intrinsic resolution of each scanner.

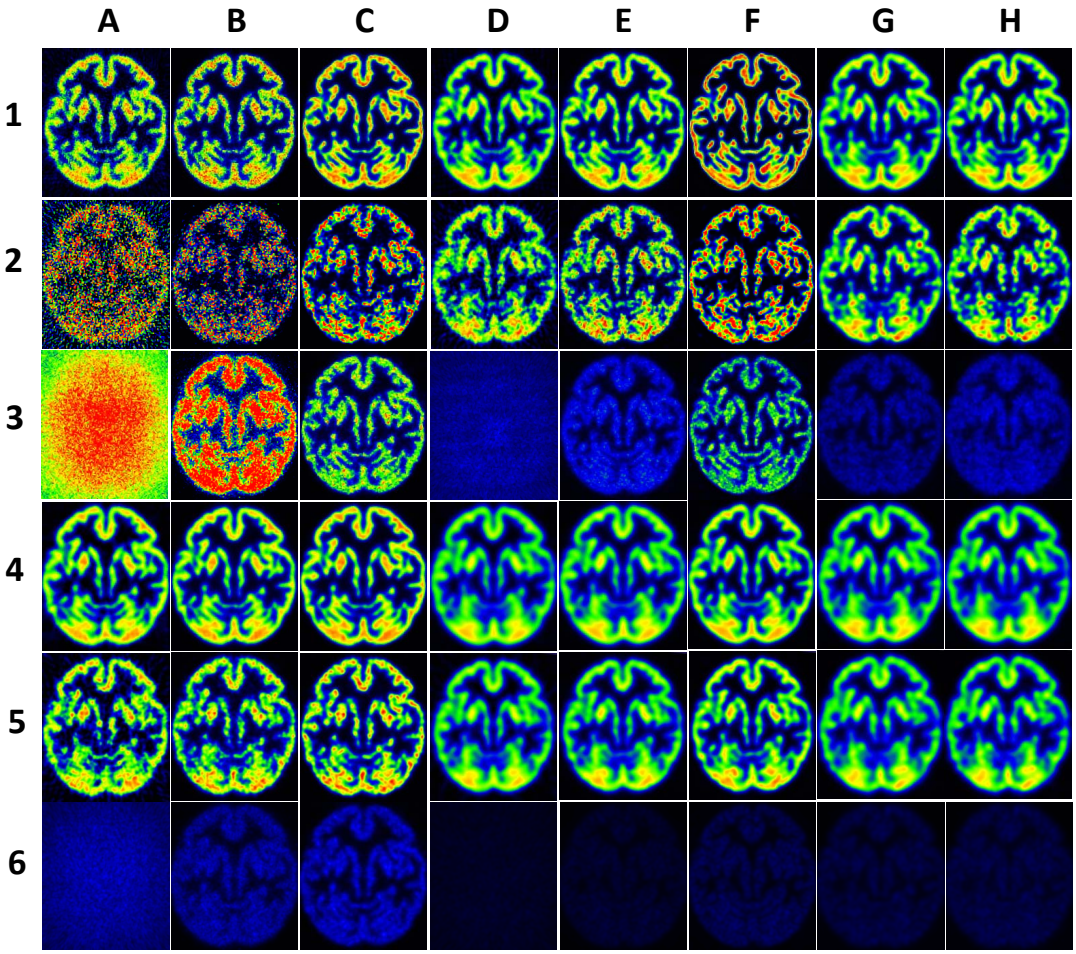


Fig. 2: Transaxial slices at the level of basal ganglio. Same color scale was used within each row. Rows 1 and 4: high count statistics images. Rows 2 and 5: low count statistics images. Rows 3 and 6: standard deviation images from low count statistics replicates. Rows 1 to 3 represent data from native non-smoothed images, and rows 4 to 6 images after post-smoothing with a scanner resolution matched 3D Gaussian filter. Colums represent different scanners and reconstruction settings as follow: **A**=HRRT, Bicubic gap-filling and 3DRP; **B**=HRRT, OP-OSEM, 10 it. 16 ss.; **C**=HRRT, PSF-OP-OSEM, 12 it., 16 ss.; **D**=GED690, FORE and FBP; **E**=GED690, OSEM, 7 it., 24 ss.; **F**=GED690, PSF-OSEM, 30 it., 24 ss.; **G**=PhITF, RAMLA, 5 it., default blob; **H**=PhITF, RAMLA, 30 it., default blob

A. Convergence of iterative reconstruction

Figures 3, 4 and 5 show the convergence curves for iterative reconstructions in PhITF, GED690 and HRRT, respectively. Convergence was studied both in high and low count statistics. Low count statistics data are shown as mean±s.d. difference over the replicates. For the GED690 and HRRT analytical reconstruction was used as reference, whereas, PhITF reconstructions were compared to high iteration iterative reconstruction, since analytical reconstruction was not readily available.

Fig. 3 shows the convergence curves for the PhITF RAMLA iterations. After five iterations there was negligible change in the image estimate, as measured by ROI mean. Furthermore, there was no significant difference in the convergence speed between high and low count statistics.

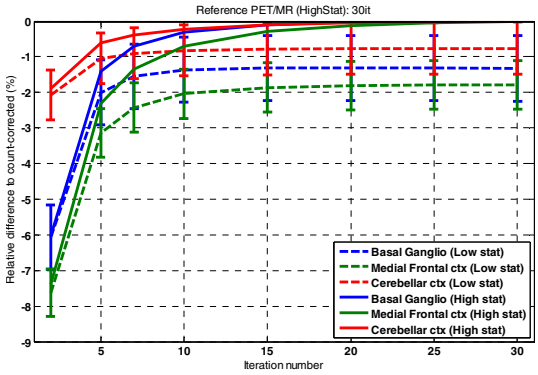


Fig. 3: Convergence of iterative RAMLA reconstruction in PhITF. Default blob parameters for brain studies were used. Solid lines represent high count statistics, and the dashed lines with the errorbars show the mean±s.d. difference in low count statistics reconstructions.

Fig. 4 shows the convergence curves for the GED690 reconstructions. GED690 VuePoint iterative reconstruction yielded overall higher mean ROI values than the corresponding analytical reconstruction, likely due to filtering applied in the FORE-FBP algorithm. There was no significant difference in the convergence between high and low count statistics, although, low count statistics yielded systematically lower ROI mean values than high count statistics. After seven iterations there was negligible change in the image estimates when regular OSEM algorithm was used. Inclusion of resolution modeling (PSF-OSEM) in the reconstruction yielded significantly higher ROI mean values as compared to analytical reconstruction in low iteration numbers (7% to 13%), and up to 15% to 23% higher values at high iteration numbers. There was no plateau in the convergence curve after 30 iterations.

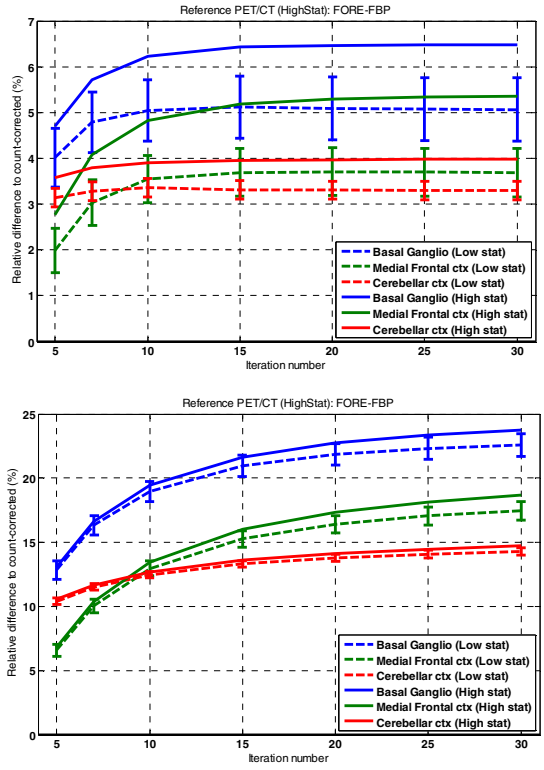


Fig. 4: Convergence of iterative OSEM (Top) and PSF-OSEM (Bottom) reconstruction in the GED690. VuePoint algorithm with 24 subsets was used. Solid lines represent high count statistics, and the dashed lines with the errorbars show the mean±s.d. difference in low count statistics reconstructions.

Fig. 5 shows the convergence curves for the HRRT OP-OSEM (top) and PSF-OP-OSEM (bottom) iterations. There was good agreement between OP-OSEM and analytical reconstruction after eight to ten iterations with regard to high count statistics data, whereas, low count statistics data converged much more quickly to a level that was lower than that from high statistics. Inclusion of resolution modeling (PSF-OP-OSEM) in the reconstruction yielded good agreement in the low iteration numbers (5% to 7% difference at iteration 12) whereas high iterations yielded 7% to 11% overshoots as compared to analytical reconstruction. There was no plateau in the convergence curves after 50 iterations, but the behaviour in low count statistics was similar with high count statistics data.

In all scanners the low count statistics data yielded systematically lower mean values than the corresponding high count statistics reconstructions. In the PhITF the difference was between 0.8% to 1.8% using RAMLA, in the GED690 the difference was between 0.8% to 1.5% using OSEM and between 0.5% to 1% using PSF-OSEM and in the HRRT the difference was between 1.5% to 6% using OP-OSEM and between 0.5% to 1% using PSF-OP-OSEM.

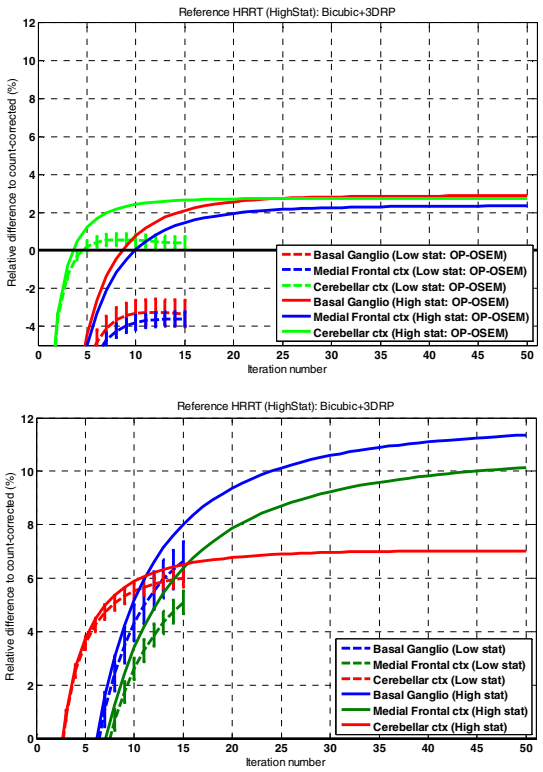


Fig. 5: Convergence of iterative OP-OSEM (Top) and PSF-OP-OSEM (Bottom) reconstruction in the HRRT scanner. 16 subsets was used. Solid lines show the difference between iterations when high statistics emission data was reconstructed, whereas, the dashed lines with the errorbars show the mean±s.d. difference between iterations in low count statistics reconstructions.

A. Asymmetry indices

Fig. 6 shows asymmetry indices between hemispheres, between front and back and between top and bottom for chosen reconstructions. Hemispheric asymmetries and asymmetries between front and back were in the range of -7% to 6%, whereas asymmetries between top and bottom were in the range of 3% to 23%. Asymmetries in high and low count statistics reconstruction were identical (data not shown).

B. Image bias relative to HRRT analytical reconstruction

Since the lack of ground truth values, image bias in the anatomical ROIs was evaluated with regard to HRRT analytical reconstruction. Fig. 7 shows the regional biases for chosen reconstructions. For native data, PhITF yielded 4% to 14% lower values, in both high and low count statistics, whereas HRRT OP-OSEM yielded slightly higher values in high statistics and lower in low statistics, but PSF-OP-OSEM values were 8% to 15% higher in either statistics. For native data, GED690 analytical reconstruction yielded 3% to 10% lower values, whereas range of difference for OSEM values was -4% to 3%. GED690 PSF-OSEM yielded highest values, with relative difference in the range of 9% to 25%.

After post-smoothing with a 3D gaussian filter all differences were in the range of -20% to 10%, with largest absolute difference (20%) in the temporal cortex with PhITF. Largest change was seen in the GED690 PSF-OSEM reconstruction, where the overshoots of 9% to 25% were changed to 2% to -5% difference to HRRT, which means good agreement. However, the overshoots seen in the HRRT PSF-OP-OSEM were not removed by post-smoothing.

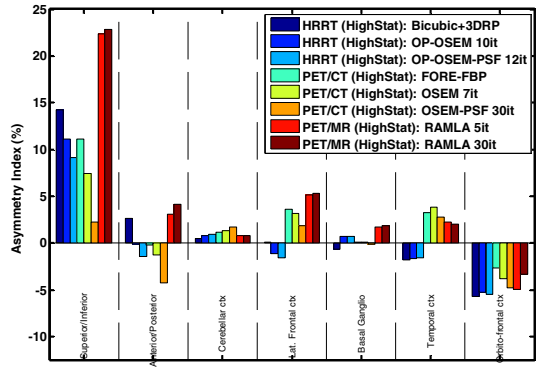


Fig. 6: Assessment of image homogeneity. Asymmetry index indicates the inhomogeneity between hemispheres, between superior/inferior and between anterior/posterior regions. All images were reconstructed from full count statistics emission data

IV. Discussion and Conclusions

In the current study we used a new brain phantom to assess the impact of scanner and reconstruction method choice to the image quantification. To our knowledge this was the first time this phantom was scanned in high-resolution PET. We generated a database of images from three PET scanners and using various reconstruction parameters. Images in the database were matched for direct comparisons within and between scanners.

We measured convergence of iterative reconstruction methods and showed that inclusion of PSF model can change the convergence behaviour. In the GED690 we saw still significant increase in the mean ROI values after 30 iterations, which can be considered as high iteration number. Similar behaviour was seen in the HRRT when PSF model was included, although, the slope of the convergence curve was probably lower. On the other hand, inclusion of the PSF model in the HRRT reconstruction was beneficial for the convergence in low count statistics, as has been shown before [4]. We have shown before in a phantom study that the OP-OSEM reconstruction suffers from negative bias in low count statistics in regions with high contrast to background [2]. In the current study we also found small negative biases, which can be appreciated from Fig. 5 (top) – as compared to high statistics, convergence

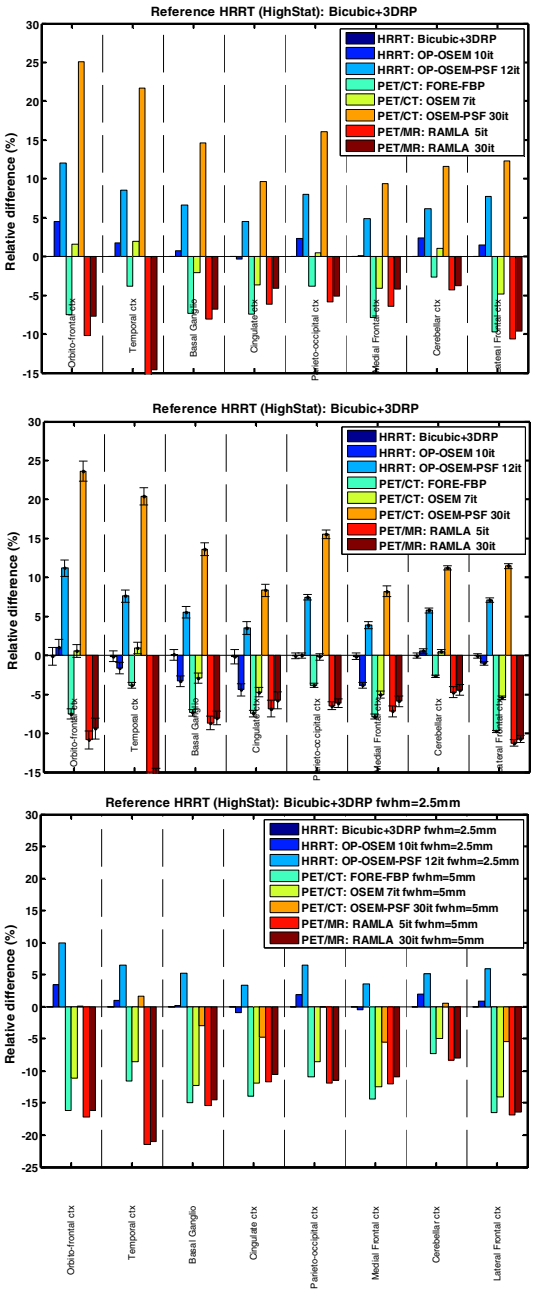


Fig. 7: Regional image bias relative to HRRT analytical reconstruction (Bicubic gap-filling + 3DRP). Top: biases in high counting statistics data and native spatial resolution. Middle: biases in low counting statistics data and native spatial resolution, bars represent the mean relative difference and errorbars the ±s.d. of the relative difference. Bottom: biases in high count statistics data after post-smoothing.

in low count statistics stops in a much earlier point, which results in underestimation of high values. It has been suggested (personal communication with Y. Jian and R. Carson, Yale University, US) that the convergence problem might be due to zero valued pixels at some image update in low count statistics and high number of subsets (c.f. [6]). Regularized-, or PSF-(OS)EM algorithms don't suffer from zero-voxel problem, since covariance of neighboring voxels quarantees positiveness. Further experiments using less subsets in OP-OSEM reconstruction in the HRRT will be made. Benefit of non-PSF reconstruction in the HRRT – if low count statistics bias can be avoided – would be better image homogeneity in high iteration numbers. Fig. 6 shows that PSF reconstruction converges in different values depending on the region, resulting in up to 5% difference between cerebellar cortex and basal ganglio, whereas non-PSF reconstruction was within 1%.

Overshoot behaviour of the PSF reconstruction was most evident in the GED690 where up to 9% difference was seen between cerebellar cortex and basal ganglio. As compared to the ROI values from the analytical HRRT reconstruction, region dependent overshoots of 10% to 25% were seen at 30 iterations of GE690 PSF reconstruction. These large overshoots are likely due to so-called edge artefacts that are typically seen in PSF reconstructed images (see Fig. 2 column F). A method of sieves that applies a post-smoothing filter of the same FWHM as was used in the PSF reconstruction has been proposed to remove the edge artefact [5]. Therefore, we post-smoothed all images with the corresponding 3D Gaussian filters and re-calculated the regional biases. The overshoots in the GED690 PSF reconstruction changed to -5% to 1% differences as compared to analytical HRRT reconstruction, whereas overshoots in the HRRT PSF reconstruction remained after post-smoothing (see Fig. 7 bottom).

Largest biases, as compared to the HRRT, were seen in the PhITF images regardless of count statistics or number of iterations in RAMLA reconstruction. The blob based RAMLA reconstruction parameters provided by the vendor are probably optimized for diagnostic use, where certain image smoothness is beneficial. However, in research use where quantitative values are the main interest PVE should be minimized. In the future study we will experiment with different blob parameter combinations to find optimal settings with regard to image contrast and noise. Furthermore, the image bias was region dependent so that the medial regions suffered from less bias than lateral and inferior regions. Largest bias was seen in the temporal cortex which in this phantom was surrounded by the skull compartment. The current reconstruction scheme in PhITF does not incorporate the skull attenuation in the attenuation correction, which might partly explain the large biases. The attenuation correction effect is elaborated in [7].

We conclude that the new anthropomorphic brain phantom is very usefull for image quality and quantitative assessments. Albeit the lack of ground truth values relative comparisons can give us important insight to the scanner performance and impact of image reconstruction parameters. Advantages of this new phantom are realistic head contour and attenuation effect from the skull. As compared to real human scans we can also quarantee immobilization of the head and stationary

distribution of the activity, thus providing us a high count statistic reference image. Inclusion of background activity within the white matter compartment would be a major improvement in the phantom design, but has proven complicated due to limitations of 3D printing. In the future work we seek to find improvements in the phantom design, and meanwhile do further experiments with the current phantom as was discussed above.

ACKNOWLEDGMENT

We would like to thank Prof. Hidehiro Iida from National Cerebral and Cardiovascular Center Research Institute, Japan, for providing the phantom for this study.

REFERENCES

[1] H. Iida, Y. Hori, K. Ishida, E. Imabayashi, H. Matsuda, M. Takahashi, H. Maruno, A. Yamamoto, K. Koshino, J. Enmi, S. Iguchi, T. Moriguchi, H. Kawashima, ja T. Zeniya, "Three-dimensional brain phantom containing bone and grey matter structures with a realistic head contour", *Ann Nucl Med*, vol. 27, nro 1, ss. 25–36, tammi 2013.

[2] J. Johansson, V. Oikonen, ja M. Teras, "Quantitative brain imaging using the new, fast iterative histogram-mode reconstruction for the HRRT PET scanner", teoksessa *IEEE Nuclear Science Symposium Conference Record, 2007. NSS '07*, 2007, vol. 5, ss. 3463–3467.

[3] U. Tuna, S. Peltonen, ja U. Ruotsalainen, "Interpolation for the gap-filling of the HRRT PET sinograms by using the slices in the direction of the radial samples", teoksessa *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 2009, ss. 3273–3279.

[4] M. Walker, "Bias in iterative reconstruction of low-statistics PET data: Benefits of a resolution model", *Physics in Medicine and Biology*, 2011.

[5] S. Stute ja C. Comtat, "Practical considerations for image-based PSF and blobs reconstruction in PET", *Phys Med Biol*, vol. 58, nro 11, ss. 3849–3870, kesä 2013.

[6] Y. Jian ja R. Carson, "Effect of Subsets on Bias and Variance in Low-Count Iterative Reconstruction", teoksessa *2013 IEEE Nuclear Science Symposium Conference Abstract(NSS/MIC)*.

[7] J. Teuho, J. Johansson, J. Lindén, V. Saunavaara, T. Tolvanen, ja M. Teräs, "Quantitative Bias in PET/MR from Attenuation Correction and Reconstruction: a Comparison with PET and PET/CT with an Anatomical Brain Phantom", teoksessa *2013 IEEE Nuclear Science Symposium Conference Abstract(NSS/MIC)*, vol. 2013.

Publication IV

Tuna, U., Johansson, J., Ruotsalainen, U., "Evaluation of analytical reconstruction with a new gap-filling method in comparison to iterative reconstruction in [^{11}C]-raclopride PET studies.", *Ann Nucl Med*, 2014, 28, 5, 417-429.

Publication V

Johansson, J., Alakurtti, K., Joutsa, J., Tohka, J., Ruotsalainen, U., Rinne, J.,O., "Comparison of manual and automatic techniques for sub-striatal segmentation in 11C-Raclopride PET studies", *Nuclear Medicine Communications*, 2016, Accepted for publication.

Original article

Nuclear Medicine Communications

Comparison of manual and automatic techniques for substriatal segmentation in ^{11}C-raclopride high-resolution PET studies

Jarkko Johansson[a,d], Kati Alakurtti[a,b], Juho Joutsa[c,d], Jussi Tohka[g,h], Ulla Ruotsalainen[e,f] and Juha O. Rinne[a,c]

Background **The striatum is the primary target in regional ^{11}C-raclopride-PET studies, and despite its small volume, it contains several functional and anatomical subregions. The outcome of the quantitative dopamine receptor study using ^{11}C-raclopride-PET depends heavily on the quality of the region-of-interest (ROI) definition of these subregions. The aim of this study was to evaluate subregional analysis techniques because new approaches have emerged, but have not yet been compared directly.**

Materials and methods **In this paper, we compared manual ROI delineation with several automatic methods. The automatic methods used either direct clustering of the PET image or individualization of chosen brain atlases on the basis of MRI or PET image normalization. State-of-the-art normalization methods and atlases were applied, including those provided in the FreeSurfer, Statistical Parametric Mapping8, and FSL software packages. Evaluation of the automatic methods was based on voxel-wise congruity with the manual delineations and the test–retest variability and reliability of the outcome measures using data from seven healthy male participants who were scanned twice with ^{11}C-raclopride-PET on the same day.**

Results **The results show that both manual and automatic methods can be used to define striatal subregions. Although most of the methods performed well with respect to the test–retest variability and reliability of binding potential, the smallest average test–retest variability and SEM were obtained using a connectivity-based atlas and PET normalization (test–retest variability = 4.5%, SEM = 0.17).**

Conclusion **The current state-of-the-art automatic ROI methods can be considered good alternatives for subjective and laborious manual segmentation in ^{11}C-raclopride-PET studies.** ***Nucl Med Commun*** **00:000–000**

Nuclear Medicine Communications 2016, **00**:000–000

Keywords: **^{11}C-raclopride, PET, region-of-interest analysis, striatum**

[a]Turku PET Centre, [b]Department of Diagnostic Radiology, [c]Division of Clinical Neurosciences, Turku University Central Hospital, [d]Turku PET Centre, University of Turku, Turku, [e]Department of Signal Processing, Tampere University of Technology, [f]BioMediTech, Tampere, Finland, [g]Department of Bioengineering and Aerospace Engineering, Universidad Carlos III de Madrid, Leganes and [h]Instituto de Investigación Sanitaria Gregorio Marañon, Madrid, Spain

Correspondence to Jarkko Johansson, MSc, Turku PET Centre, University of Turku, PO Box 52, FIN-20521 Turku, Finland
Tel: +358 231 330 564; fax: +358 223 18191; e-mail: jarkko.johansson@tyks.fi

Received 23 January 2016 **Revised** 31 March 2016 **Accepted** 17 May 2016

Introduction

PET with ^{11}C-raclopride provides a widely validated method for assessing baseline levels of dopamine (DA) type 2 receptor (DA2R) availability [1] as well as DA2R occupancy induced by either pharmacological [2] or nonpharmacological [3,4] stimuli. In humans, the DA2R are most abundant in the striatum, where they act as modulators of various functions, such as locomotion, and reward system, as well as many high-order cognitive functions such as working memory [5]. The reward system is considered to be linked to addictive behaviors such as substance abuse [6] and pathological gambling [7], and thereby linked to dopaminergic signaling. Thus, ^{11}C-raclopride-PET can be used in the research of numerous facets of human behavior and cognition in health and in disease (see Egerton *et al.* [4] for review). As a downside, methodological caveats such as erroneous region-of-interest (ROI) delineation can significantly hamper the reliability and sensitivity of ^{11}C-raclopride-PET because of the small size of striatal substructures relative to the spatial resolution of PET. In the current work, we tackled the ROI-delineation issue through investigation of numerous ROI-delineation techniques in the analysis of high-resolution ^{11}C-raclopride-PET.

Inaccuracies in ROI delineation can hamper the sensitivity of the PET assessment, and in the worst case, generate biased inferences (cf. [4]). In a PET assessment of specific receptor binding, decreased sensitivity can be because of oversized ROIs compared with the receptor population or because of intrapatient or interpatient mismatches in region definitions. Oversized ROIs can

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (*www.nuclearmedicinecomm.com*).

DOI: 10.1097/MNM.0000000000000559

yield exaggerated partial-volume effect (PVE), resulting in underestimated PET radioactivity, whereas direct PET-based segmentation can in the worst case result in biased outcome (cf. [4]). Furthermore, inaccuracies in the ROI border placement between adjacent regions can result in attenuated stimulus response in a single assessment or decreased statistical power in a group-level assessment because of additional methodological variability between patients.

The concept of substriatal ROI-delineation that is based on functional rather than structural territories was first introduced by Mawlawi *et al.* [8] and later reviewed and validated by Martinez *et al.* [9] using D-amphetamine-induced striatal activation. Martinez *et al.* [9] defined functional territories on the basis of experimental animal studies that included three striatal subregions: the limbic striatum (LSTR), involved in drive and motivation; the associative striatum (ASTR), involved in cognition; and the sensorimotor striatum (SMST), involved in locomotion. Furthermore, they showed significant differences in the response to amphetamine-stimulus between the subregions using ^{11}C-raclopride-PET. The concept of functional subdivision is nowadays established and the ROI delineations are mostly generated by hand using the guidelines of Mawlawi and colleagues, Martinez and colleagues. More recently, Tziortzi *et al.* [10] have refined the guidelines of Mawlawi and colleagues in the light of current understanding of the striatal functional organization in humans. Nevertheless, manual ROI delineation inevitably contains a subjective component that hampers the reproducibility of ^{11}C-raclopride-PET analysis. Thus, automated methods are called upon that provide precise and repeatable striatal ROI delineation.

The automated or operator-independent methods for striatal ROI-delineation can be based on (a) individualization of a ROI set defined in a standard space through image deformation (see Del Campo *et al.* [11] for review); (b) by direct automatic clustering of the PET image ([12]); or (c) individual measurement of corticostriatal connectivity as suggested by Tziortzi *et al.* [13]. In the current work, we examined the approaches that did not require diffusion tensor imaging measurements. The first approach based on the individualization of a template ROI set has been implemented successfully using conventional PET scanning (see Del Campo *et al.* [11] for review). In the study by Del Campo *et al.* [11], a probabilistic atlas was generated on the basis of manual ROI delineations and subsequently individualized using nonrigid transformations. Their conclusion was that an automated method based on a probabilistic atlas can provide an accurate and efficient alternative to manual ROI drawing. The second approach using direct PET image clustering has been implemented previously for high-resolution PET data (see Farinha *et al.* [14] for validation). In the study by Farinha *et al.* [14], the direct PET image clustering was used to generate five striatal subregions in both striata instead of three for the first time.

The primary aim of the current study was to explore the characteristics of various ROI-delineation methods used in the analysis of high-resolution ^{11}C-raclopride-PET as new methodologies have emerged, but have not yet been evaluated. ^{11}C-raclopride-PET has widespread usage and recent investigations suggest the benefit of connectivity-based (CB) ROI methods over the structure-based manual ROI delineation [13]. In particular, the potential advantage of high-resolution PET is not fully exploited if the ROI methods are inaccurate. We approached the ROI-method optimization task through evaluation of the current state-of-the-art ROI delineation methods with respect to their test–retest characteristics. A test–retest data set of seven healthy male participants scanned twice (during the same day) with the high-resolution research tomography (HRRT) was collected previously to investigate the short-term repeatability and reliability of the ^{11}C-raclopride assessment in high resolution and at resting state [15]. The ultimate aim of the test–retest setup is to substitute testing against an unknown ground truth in methodological testing, and it has therefore been recommended for methodological comparison studies (see Weir [16] for review). The repeated scans under the presumably same condition enable estimation of method-wise bias as well as reliability through analysis of variance. It was shown in our previous report [15] that the binding potentials (BPs) within the manually delineated structure-based (SB) ROIs (on the basis of rules by Martinez and colleagues) were not biased, crossing the trials for the current data set; thus, it is reasonable to expect nonbiased estimates using any of the tested methods. Furthermore, the test–retest setup can be used to estimate the dissection between the inherent, true variation between patients from that because of methodological imprecision.

Methods

Overview

The primary objective of the current work was to compare the striatal ROI delineation approaches in the analysis of high-resolution ^{11}C-raclopride-PET data to provide practicable guidelines for realistic imaging conditions. Analysis of the ^{11}C-raclopride-PET data most commonly uses the SRTM [17] with the cerebellar cortex as the reference region to obtain the BP_{ND} (see Innis *et al.* [18] for nomenclature) estimates; thus, the SRTM-BP_{ND} was considered the primary parameter of interest in the test–retest evaluations.

Participants

The current study reuses human ^{11}C-raclopride-PET measurements acquired originally for a different purpose as reported in [15]. The original study protocol was approved by the Ethics Committee of the Hospital

District of Southwestern Finland. The study participants were provided written information on all of the relevant issues involved in the study. Written consent, not limited for single usage of the data, was obtained from each participant. The study was carried out according to the Declaration of Helsinki.

All of the participants were right handed and non-smokers. The age, height, and weight of the participants were 24.5 ± 3.5 years, 185.5 ± 12.5 cm, and 74 ± 14 kg, respectively (mean ± SD). To exclude any structural brain abnormalities and obtain anatomical references, all participants underwent 1.5 T MRI. Each participant underwent two ^{11}C-raclopride-PET scans in a resting condition on the same day between 10 : 00 a.m. and 6 : 00 p.m., the tracer injections being at least 2.5 h apart.

The BPs of these participants have been reported previously [15]. However, in the current study, the data set was exploited for the ROI method comparison, rather than for research of the BP *per se*. None of the results in the current study are identical to the previously reported results because of small differences in the data preprocessing and manual ROI delineation. However, there was a strong correlation between the results from these two studies (data not shown) and the current results should not be considered as individual observations in a meta-analysis.

MRI

The MRI was performed using a 1.5 T MRI system (Gyroscan Intera CV Nova Dual; Philips Medical Systems, Best, the Netherlands) with a SENSE head coil. Transversal T1-FFE three dimensional (3D) images were acquired with isotropic 1 mm voxel size, repetition time 25 ms, echo time 4.6 ms, and flip angle 30°. Parallel imaging was used with factor 2.

PET imaging

The ^{11}C-raclopride preparation and PET imaging have been described previously [15]. PET measurements were performed using the HRRT scanner (Siemens Medical Solutions, Knoxville, Tennessee, USA), a brain-dedicated high-resolution PET research tomograph. The HRRT system is a dual-crystal-layer scanner capable of depth-of-interaction measurement of the coincidence photons. The intrinsic resolution of the HRRT system is ~ 2.5 mm, whereas the measured spatial resolution [full-width-at-half-maximum (FWHM), of a point-source] varies between 2.5 and 3.5 mm within a field-of-view covering most of the brain [19]. In the current work, a transmission scan was acquired before an ^{11}C-raclopride injection for attenuation-map calculation. PET scanning in listmode was initiated at the time of tracer injection and continued until 55 min from injection time. A dynamic series was generated from the listmode data using the following frame sequence: 2×0.5, 9×1, 3×2, 3×3, and 6×5 min. Image reconstructions were performed using the ordinary-Poisson ordered-subsets expectation maximization (OP-OSEM) algorithm [20], with resolution modeling (RM-OP-OSEM) [21] based on external measurement of the scanner point-spread function. The number of iterations in RM-OP-OSEM reconstruction was 10, whereas the number of subsets was 16, and the image voxel size was $\sim 1.22 \times 1.22 \times 1.22$ mm^3. Tissue attenuation maps were reconstructed using the maximum-a-posteriori for transmission data repetition time algorithm with the standard human brain priors for bone, soft tissue, noise, water, or air [22]. Scattered events were estimated using the single-scatter simulation algorithm [23], whereas randoms were estimated from the block singles using a variance reduction algorithm [24]. In the RM-OP-OSEM algorithm, the emission data were not precorrected, but the correction terms were included in the update equation to avoid bias because of zero truncation of the precorrected data [20].

Image preprocessing

Image preprocessing steps are shown in a diagram in Fig. 1. Image preprocessing was performed in Statistical Parametric Mapping (SPM) software (version 8; Wellcome Trust Centre for Neuroimaging, London, UK; *http://www.fil.ion.ucl.ac.uk/spm/*). First, the SPM8 realign function was used for correcting the dynamic PET imaging data for misalignments between scans as well as between frames using the normalized mutual-information algorithm. The first frame with a decent signal (frame 5) was chosen as a reference in both sessions. Initially, each frame in both PET sessions was registered within session (to frame 5 position), followed by registration of the second PET session data to the first PET session orientation (frame 5 of session 2 to the frame 5 of session 1). After the first pass, a mean image of all frames (from both sessions) was created and each frame was re-registered to the mean image. The two-pass procedure was used to minimize misalignment between PET sessions that might hamper the test–retest characteristics in ROI-based analysis. Registrations were visually confirmed. The subsequent analyses were carried out on the basis of the assumption that the data from the two PET sessions were spatially aligned, that is, the ROI methods that do not directly use PET data were not repeated for each PET, but the same ROIs were used for both sessions. The two-pass registration process should help to protect against any systematic biases in the registration between the two PET sessions.

Furthermore, an MRI-based procedure was used to standardize the head posture in PET data. First, the MRI data were registered with the individual mean PET image (cf. previous paragraph) using the SPM8 coreg-function and normalized mutual-information optimization. Registrations were confirmed visually. Second, the MRI data were registered with an MNI152 template

Fig. 1

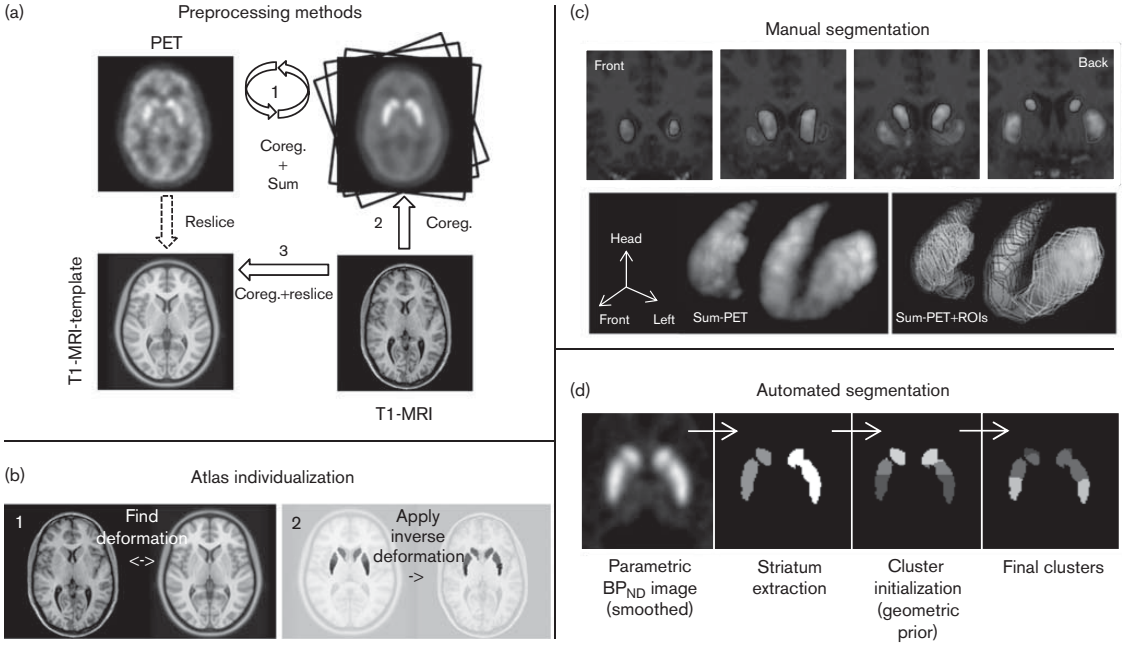


Illustration of the data preprocessing and the region-of-interest (ROI) methods. Part (a) shows how the dynamic PET data and the MRI data were preprocessed to standardize the data for subsequent ROI extraction. First, the PET frames were coregistered to a single reference frame and sum image of all frames was created, after which all frames were coregistered to the sum image and the sum image was recalculated. Second, the individual T1-MRI-image was coregistered to the sum image, followed by coregistration to the MNI template. Finally, the dynamic-PET and MRI data were resliced into MNI-template matrix and voxel size according to the above transformations. Part (b) shows how the ROIs were generated using the atlas individualization approach. Atlas-based methods used various mappings from standard space (atlas-space) to individual space, and various atlases (Table 1). Part (c) shows how the ROIs were generated using manual segmentation. Manual segmentation was performed on the MRI + sum-PET fused image according to anatomical guidelines in coronal slices. The lower panel shows the manually delineated contours in 3D. Part (d) shows how the ROIs were extracted using automatic segmentation. Automatic PET image clustering proceeded in four steps: (1) striatum extraction, (2) edge-preserving smoothing, (3) spatial initialization, and (4) k-means clustering. Coreg, coregistered; 3D, three dimensional.

image using rigid registration (coreg-function). The coreg-function enabled simultaneous registration of pre-registered PET data through transformation matrix manipulation. Both the MRI and the PET data were resliced into $1.5 \times 1.5 \times 1.5$ mm voxel size of the MNI152 template image. The rigid registration does not provide exact spatial matching between individuals, but works as a means to standardize the orientation of the anterior commissure – posterior commissure line. Standardization of the anterior commissure-posterior commissure line orientation may facilitate the manual ROI delineation as the structures appear in the same orientation as in the typical brain atlases, whereas the head posture during PET imaging can be somewhat tilted.

Nonrigid image deformation procedures

Spatial normalization parameters (deformation fields; nonrigid mappings) were obtained using the unified segmentation algorithm in SPM8 [25] and using a combination of the FLIRT and FNIRT functions in FSL5 (FMRIB, Oxford, UK; *http://fsl.fmrib.ox.ac.uk/*) [26]. First, the spatial normalization mappings were applied to the individual mean PET images. After visual inspection, the FSL-based normalized PET data were found to be slightly more consistent compared with the SPM-based data; thus, FSL normalization was chosen for PET-template image generation. A raclopride-specific template image was created from the seven normalized ^{11}C-raclopride mean images as an average. PET-based image normalization was performed using the SPM8 normalise-function [27] and the novel raclopride-specific template image as a target. Both the source and the target image data were presmoothed using a 3D Gaussian filter with 4 mm (FWHM) kernel size in each direction.

In the current work, the aim of nonrigid spatial matching was to create an individualizing spatial mapping from the atlas-defining space to the individual coordinates. Atlas-based ROI delineation can be achieved through either mapping the PET data into the atlas-defining space or mapping the atlas into individual space. In the current

work, the latter approach was chosen to minimize the need for interpolation of the PET data, and to more readily obtain quality assurance of the spatial deformations through visual inspection of the ROIs in the individual space. The inverse deformation (i.e. deformation from the standard space to the individual space) was obtained directly in the unified segmentation algorithm in SPM, whereas the FSL-based deformation fields were explicitly inversed using the invwarp function of FSL. The PET-based SPM normalizations were inversed using the deformations-function and the Inverse-component in SPM8. The three individualizing mappings are hereafter denoted as MRIF, MRIS, and PETS for MRI-based mappings using FSL and SPM and PET-based mapping using SPM, respectively.

ROI methods

The manual and automatic ROI methods for substriatal segmentation examined in this study are summarized in Table 1 and shown in a diagram in Fig. 1. The methods examined in the current study had differences in the naming and definitions of the striatal subregions. To allow for a comparison between methods, common regions consisting of the whole STR, the LSTR, the ASTR, and the SMST were formed as primary targets. Table 2 summarizes the initial regions and combinations for each method.

Manual segmentation

Manual segmentation of the striatum was performed by two operators, one with more experience (MANSEG1) and other with less experience, but instructed by the first (MANSEG2). Individual T1-weighted MRI fused to PET sum image (both sessions together) was used as a reference. Thus, the ROI delineation was performed only once by both operators and the same ROIs (within operator) were used to extract data from both sessions using coregistered PET images (cf. previous paragraphs). The anatomical landmarks described by Mawlawi *et al.* [8], Martinez *et al.* [9] were applied in the ROI delineation and all ROIs were drawn on coronal slices. The PET sum image was used as additional guidance in the striatum border search to minimize PVE. MANSEG included five anatomical substructures per hemisphere, which were combined into the three functional substructures: the ventral striatum as such formed the LSTR, the precommissural portions of the caudate and putamen formed the ASTR, and the postcommissural portions of the caudate and putamen formed the SMST (Table 2).

Atlas-based segmentation

The atlas individualization process is shown in a diagram in Fig. 1. Three atlases were considered in the atlas-based striatal segmentation: the probabilistic structural atlases by Fischl *et al.* [28] and by Tziortzi *et al.* [10] and a connectivity-based atlas by Tziortzi *et al.* [13]. The latter two are included in the FSL package (striatum-structural-2 mm and striatum-con-label-thr50-7sub-2 mm), whereas the first is a FreeSurfer (Laboratory for Computational Neuroimaging, Charlestown, Massachusetts, USA; *http://surfer.nmr.mgh.harvard.edu/*) built-in atlas. The structural atlas in FreeSurfer is based on automated labeling of several brains, whereas the structural FSL atlas was obtained by directly segmenting the MNI152 template according to the guidelines in [10]. The SB atlas segmentations yielded caudate, putamen and nucleus accumbens or ventral striatum ROIs, which were transformed into STR (all regions) and LSTR (nucleus accumbens or ventral striatum) for comparison (Table 2). The structural atlases were individualized only within the host software packages and are denoted as FSSEG for FreeSurfer and as STRUCTMRIF for FSL.

The CB atlas was generated using diffusion tensor imaging and a tractography technique to find the projection territories in the striatum that innervate certain cortical regions [13]. Tziortzi and colleagues used probabilistic tractography in 12 healthy male volunteers to investigate the projections between the cortex and the striatum

Table 1 Summary of the manual and automatic ROI methods for the substriatal segmentation examined in this study

Method	Source	Summary	Atlas/guidelines
Manual segmentation (MANSEG1/2)	PET + MRI	Manual ROI delineation on fused PET + MRI	No atlas. Guidelines in (Mawlawi *et al.*) [8] were applied
Automatic striatum clustering (PETSEG)	PET	Two-step direct segmentation: (1) Striatum extraction using MRF (2) Clustering using k-means	No atlas. Initialization of k-means clustering based on location.
Structure-based atlas individualization with FreeSurfer (FSSEG)	MRI	Atlas-based segmentation provided in FreeSurfer	Probabilistic structural atlas (Fischl *et al.*) [28]
Structure-based atlas individualization with FSL (STRUCTMRIF)	MRI	Structural atlas individualization based on FSL MRI normalization	Segmentation of MNI152 template (Tziortzi *et al.*) [10]
Connectivity-based atlas individualization/FSL (CONNMRIF)	MRI	Connectivity atlas individualization based on FSL MRI normalization	DTI tractography analysis between cortical regions and striatum provided seven substructures (Tziortzi *et al.*) [13]
Connectivity-based atlas individualization with SPM8 and MRI (CONNMRIS)	MRI	Connectivity atlas individualization based on SPM8 MRI normalization	
Connectivity atlas individualization with SPM8 and PET (CONNPET)	PET	Connectivity atlas individualization based on SPM8 PET normalization	

ROI, region-of-interest; SPM8, Statistical Parametric Mapping.

Table 2 Initial striatal subregions provided by each method and combinations

Method	MANSEG1, MANSEG2, PETSEG	FSSEG, STRUCTMRIF	CONNMRIF, CONNMRIS, CONNPET
Region			
Limbic striatum (LSTR)	VST	NACC/VST	LSTR
Associative striatum (ASTR)	CAUA + PUTA	–	EXE
Sensorimotor striatum (SMST)	CAUP + PUTP	–	CAM + ROM + PAR + TEM + OCC
Caudate (CAU)	CAUA + CAUP	CAU	–
Putamen (PUT)	PUTA + PUTP	PUT	–
Anterior caudate (CAUA)	CAUA	–	–
Posterior caudate (CAUP)	CAUP	–	–
Anterior putamen (PUTA)	PUTA	–	–
Posterior putamen (PUTP)	PUTP	–	–
Ventral striatum (VST)	VST	VST (FSL)	–
Nucleus accumbens (NACC)	–	NACC (FS)	–
Executive cortex area (EXE)	–	–	EXE
Caudal motor cortex area (CAM)	–	–	CAM
Rostral motor cortex area (ROM)	–	–	ROM
Parietal cortex area (PAR)	–	–	PAR
Temporal cortex area (TEM)	–	–	TEM
Occipital cortex area (OCC)	–	–	OCC

CONNMRIF/MRIS/PET, connectivity-based atlas individualized using FSL-MRI, SPM-MRI or SPM-PET image normalization; FSSEG, FreeSurfer segmentation; MANSEG1/2, manual segmentation PETSEG, automatic PET image clustering; STRUCTMRIF, structural atlas segmentation in FSL.

yielding CB subdivision of the striatum in each participant. As a result of the connectivity analysis, Tziortzi and colleagues generated atlases at different probability thresholds. For the current study, we chose a seven-substructure atlas with a 2 mm voxel size that was obtained at a 50% probability threshold (striatum-con-label-thr50-7sub-2 mm). The seven substructures were named on the basis of the cortical regions that they innervate: limbic, executive, rostral and caudal motor, parietal, temporal, and occipital. For the CB atlas ROIs, the motor, parietal, temporal, and occipital components were combined to form the SMST, whereas the executive component was considered the ASTR and limbic was considered the LSRT (Table 2). The connectivity atlas was individualized using the individualizing mappings as described above and are denoted as CONNMRIF, CONNMRIS, and CONNPET for MRI-FSL, MRI-SPM, and PET-SPM normalizations, respectively.

Direct PET image segmentation

The principle of the direct PET image segmentation process is shown in a diagram in Fig. 1. An automatic PET image clustering algorithm uses the known heterogeneity of the DA receptor distribution [29] to find clusters of distinct ^{11}C-raclopride binding in the striatum. Similar to the methods of Juslin and Tohka [30] and Farinha *et al.* [14], the algorithm consists of two main steps: (a) initial striatum extraction and (b) weighted kernel k-means clustering [31]. In the current study, the initial step was implemented using a Markov random field (MRF) model-based extraction similar to that of other authors [30,32] (see Supplementary Document, Supplemental digital content 1, *http://links.lww.com/NMC/A60*, which describes the method MRF-based method), instead of a possibly more error-prone deformable surface model approach that we have used in [14]. The benefit of using MRF-based striatum extraction lies in the inclusion of the PVE kernel in the image model inspired by MRI segmentation approaches [32]. Before the clustering step, the BP_{ND} image was smoothed using an edge-preserving 3D Gaussian filter with FWHM = 2.5 mm and the result of the striatum extraction step as a striatal mask (Fig. 1). In the second step, the weighted kernel k-means algorithm was applied to partition the connection graph into five segments using a geometrical division of the striatum into precommissural and postcommissural caudate and putamen and ventral striatum as initialization, instead of a random initialization approach [14]. The automatic PET image segmentation results are denoted as PETSEG throughout the report. PETSEG yielded the same substructures as the MANSEG; thus, the same combination strategy was used to obtain LSTR, ASTR, and SMST (Table 2).

Pharmacokinetic modeling methods

Pharmacokinetic modeling was performed for ROI time–activity course (TAC) data using SRTM [17]. SRTM yields a BP of the tracer that is relative to the nondisplaceable binding as measured in the reference tissue and is therefore denoted as BP_{ND} [18]. The ROI-TAC modeling was performed using in-house software fit_srtm, version 3.0.7 for nonlinear SRTM (Turku PET Centre, Turku, Finland). An atlas-based automated method for cerebellar ROI generation was chosen for the reference region TAC extraction. The reference region method was chosen on the basis of cerebellar distribution volume test–retest characteristics (see Supplementary Table, Supplemental digital content 2, *http://links.lww.com/NMC/A61* for cerebellar test–retest characteristics).

Statistical methods

Spatial agreement between MANSEG1 and each of the segmentation methods was determined at the ROI voxel

set level, that is, the MANSEG1 ROIs were considered as a reference, although not necessarily ground truth. The ROI voxel set similarity was measured using the Jaccard coefficient expressed as the ratio between the size of the intersection and the size of the union of the voxel sets:

$$J(A,B) = (\mid A \cap B \mid)/(\mid A \cup B \mid),$$

where the Jaccard coefficient takes values between 0 (no agreement) and 1 (perfect match) and is a widely used performance measure for evaluating image-segmentation algorithms.

Furthermore, agreement of the BPs (BP_{ND} as obtained through ROI-TAC SRTM analysis) between MANSEG1 and each of the segmentation methods was determined using the limits of agreement (LOA), expressed as the mean of pairwise differences and 95% confidence interval (CI). Moreover, the Pearson's product–moment correlation coefficient was calculated between MANSEG1-BP_{ND} and each of the methods to assess the similarity of the BP_{ND} rank order.

The test–retest setting enabled method-independent estimation of the repeatability and reliability of each method. That is, instead of choosing one of the possibly erroneous methods as a reference, each method can be assessed independently. Therefore, the test–retest setting has been suggested for method comparison studies (see Weir [16] for review). In the current study, the test–retest characteristics were considered as the primary figure of merit in the evaluation of BP_{ND} estimates (as obtained through various ROI delineations) and the test–retest characteristics were also calculated for the ROI volumes when applicable. Using MRI-based atlas-individualization methods, there was only one ROI set for each individual; thus, no replication characteristics for ROI volume could be formulated. While using the direct PET-image segmentation, PET-based atlas individualization, and MANSEG, there were two independent ROI sets for each individual and formulation of the ROI volume test–retest characteristics was possible.

The repeatability of BP_{ND} and ROI volume (when applicable) were calculated using the test–retest variability (TRV) [16], which relates the difference in the two measurement outcomes to that of their mean. The per cent TRV can be expressed as:

$$\mathrm{TRV}(\%) = 100\% \times (2 \mid X_1 - X_2 \mid)/(X_1 + X_2),$$

where X_1 is the measurement outcome in PET-session 1 and X_2 is the measurement outcome in PET-session 2. The TRV (%) is reported as the mean ± SD over the seven participants for each method.

A common measure of test–retest reliability is the intraclass correlation coefficient (ICC), which is based on the one-way random-effects analysis of variance model, denoted as ICC(3,2) in Shrout and Fleiss [33]:

$$\mathrm{ICC} = \frac{(\mathrm{MS_B} - \mathrm{MS_w})}{[\mathrm{MS_B} + (k-1)\mathrm{MS_w}]},$$

where MS_B indicates the between participants and MS_w is the within-participant mean of the sum of squared differences, and $k = 2$. It is apparent that ICC is a relative measure of reliability and depends heavily on the current samples' variability. Generally, large between-participant variability (large MS_B) results in ICC closer to one (good reliability), although the within-participant repeatability might be poor. In method comparison studies, this feature of the ICC contains a particular caveat because of the possibly method-dependent outcome of MS_B. That is, one of the methods can exaggerate the between-participant variability and thereby wrongly benefit from that as high ICC. In the current study, the manual approach embodies a considerable subjective component that can result in exaggerated between-participant variability in the ROI delineation; thus, a more robust measure of reliability is required.

It has been noted that an absolute measure of reliability would be more appropriate in method comparison studies [16]. Moreover, an absolute measure would allow for the calculation of the CI of the minimal detectable change (single participant). In the current study, the SEM was applied as an absolute measure of reliability, as expressed in [16]:

$$\mathrm{SEM} = \mathrm{SD}\sqrt{(1-\mathrm{ICC})},$$

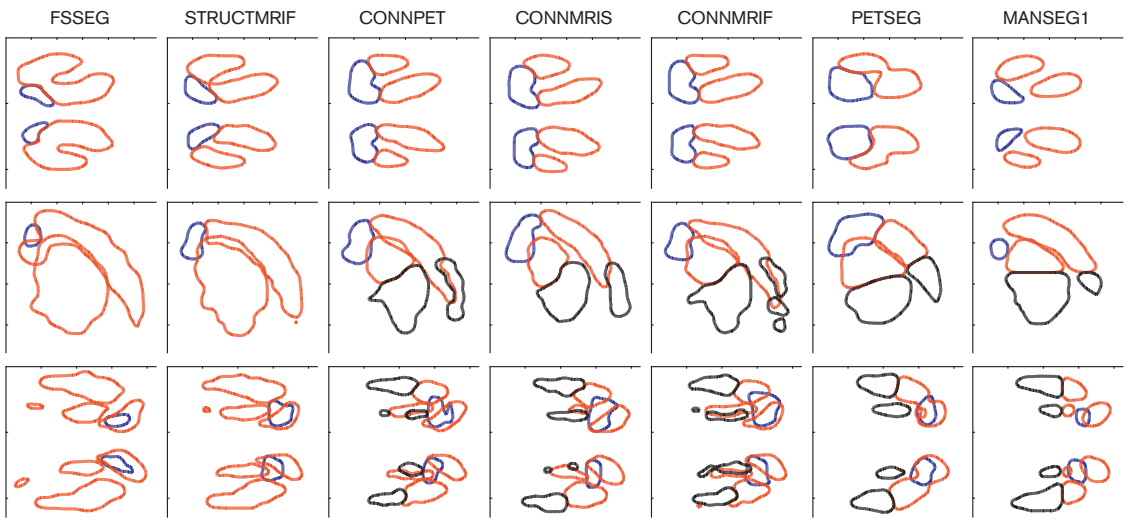
where SD indicates the standard deviation of binding parameters from all participants. The minimal detectable change (MD) was calculated on the basis of SEM for a CI of 95% [16].

$$\mathrm{MD} = \mathrm{SEM} \times 1.96 \times \sqrt{2},$$

where MD shows how much two measurements of the same individual must differ from each other to be considered as a true change; thus, a decrease in MD could be interpreted as improved sensitivity. MD (%) was calculated relative to the mean BP_{ND}.

The repeatability of the ROI volumes was calculated when applicable; for instance, PETSEG and CONNPET were made independently on PET1 and PET2, and MANSEG was repeated by two operators. Furthermore, coefficients of variation were calculated from regional distribution volumes. Descriptive statistics include the arithmetic mean ± SD and the coefficient of variation [CoV (%) = 100%S/X] (where X represents the mean value of X) for binding parameters.

Fig. 2



Approximate contours of the limbic striatum (blue), associative striatum (red), and sensorimotor striatum (black). Columns from left to right represent coronal, sagittal, and transaxial projections, respectively. Rows from top to bottom represent manual segmentation (MANSEG1), automatic PET image clustering (PETSEG), connectivity atlas individualization using FSL-MRI (CONNMRIF), SPM8-MRI (CONNMRIS), or SPM8-PET (CONNPET) normalization, and structural atlas individualization using FSL-MRI (STRUCTMRIF) normalzation or Freesurfer (FSSEG) segmentation results, respectively. The sagittal and transaxial projections represent combinations from multiple cutting planes to show all structures at once.

Results

Agreement of ROI volumes

Figure 2 shows the ROI delineations generated by the various methods (of the combined regions). There was reasonable agreement between the methods with respect to the location of the striatum. However, manual delineation (particularly MANSEG1) yielded systematically smaller volume for the whole striatum ROI compared with the automated methods (see Supplementary Figure, Supplemental digital content 3, *http://links.lww.com/NMC/A62*, which shows the method-wise volumes). The average volume of the whole striatum was 13 cm^3 for the MANSEG1, whereas average volumes generated with automated methods were in the range of 16–21 cm^3. Direct PET segmentation was associated with smaller average striatal volumes compared with other automated methods. A strong negative linear correlation ($R = -0.93$; see Supplementary Figure, Supplemental digital content 4, *http://links.lww.com/NMC/A63*, which shows the linear regression lines) existed between the average striatal volume (by method) and the average BP_{ND} (by method).

Variability in the striatal volume between participants was assessed using CoV (%) as presented in Table 3. MANSEG and direct PET segmentation yielded larger between participant variability compared with atlas-based automated methods, with approximately two-fold CoV (%) for MANSEG1 compared with that of atlas-based methods. Furthermore, the within-participant variability in ROI volumes was calculated when applicable [see Table 3 TRV (%) volume]. The interoperator variability in manual ROI volume was in the range of 16–44% [average TRV (%)], whereas the direct PET image segmentation showed TRV (%) of 7–17% and the CB atlas individualization using PET-normalization (CONNPET) as small as 2–3% [average TRV (%)].

Illustration of the ROI contours in Fig. 2 showed systematic intermethod differences in the definition of the LSTR ROIs that resulted in poor spatial agreement as measured using the Jaccard coefficient and Pearson's correlation coefficient (Table 3) relative to the MANSEG1, whereas the agreement was somewhat better in the associative and SMST. In particular, the direct PET image segmentation and CB atlas methods yielded large deviation from the SB LSTR ROIs generated either manually or using atlas individualization.

Agreement of the BP_{ND} estimates

Intermethod variability in ROI volumes was manifested in BP_{ND} estimates. As was noted earlier, the average BP_{ND} estimate was strongly associated with the striatal volume in the whole striatum and also in the associate striatum, but not in the limbic or SMST (see Supplementary Figure, Supplemental digital content 4, *http://links.lww.com/NMC/A63*, which shows the linear regression lines). However, Pearson's correlation coefficient showed strong intermethod correlation between the

Table 3 Regional (combined regions) BP_{ND}, test–retest repeatability, and reliability values of BP_{ND} and ROI volumes and agreement with MANSEG1

		Volume		Agreement (volume)		BP_{ND}		TRV(%) (BP_{ND})		Reliability (BP_{ND})			Agreement (BP_{ND})	
	Method	CoV (%)	TRV (%)	Jaccard	*R*	Mean ± SD	CoV (%)	Mean ± SD	Range	ICC	SEM	MD (%)	LOA (range)	*R*
Limbic striatum	MANSEG1	18				3.76 ± 0.55	14.6	4.55 ± 3.19	−1.8 to −9.1	0.94	0.14	10.2		
	MANSEG2	16.3	21.1 ± 12.8	0.46	0.26	3.65 ± 0.51	13.9	4.66 ± 5.32	−0.6 to −15.7	0.9	0.16	12.4	(−0.28 to 0.06)	0.95
	PETSEG	17.7	17.4 ± 17.6	0.16	−0.03	4.49 ± 0.47	10.6	7.28 ± 3.50	−11.6 to −11.2	0.73	0.24	15.1	(0.46–1.00)	0.87
	STRUCTMRIF	8.3		0.33	−0.48	3.62 ± 0.32	8.8	4.36 ± 2.73	−4.9 to −7.6	0.86	0.12	9.2	(−0.41 to 0.13)	0.95
	CONNMRIF	8.2		0.21	−0.54	4.05 ± 0.41	10.2	4.30 ± 2.50	−6 to −9.3	0.9	0.13	8.9	(0.09–0.49)	0.96
	CONNPET	10.6	2.6 ± 1.7	0.2	−0.34	3.96 ± 0.35	8.9	3.42 ± 1.81	−5.4 to −5.6	0.92	0.1	7	(−0.04 to 0.44)	0.95
	FSSEG	9.4		0.18	0.08	3.24 ± 0.30	9.2	5.33 ± 5.44	−10.2 to −15.6	0.74	0.15	13.1	(−0.87 to −0.17)	0.81
Associative striatum	MANSEG1	14.3				5.07 ± 0.50	9.8	5.42 ± 3.51	−9 to −9	0.81	0.22	11.9		
	MANSEG2	14.2	16.3 ± 11.1	0.61	0.7	4.65 ± 0.38	8.1	4.92 ± 3.83	−8.5 to −9.1	0.74	0.19	11.5	(−0.58 to −0.24)	0.96
	PETSEG	13.2	7.2 ± 3.5	0.51	0.67	4.98 ± 0.50	10	7.74 ± 5.05	−14.1 to −12.2	0.63	0.31	17	(−0.22 to 0.04)	0.97
	CONNMRIF	7.2		0.47	0.74	4.53 ± 0.25	5.6	5.26 ± 3.69	−8.8 to −10	0.41	0.2	12	(−0.84 to −0.24)	0.88
	CONNPET	9.8	2.2 ± 1.4	0.47	0.73	4.56 ± 0.35	7.7	5.08 ± 3.64	−9.5 to −9.4	0.71	0.19	11.6	(−0.70 to −0.32)	0.96
Sensorimotor striatum	MANSEG1	21.9				5.47 ± 0.59	10.8	4.34 ± 2.66	−5.8 to −8.1	0.91	0.18	9		
	MANSEG2	9.7	43.7 ± 16.6	0.56	0.75	4.82 ± 0.39	8.1	4.84 ± 2.97	−5.8 to −10	0.78	0.18	10.5	(−0.94 to −0.36)	0.91
	PETSEG	18.5	16.1 ± 13.3	0.53	0.68	5.20 ± 0.51	9.8	7.01 ± 4.50	−13.6 to −11.6	0.7	0.28	15	(−0.46 to −0.08)	0.95
	CONNMRIF	8		0.33	0.19	4.23 ± 0.25	6	5.16 ± 2.47	−8.3 to −7.1	0.58	0.17	10.9	(−1.86 to −0.62)	0.1
	CONNPET	8.3	2.2 ± 1.0	0.36	0.18	4.38 ± 0.34	7.8	4.90 ± 2.52	−8.2 to −6.5	0.78	0.16	10.1	(−1.36 to −0.82)	0.98

BP_{ND} (mean ± SD from both scans), SRTM-based binding potential using the PETS reference region TAC; CONNMRIF/MRIS/PET, connectivity atlas segmentation based on FSL-MRI, SPM-MRI, or SPM-PET image normalization; CoV (%), coefficient of variation; FSSEG, Freesurfer segmentation; ICC, intraclass correlation coefficient; LOA, limits of agreement (relative to MANSEG1); MANSEG1/2, manual segmentation; MD (%), minimal detectable change; PETSEG, automatic PET image clustering; R, Pearson's correlation coefficient (relative to MANSEG1); ROI, region-of-interest; STRUCTMRIF, structural atlas segmentation in FSL; TRV (%), test–retest variability.

BP_{ND} estimates (except for CONNMRIF in SMST, see Table 3), whereas, the LOA range (Table 3) showed both negative and positive systematic biases relative to MANSEG1. Even between MANSEG, significant differences were found in the associative and sensorimotor striata.

Repeatability and reliability of the BP_{ND} estimates

Intramethod within-participant variability of the BP_{ND} was smallest using CONNPET [see TRV (%) (BP_{ND}) in Table 3], followed by MANSEG1 and MANSEG2. However, differences in between-participant variability (repeatability) were mostly small, whereas that of PETSEG was somewhat larger. The repeatability of PETSEG and, subsequently, its reliability was ruined by a single outlier scan with a markedly lower dose (3.3 MBq/kg compared with the group mean ± SD of 5.6 ± 1.6 MBq/kg) in the second scan. Although the outlying participant had exceptionally poor-quality PET data, the sensitivity of PETSEG with respect to noise must be acknowledged.

Intramethod between-participant variability of the BP_{ND} was smallest using CONNMRIF [see CoV (%) (BP_{ND}) in Table 3], followed by CONNPET. There was up to a 61% increase in CoV (%) between the smallest (CONNMRIF) and the largest variability (MANSEG1). However, the large between-participant variability resulted in high ICC, as expected. The highest ICC was obtained using MANSEG1, followed by MANSEG2 and CONNPET. The SEM, however, indicated a different order in performance – the smallest SEM (best reliability) was obtained using CONNPET, followed by CONNMRIF and then MANSEG, although the differences were not very large. Only PETSEG showed markedly higher SEM compared with the other methods, because of one outlier scan with significantly poorer image quality. The average MD was close to 10% (smallest using CONNPET; see Table 3), except for PETSEG (15%).

Test–retest characteristics within the initial subregions are presented in the Supplementary Table, Supplemental digital content 5, *http://links.lww.com/NMC/A64*. For the manual delineation and direct PET image segmentation, the five structural striatal subregions showed mostly similar test–retest characteristics as the combined functional subregions, except for the posterior caudate, which showed somewhat poorer performance for MANSEG1 and PETSEG compared with the combined regions. For the CB atlas ROIs, the five subregions that constituted the SMST ROI showed varying test–retest performance that depended on the ROI size. The largest subregions (caudal and rostral motor, and parietal) showed comparable test–retest performance with the combined regions, whereas the smallest subregions (occipital and temporal) showed poorer performance.

Discussion

This study compared several ROI-delineation methods for high-resolution ^{11}C-raclopride-PET. Manual ROI-delineation based on rules by Mawlawi *et al.* [8] and Martinez *et al.* [9] can be considered a gold standard because of its widespread usage. However, the manual ROI delineation approaches are associated with considerable operator-dependent variability and high cost, and alternative methods are constantly being developed. The popularity of automated methods for functional striatal subdivision has been limited because of the complexity of the segmentation task. Nevertheless, few previous studies have shown the feasibility of automated segmentation within striatum. For the current study, an atlas-based approach and direct PET image clustering algorithms were implemented on the basis of promising results in previous studies [11,14]. In the current study, the automated methods were not only compared against the manual ROI delineation as is commonly done, but the methodological performance was primarily assessed using the test–retest protocol. The test–retest protocol has been suggested for method comparison studies (see Weir [16] for review) for its method-independent performance evaluation. In addition, a secondary analysis was carried out through a comparison with the manual ROI method to enable descriptive analysis of the method performance.

Manual segmentation

For the data presented in this study, the interoperator variability in manual ROI delineation was considerable, yielding statistically significant differences between BP_{ND} estimates (Table 3). Although the linear regression analysis showed good agreement of the BP_{ND} estimates, the LOA range indicated statistically significant interoperator differences in the associative and SMST, but not in the LSTR (Table 3). Interoperator differences may be attributed to subjective striatum-border search, that is, the fused PET and MRI data can be interpreted differently according to operators' experience and habits. The rules of Mawlawi and colleagues, Martinez and colleagues describe the structural cues only on the basis of T1W MRI-data, but it has been considered that use of PET-image fusion as a guideline would help to protect against unwanted PVE. Apparently, for the data presented in this study, the more experienced operator (MANSEG1) was more conservative with respect to the striatum border placement than the novice operator (MANSEG2). This was indicated by the systematically smaller ROI volumes of the experienced operator compared with the beginner (see Supplementary Figure, Supplemental digital content 3, *http://links.lww.com/NMC/A62*, which shows the striatal volumes). Smaller ROIs of the experienced operator yielded systematically higher BP_{ND} estimates, thus implying lesser PVE contamination in MANSEG1 ROIs. In principle, lesser PVE contamination is a desirable property, but often PVE-compensation methods are associated with increased methodological variation. For the data presented in this study, the between-participant variability for both the ROI volume and BP_{ND} estimates was highest using MANSEG1 ROIs, followed by MANSEG2. In the absence of ground truth data, it cannot be unequivocally inferred whether the increased variability is because of true variability between participants that is more accurately measured using smaller ROIs or whether it is more because of erroneous variability in ROI delineations. As was discussed before, the commonly used test–retest reliability measure of ICC strongly favors large variability between patients, be it erroneous or true variation. Consequently, the ICC measures for MANSEG were superior to those using automated methods and superior using MANSEG1 ROIs compared with using MANSEG2 ROIs. However, the performance rank order was rather different according to the SEM, which is an absolute measure of reliability taking into account the standard deviation of the BP_{ND} estimates, that is, the between-participant variability. The SEM was marginally superior using MANSEG2 ROIs and more so using the atlas-based automatic ROIs in comparison with MANSEG1. SEM has been suggested for measuring the reliability instead of ICC as it is less vulnerable to increased methodological variation (see Weir [16] for review). Thus, according to the SEM, the more conservative border placement of the more experienced operator may have been associated with increased methodological variation, albeit the impact of PVE was likely smaller.

Atlas-based methods

The atlas-based methods showed good overall performance; there were no complete failures and general agreement with manual delineation was reasonable. Albeit the ROIs did not match exactly with the MANSEG (see Jaccard and Pearson's correlation coefficient in Table 3), the BP_{ND} estimates showed a strong correlation in the linear regression analysis (except for CONNMRIF in SMST). Thus, for the data presented in this study, the atlas-based automatic methods could replicate the between-participant rank order of BP_{ND} estimates of the manual ROI delineation, whereas the LOA range showed statistically significant differences between the BP_{ND} estimates. In the LSTR, the SB atlas ROIs showed small underestimation (FSSEG) or no bias (STRUCTMRIF), whereas the CB atlas ROIs showed overestimation (CONNMRIF) or nonsignificant bias (CONNPET) compared with MANSEG1, but in the associative and sensorimotor striata, the CB atlas ROIs showed systematic underestimations compared with MANSEG1. As discussed earlier, differences in the associative and sensorimotor striata may be mostly attributed to the ROI size rather than gross differences in the ROI placement, whereas there was a clear difference in the LSTR ROI placement between SB and CB ROIs.

That is, the placement of LSTR ROI was similar using manual ROI delineation and SB atlas ROIs, but the CB atlas ROIs were placed more toward the caudal ventral putamen and less toward the rostral dorsal caudate (see Fig. 2 for ROI contours). Thus, a visible difference in the LSTR ROI placement was shown using the CB atlas compared with structure-based approaches. The data presented here cannot be used to make inferences on whether the SB or the CB approach would yield LSTR ROIs that better delineate the functional organization. However, the data presented in this study show the feasibility of the novel CB atlas in automated ROI analysis, whereas the analysis carried out by Tziortzi *et al.* [13] suggests improved regional selectivity of CB ROIs over the SB ROIs for measuring the DA response.

Impact of the normalization method in the atlas-based approach

The data presented in this study suggest flexibility with respect to the choice of the normalization procedure. In the current study, we used individual T1W MRI data as well as PET data to find a mapping from the atlas defining space to the individual space where the ROIs were analyzed. The individualizing mappings on the basis of MRI data were generated using state-of-the-art algorithms implemented in SPM8 and FSL5 for nonrigid image matching. Differences between the ROIs from SPM8 and FSL5 were small (see Supplementary Table, Supplemental digital content 5, *http://links.lww.com/NMC/A64* for substriatal BP_{ND} in initial regions). Furthermore, the normalizing (inverse of individualizing) mappings from FSL5 MRI matching were exploited to generate an ^{11}C-raclopride-specific PET template. The PET template was generated using the PET sum images from the current study, and the PET-based individualizing mappings were estimated using an algorithm implemented in SPM8. Albeit the differences were not large, the PET-based individualization of the CB atlas performed better than the MRI-based one in the test–retest evaluation. Both the within-participant variability [TRV (%) see Table 3] and SEM were slightly superior using PET-based mappings. It can be discussed whether the PET-based normalization would have benefited from the fact that the PET template was created using the same data as those used to find the individualizing mappings. The concern is, however, alleviated by the fact that fourteen PET sum images were averaged in the template formation; thus, comparable performance could be expected for independent datasets. It has also been noted by others that the striatum is a challenging target for MRI-based normalization and PET-based normalization has been suggested instead [34].

Direct PET image segmentation

Direct PET image segmentation algorithm performed well in general, but on a single occasion, the precision of the initial striatum extraction was not adequate. Closer analysis of the data showed a visible difference in the image quality of the outlier scan compared with the others, attributed to a significantly lower radioactive dose as compared with the average dose (3.3 MBq/kg compared with the group mean ± SD of 5.6 ± 1.6 MBq/kg). Various parameter combinations were tested to improve the segmentation outcome, but without significant improvement. Applying larger kernel size (FWHM > 5 mm) in the presegmentation smoothing step (cf. Fig. 1) would have enabled extraction of the striatum more robustly, but at the cost of inaccurate segmentation particularly in the thin structure of the rostral caudate nucleus. Therefore, it was deemed necessary not to increase the smoothing kernel size to maintain the benefit from high-resolution PET imaging. Otherwise, the initial striatum extraction and subsequent clustering performed well, but the sensitivity of the current method with respect to image noise must be acknowledged. Conservative smoothing combined with the MRF-based RM in the striatum extraction yielded a small average striatal volume compared with other automated methods. The average striatal volume from PETSEG ROIs was second smallest after MANSEG1, suggesting excellent performance with respect to protection against PVE. Unfortunately, for the data presented here, the outlier scan ruined the overall test–retest performance of PETSEG, but the method showed great potential that might be exploited in future studies. For instance, combining (averaging) the BP_{ND} images of the same patient (if several scans made) would enhance the image SNR considerably and the outcome of direct PET image segmentation would be more reliable. In the current study, however, the aim was to investigate the replicability of the segmentations. A question that remains is whether the direct image clustering is sensitive to changes in BP_{ND} distribution. It has been shown that differences in the response to amphetamine-induced stimulation exist between striatal subregions [9,13], but the direct PET image segmentation has not been applied to such data. The study by Egerton *et al.* [4] showed the potential risk of using direct PET image segmentation in a comparison between baseline and task-induced stimulation when the ROIs were generated independently. The potential hazard in changing uptake distribution assessments is related to changing ROI definitions over sessions and thereby underestimation or overestimation of the activation effect. Thus, it might be advisable to combine the BP_{ND} data before image clustering not only for improved SNR but also for bias-free ROI analysis. Another question that remains is whether the direct PET image clustering yields subregions that are relevant with respect to the functional organization of the striatum. It is often assumed that brain anatomy follows function, or in this case, the receptor populations follow function, but the correspondence is seldom one to one. The data presented in this study showed a visible difference in the PETSEG LSTR ROIs compared with both SB and CB

delineations (see Fig. 2 for substriatal contours). Different initializations were tested (including manual ROIs), but the clustering outcome was nearly identical, implying that there is a strong gradient in the BP_{ND} estimates within ventral portions of the rostral caudate and the putamen that defined the LSTR cluster. The PETSEG LSTR ROIs extended more dorsally within the caudatus compared with other methods and less caudally within the putamen compared with CB ROIs, whereas the associative and SMST ROIs matched well with the CB ROIs (except for the rostral caudatus). It will remain for the future studies to show whether and how well the gradients in BP_{ND} distribution correspond with response to DA stimulation. In turn, it might be worthwhile to pursue the possibilities of direct PET image clustering in studies investigating changes in the striatal volume *per se*. It has been shown using ^{11}C-raclopride that there is an age-related decline in D2-receptor density on the one hand [35], and a decrease in the striatal volume as measured using MRI on the other [36]. Consequently, the interplay between the striatal volume loss and decline of the D2-receptor binding has been acknowledged as a potential confound because of increased PVE (Morris and colleagues). In the study by Morris *et al.* [37], PVE compensation was used for the PET data and a clearly slower decline in D2-receptor density compared with earlier studies was shown, implying a confound because of inadequate PET analysis in some earlier studies [37]. Thus, the direct PET image segmentation might be a viable option for improving the PET analysis in aging research.

Choice between ROI methods in the striatum

As was discussed earlier, the outcome from manual ROI delineations was possibly less contaminated by PVE compared with the atlas-based methods and to some degree compared with direct PET image segmentation, but likely at the cost of higher methodological variation. In addition, there was significant interoperator variability that would effectively prevent multioperator ROI-delineation in high-resolution ^{11}C-raclopride studies. Albeit the intraoperator variability was not directly measured in the current study, the elevated between-participant variability suggests variation in manual ROI delineations within operator. Be it true or erroneous, a slightly elevated variation of BP_{ND} was also observed between repetitions using manual ROIs. Test–retest within-participant variability is commonly considered to reflect the relative (BP_{ND}) change that can be detected using a given method. However, small within-participant variability can be obtained without high specificity, curbing the sensitivity of the assessment. Thus, it is important to consider the interplay of the within-participant and between-participant variability in the method performance evaluation. For the data presented here, the macroparameter MD (%) (Table 3) was calculated on the basis of SEM to designate the per cent change in individual BP_{ND} estimates that can be considered as true change. Thus, smaller MD (%) could be interpreted as improved sensitivity. For the data presented here, the average MD (%) was close to 10% in combined regions for all methods, except PETSEG (15%), with the smallest average using CB atlas-based and PET-based normalization (Table 3), thus implying that at the individual level, the sensitivity of other methods except PETSEG was very similar, albeit the impact of PVE was likely different. Furthermore, in group-level analysis, the BP_{ND} between-participant variability (SD of BP_{ND} estimates over participants) is often used in the statistical power calculations in place of true cohort variability. That is, the BP_{ND} SD and mean BP_{ND} (Table 3) are commonly used in the sample-size calculations. We used G*Power (version 3.1.9.2; University of Kiel, Kiel, Germany) to calculate the effect size on the basis of the mean and SD of BP_{ND} in Table 3, and subsequently, the total sample size for the detection of a 10% decrease in group mean BP_{ND} with a one-tailed paired *t*-test (with $\alpha = 0.05$, and power = 0.95). For the data presented here, the sample size was 8–10 participants using CONNPET, whereas for the MANSEG, the minimum sample sizes were 13–25 (MANSEG1) and 9–23 (MANSEG2) participants. Thus, the CB-atlas method showed superior sensitivity over manual ROI delineation at both the individual-level and group-level analysis. However, the data presented here do not provide direct validation of the methods; the relatively improved consistency shown here for the CB-atlas method suggests a potential benefit over manual ROI delineation.

As was discussed earlier, the PET-based normalization might have offered some improvement in the atlas individualization compared with the more commonly used MRI-based normalization. The striatum is a demanding target for MRI-based normalization because of a mixture of gray and white matter cells (thus the name striatum) and PET-based normalization has been acknowledged as a viable alternative [34]. In the current study, however, the PET data were acquired using a high-resolution PET scanner and the PET template was generated using the same data set, rendering the generalization of the current results rather difficult. That is, poor spatial resolution of the PET data from conventional scanners might hamper the PET-based normalization to a degree that would favor the usage of high-resolution MRI-based normalization instead. However, the source and target data in PET-normalization were presmoothed using a 4 mm (FWHM) kernel size that can be considered a means of resolution matching (cf. [38]). Under certain circumstances, the direct PET image segmentation algorithm may also be an appealing alternative, although the validity of the (BP_{ND}) gradient-based clustering may require further validation.

Reference region ROI generation

In addition to the striatal segmentation, the cerebellar ROI delineation was evaluated in the current study because of its broad usage in the analysis of ^{11}C-raclopride-PET and reference tissue-based modeling (see Supplementary Table, Supplemental digital content 2, *http://links.lww.com/NMC/A61* for cerebellar test–retest characteristics). Currently, the manual ROI delineation can be considered a gold standard for its widespread usage, but as for the striatal ROI drawing, there can be considerable differences between operators' opinions. However, differences in the ROI delineations per se may have lesser significance within the cerebellum, where the ^{11}C-raclopride uptake is nonspecific, but the noise characteristics of the reference-tissue-TAC become likely more important because of the model fitting process. Manual ROI delineation uses T1W MRI data in the search of the cerebellar cortex border typically from transaxial images, frequently with the help of fused PET data to avoid blood-rich territories such as veins. The process of manual cerebellar cortex delineation is laborious and it is therefore often limited to very few transaxial slices, resulting in small ROIs and noisy TACs. Thus, a robust automated method for cerebellar ROI-delineation would be highly beneficial for improved reference tissue-based modeling. Algorithms implemented in the FreeSurfer software have become a popular choice in automated ROI generation, but in a recent study by Schain *et al.* [39], the cerebellar FreeSurfer ROI was deemed suboptimal for reference tissue-based modeling. In their study, the comparison was, however, made only with MANSEG and not using a method-independent measure such as the test–retest characteristics in the current study. Furthermore, their suggestion was that the performance of the auto-ROI generation would be ligand dependent [39]. For the data presented here, the average cerebellar FreeSurfer ROI covered over 10-fold larger volume compared with MANSEG1 and approximately six-fold larger volume compared with atlas-based ROIs (see Supplementary Table, Supplemental digital content 2, *http://links.lww.com/NMC/A61* for cerebellar ROI characteristics). The atlas-based ROIs were deliberately limited to a specific cerebellar subregion, explaining the vast difference with FreeSurfer ROIs. Albeit the ROI volumes varied extensively, the outcome of full pharmacokinetic modeling with arterial input (distribution volume) showed a strong correlation between the methods (see Supplementary Table, Supplemental digital content 2, *http://links.lww.com/NMC/A61* for cerebellar ROI characteristics), implying a small role of the ROI delineation in the quantification of nonspecific uptake TAC. However, the repeatability of the outcome was superior using the larger ROIs of automated methods compared with those from MANSEG [see Supplementary Table, Supplemental digital content 2, *http://links.lww.com/NMC/A61* for TRV (%)], suggesting a benefit from improved SNR. In the current study, we did not notice any bias from additional spill-out signal originating from blood-rich areas using the automated methods, likely because of different characteristics of ^{11}C-raclopride uptake compared with ligands tested by Schain *et al.* [39].

Conclusion

The data presented in this study support implementation of fully automated ROI generation for ^{11}C-raclopride-PET analysis. It was shown that ROIs generated using CB atlas individualization can yield BP_{ND} estimates in substriatal domains that highly correlate with those obtained using manual ROIs. Furthermore, the ROIs generated automatically from CB atlas showed superior test–retest characteristics for both within-participant and between-participant variability, and consequently enhanced sensitivity for BP_{ND} alterations both at the individual and at the group level compared with manual ROIs. In addition, automated reference-region ROI generation was superior to that by manual ROI delineation according to improved test–retest characteristics.

The CB substriatal atlas and cerebellar atlas individualization were successful using either MRI-based or PET-based normalization. The PET-based normalization was slightly superior to that by MRI normalization in the striatum possibly partly because of the high resolution of the PET data presented here. However, the data presented here support experimentation with fully PET-based automated ROI generation using other scanners as well. Although the direct PET image clustering algorithm showed great potential, it should be applied with caution – robustness of the current implementation was questioned here by a failure to segment a single poor-quality image. If the image quality can be guaranteed through for instance BP_{ND} averaging, the direct PET image segmentation might provide an interesting alternative.

It will remain for future studies to show whether the CB definition of the striatal subdomains or those based on direct PET image segmentation can provide similar or possibly improved ROI delineations with resepct to true functional organization of the striatum compared with those obtained using conventional manual ROI drawing. Furthermore, a similar comparison is in place using different patient groups that might show larger variability in the striatal anatomy. The data presented here are representative for young healthy male participants and the performance may be slightly different for patients with anomalies or for instance enlarged ventricles.

Acknowledgements

This study was supported by the Hospital District of Southwest Finland (13133, Johansson). J.T.'s work is funded by the Universidad Carlos III de Madrid, the European Union's Seventh Framework Programme for research, technological development and demonstration

under grant agreement no. 600371, el Ministerio de Economía y Competitividad (COFUND2013-40258) and Banco Santander.

Conflicts of interest

Dr Joutsa has received lecturer honoraria from Boehringer-Ingelheim, a research grant from Lundbeck, and a travel grant from Abbvie. For the remaining authors there are no conflicts of interest.

References

1 Farde L, Hall H, Ehrin E, Sedvall G. Quantitative analysis of D2 dopamine receptor binding in the living human brain by PET. *Science* 1986; **231**:258–261.

2 Farde L, Nordstrom AL. PET analysis indicates atypical central dopamine receptor occupancy in clozapine-treated patients. *Br J Psychiatry Suppl* 1992; **17**:30–33.

3 Koepp MJ, Gunn RN, Lawrence AD, Cunningham VJ, Dagher A, Jones T, *et al.* Evidence for striatal dopamine release during a video game. *Nature* 1998; **393**:266–268.

4 Egerton A, Mehta MA, Montgomery AJ, Lappin JM, Howes OD, Reeves SJ, *et al.* The dopaminergic basis of human behaviors: a review of molecular imaging studies. *Neurosci Biobehav Rev* 2009; **33**:1109–1132.

5 Birren JE, Schaie KW. *Handbook of the psychology of aging*, 5th ed. San Diego, Calif: Academic Press; 2001.

6 Koob GF, Bloom FE. Cellular and molecular mechanisms of drug dependence. *Science* 1988; **242**:715–723.

7 Joutsa J, Johansson J, Niemelä S, Ollikainen A, Hirvonen MM, Piepponen P, *et al.* Mesolimbic dopamine release is linked to symptom severity in pathological gambling. *Neuroimage* 2012; **60**:1992–1999.

8 Mawlawi O, Martinez D, Slifstein M, Broft A, Chatterjee R, Hwang DR, *et al.* Imaging human mesolimbic dopamine transmission with positron emission tomography: I. Accuracy and precision of D(2) receptor parameter measurements in ventral striatum. *J Cereb Blood Flow Metab* 2001; **21**:1034–1057.

9 Martinez D, Slifstein M, Broft A, Mawlawi O, Hwang DR, Huang Y, *et al.* Imaging human mesolimbic dopamine transmission with positron emission tomography. Part II: amphetamine-induced dopamine release in the functional subdivisions of the striatum. *J Cereb Blood Flow Metab* 2003; **23**:285–300.

10 Tziortzi AC, Searle GE, Tzimopoulou S, Salinas C, Beaver JD, Jenkinson M, *et al.* Imaging dopamine receptors in humans with [^{11}C]-(+)-PHNO: dissection of D3 signal and anatomy. *Neuroimage* 2011; **54**:264–277.

11 Del Campo N, Tait RJ, Acosta-Cabronero J, Hong YT, Izquierdo-Garcia D, Smith R, *et al.* Quantification of receptor-ligand binding potential in substriatal domains using probabilistic and template regions of interest. *Neuroimage* 2011; **55**:101–112.

12 Tohka J, Wallius E, Hirvonen J, Hietala J, Ruotsalainen U. Automatic extraction of caudate and putamen in ^{11}C-raclopride PET using deformable surface models and normalized cuts. *IEEE Trans Nucl Sci* 2006; **53**:220–227.

13 Tziortzi AC, Haber SN, Searle GE, Tsoumpas C, Long CJ, Shotbolt P, *et al.* Connectivity-based functional analysis of dopamine release in the striatum using diffusion-weighted MRI and positron emission tomography. *Cereb Cortex* 2014; **24**:1165–1177.

14 Farinha RJ, Ruotsalainen U, Hirvonen J, Tuominen L, Hietala J, Fonseca JM, Tohka J. Segmentation of striatal brain structures from high resolution PET images. *Int J Biomed Imaging* 2009; **2009**:156234.

15 Alakurtti K, Aalto S, Johansson JJ, Nagren K, Tuokkola T, Oikonen V, *et al.* Reproducibility of striatal and thalamic dopamine D2 receptor binding using [^{11}C]raclopride with high-resolution positron emission tomography. *J Cereb Blood Flow Metab* 2011; **31**:155–165.

16 Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005; **19**:231–240.

17 Lammertsma AA, Hume SP. Simplified reference tissue model for PET receptor studies. *Neuroimage* 1996; **4** (Pt 1):153–158.

18 Innis RB, Cunningham VJ, Delforge J, Fujita M, Gjedde A, Gunn RN, *et al.* Consensus nomenclature for in vivo imaging of reversibly binding radioligands. *J Cereb Blood Flow Metab* 2007; **27**:1533–1539.

19 De Jong HW, van Velden FH, Kloet RW, Buijs FL, Boellaard R, Lammertsma AA. Performance evaluation of the ECAT HRRT: an LSO-LYSO double layer high resolution, high sensitivity scanner. *Phys Med Biol* 2007; **52**:1505–1526.

20 Comtat C, Bataille F, Michel C, Jones JP, Sibomana M, Janeiro L, *et al.* OSEM-3D reconstruction strategies for the ECAT HRRT. *IEEE Nucl Sci Symp Conf Record* 2004; **6**:3492–3496.

21 Comtat C, Sureau FC, Sibomana M, Hong IK, Sjoholm N, Trebossen R. Image based resolution modeling for the HRRT OSEM reconstructions software. *IEEE Nucl Sci Symp Conf Record* 2008; **8**:4120–4123.

22 Nuyts J, Dupont P, Stroobants S, Maes A, Mortelmans L, Suetens P. Evaluation of maximum-likelihood based attenuation correction in positron emission tomography. *IEEE Trans Nucl Sci* 1999; **46**:1136–1141.

23 Watson CC. New, faster, image-based scatter correction for 3D PET. *IEEE Trans Nucl Sci* 2000; **47**:1587–1594.

24 Byars LG, Sibomana M, Burbar Z, Jones J, Panin V, Barker WC, *et al.* Variance reduction on randoms from coincidence histograms for the HRRT. *IEEE Nucl Sci Symp Conf Record* 2005; **5**:2622–2626.

25 Ashburner J, Friston KJ. Unified segmentation. *Neuroimage* 2005; **26**:839–851.

26 Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *Neuroimage* 2012; **62**:782–790.

27 Ashburner J, Friston KJ. Nonlinear spatial normalization using basis functions. *Hum Brain Mapp* 1999; **7**:254–266.

28 Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, *et al.* Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 2002; **33**:341–355.

29 Alakurtti K, Johansson JJ, Tuokkola T, Någren K, Rinne JO. Rostrocaudal gradients of dopamine D2/3 receptor binding in striatal subregions measured with [(^{11})C]raclopride and high-resolution positron emission tomography. *Neuroimage* 2013; **82**:252–259.

30 Juslin A, Tohka J. Unsupervised segmentation of cardiac PET transmission images for automatic heart volume extraction. *Conf Proc IEEE Eng Med Biol Soc* 2006; **1**:1077–1080.

31 Dhillon IS, Guan Y, Kulis B. Weighted graph cuts without eigenvectors a multilevel approach. *IEEE Trans Pattern Anal Mach Intell* 2007; **29**:1944–1957.

32 Tohka J, Kivimaki A, Reilhac A, Mykkanen J, Ruotsalainen U. Assessment of brain surface extraction from PET images using Monte Carlo Simulations. *IEEE Trans Nucl Sci* 2004; **51**:2641–2648.

33 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**:420–428.

34 Kuhn FP, Warnock GI, Burger C, Ledermann K, Martin-Soelch C, Buck A. Comparison of PET template-based and MRI-based image processing in the quantitative analysis of C11-raclopride PET. *EJNMMI Res* 2014; **4**:7.

35 Rinne JO, Hietala J, Ruotsalainen U, Sako E, Laihinen A, Nagren K, *et al.* Decrease in human striatal dopamine D2 receptor density with age: a PET study with [^{11}C]raclopride. *J Cereb Blood Flow Metab* 1993; **13**:310–314.

36 McDonald WM, Husain M, Doraiswamy PM, Figiel G, Boyko O, Krishnan KR. A magnetic resonance image study of age-related changes in human putamen nuclei. *Neuroreport* 1991; **2**:57–60.

37 Morris ED, Chefer SI, Lane MA, Muzic RF Jr, Wong DF, Dannals RF, *et al.* Loss of D2 receptor binding with age in rhesus monkeys: importance of correction for differences in striatal size. *J Cereb Blood Flow Metab* 1999; **19**:218–229.

38 Van Velden FH, Kloet RW, van Berckel BN, Buijs FL, Luurtsema G, Lammertsma AA, Boellaard R. HRRT versus HR+ human brain PET studies: an interscanner test-retest study. *J Nucl Med* 2009; **50**:693–702.

39 Schain M, Varnäs K, Cselényi Z, Halldin C, Farde L, Varrone A. Evaluation of two automated methods for PET region of interest analysis. *Neuroinformatics* 2014; **12**:551–562.

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O.B. 527
FI-33101 Tampere, Finland

ISBN 978-952-15-3778-3
ISSN 1459-2045