

ORIGINAL RESEARCH

Open Access



Model selection criteria for dynamic brain PET studies

Sandeep S. V. Golla^{1*†}, Sofie M. Adriaanse^{1†}, Maqsood Yaqub¹, Albert D. Windhorst¹, Adriaan A. Lammertsma¹, Bart N. M. van Berckel¹ and Ronald Boellaard^{1,2}

* Correspondence: s.golla@vumc.nl
Sandeep S. V. Golla and Sofie M. Adriaanse contributed equally to this work.

[†]Equal contributors

¹Department of Radiology and Nuclear Medicine, VU University Medical Center, P.O. 7057, 1007, MB, Amsterdam, The Netherlands
Full list of author information is available at the end of the article

Abstract

Background: Several criteria exist to identify the optimal model for quantification of tracer kinetics. The purpose of this study was to evaluate the correspondence in kinetic model preference identification for brain PET studies among five model selection criteria: Akaike Information Criterion (AIC), AIC unbiased (AIC_C), model selection criterion (MSC), Schwartz Criterion (SC), and F-test.

Materials and Methods: Six tracers were evaluated: [¹¹C]FMZ, [¹¹C]GMOM, [¹¹C]PK11195, [¹¹C]Raclopride, [¹⁸F]FDG, and [¹¹C]PHT, including data from five subjects per tracer. Time activity curves (TACs) were analysed using six plasma input models: reversible single-tissue model (1T2k), irreversible two-tissue model (2T3k), and reversible two-tissue model (2T4k), all with and without blood volume fraction parameter (V_B). For each tracer and criterion, the percentage of TACs preferring a certain model was calculated.

Results: For all radiotracers, strong agreement was seen across the model selection criteria. The F-test was considered as the reference, as it is a frequently used hypothesis test. The F-test confirmed the AIC preferred model in 87% of all cases. The strongest (but minimal) disagreement across regional TACs was found when comparing AIC with AIC_C. Despite these regional discrepancies, same preferred kinetic model was obtained using all criteria, with an exception of one FMZ subject.

Conclusion: In conclusion, all five model selection criteria resulted in similar conclusions with only minor differences that did not affect overall model selection.

Keywords: Positron emission tomography, Brain imaging, Molecular imaging, Pharmacokinetics, Imaging

Background

The full potential of a novel PET tracer can only be achieved when the uptake or clearance of the tracer can be extracted from the normal activity distribution of the dynamic PET data. To evaluate this, a pharmacokinetic model describing the tracer in vivo kinetics is necessary. With a kinetic model, the in vivo physiological behaviour of the tracer both in tissue and blood can be described. The model that best describes the kinetic behaviour of the tracer may be used to reproducibly and reliably quantify tracer distribution or binding. Finding the optimal model can, however, be a challenge.

Several criteria exist to identify the preferred kinetic model that provides the best fit to the time course of the tissue tracer concentration [1] [2] [3] [4] [5]. Each of these

criteria ranks the quality of the model fit and thus provides a means for model selection. These criteria are all based on the goodness of fit (sum of squared differences between measured and estimated data, also termed residuals) and the complexity of the model (number of fit parameters). In addition, both number of frames and weighting factors applied to residuals are taken into account. The various criteria have substantial similarities in their equations and parameters used (Eqs. 1–4). For example, Eq. 3 can be written as the difference between \ln of nominator and denominator; this now is quite similar to other equations. The only difference is in the penalty parameter (added based on the complexity of model) in these equations. We therefore hypothesize that these criteria would result in (overall) comparable model preferences in actual clinical data. To the best of our knowledge, a direct comparison of their performance in clinical PET brain studies has never been reported and the selection of the criterion used is frequently a matter of debate.

The purpose of the present study was to evaluate the similarity and/or difference in model preference for several commonly used criteria: Akaike Information Criterion (AIC) [6], [1], AIC unbiased (AIC_C) [7], [2], Model Selection Criterion (MSC) [8], [3], Schwartz Criterion (SC) [9], [4], and finally validate them against the F-test [10] [5]. To cover a range of kinetics, these criteria were used to identify the preferred model for six tracers with known differences in kinetic (in vivo) properties.

Methods

Six neuroPET tracers with different kinetic properties were evaluated: [¹¹C]Flumazenil ([¹¹C]FMZ), [¹¹C]N-(2-chloro-5-thiomethylphenyl)-N'-(3-methoxy-phenyl)-N'-methylguanidine ([¹¹C]GMOM), [¹¹C](1-[2-chlorophenyl]-N-methyl-N-[1-methyl-propyl]-3-isoquinoline carboxamide) ([¹¹C]PK11195), [¹¹C]Raclopride, [¹⁸F]Fluorodeoxyglucose ([¹⁸F]FDG), and [¹¹C]Phenytoin ([¹¹C]PHT). GABA (A) receptor binding, NMDA receptor binding, TSPO binding, dopamine receptor binding, metabolic rate of glucose, and function of Pgp transporters can be deduced using the following PET tracers: [¹¹C]FMZ, [¹¹C]GMOM, [¹¹C]PK11195, [¹¹C]raclopride, [¹⁸F]FDG, and [¹¹C]PHT, respectively. Additional file 1: Figure S1 illustrates typical whole brain gray matter time activity curves (TACs) of the six tracers. A full description of the original studies has been reported previously for [¹¹C]FMZ [11], [¹¹C]PK11195 [12], [¹¹C]Raclopride [13], [¹⁸F]FDG [14], [¹¹C]GMOM [15], and [¹¹C]PHT [1]. For each tracer, PET scan data from five subjects were included. Subjects were randomly selected from existing databases, and researchers had no information on group status (both healthy subjects and patients). Five subjects per tracer reflect the number of subjects often included in first-time studies examining tracer kinetics. Each of these subjects had a T1-weighted magnetic resonance imaging (MRI) scan and a metabolite corrected arterial input function was also available. For scanner properties, attenuation correction, scan duration, and reconstruction methods, see Table 1. All scans were corrected for dead time, randoms, scatter, and decay. The original studies had all been approved by the Medical Ethics Review Committee of the VU University Medical Center. All subjects had provided written informed consent after complete explanation of the study procedures.

For anatomical delineation of volumes of interest (VOIs), MRI was used. Subjects' T1-weighted MRI scan was co-registered with the summed PET data using VINCI software (Cologne, Germany). The co-registered MRI scan was segmented into gray matter, white matter, and extra-cerebral fluid, and TACs were extracted using PVElab

Table 1 Overview of retrospective datasets used

	[¹¹ C]FMZ	[¹¹ C]PK11195	[¹¹ C]Raclopride	[¹⁸ F]FDG	[¹¹ C]GMOM	[¹¹ C]PHT
Scanner	ECAT EXACT HR+	ECAT EXACT HR+	ECAT EXACT HR+	ECAT EXACT HR+	Gemini TF-64 PET/CT	Gemini TF-64 PET/CT
Attenuation correction	10 min two-dimensional transmission scan	10 min two-dimensional transmission scan	10 min two-dimensional transmission scan	10 min two-dimensional transmission scan	Low-dose CT	Low-dose CT
Scan duration (min)	60	60.5	60	60	90	60
Number of frames	16	22	20	39	22	19
Reconstruction method	OSEM	FORE+ 2D-filtered back projection algorithm	FORE+ 2D-filtered back projection algorithm	OSEM	3-D RAMLA	3-D RAMLA
MRI	Siemens Sonata 1.5 T	Siemens 1.0 T IMPACT	Siemens Sonata 1.5 T	Siemens Magnetom Vision	Siemens Sonata 1.5 T	Siemens Sonata 1.5 T

ECAT EXACT HR+ scanner (CTI/Siemens, Knoxville, TN, USA). Gemini TF-64 PET/CT scanner (Philips Medical Systems, Cleveland, OH, USA). Philips Intera 1.5 T scanner (Philips Medical Systems, Best, The Netherlands). Siemens Sonata 1.5 T & Siemens 1.0 T IMPACT (Siemens, Erlangen, Germany)

3-D RAMLA three-dimensional row action maximum likelihood reconstruction algorithm, OSEM ordered subset expectation maximization

[16] in combination with the Hammers template [17]. Sixty-eight gray matter VOIs were delineated onto MRI, i.e. 68 TACs were extracted per subject. TACs were analysed using six plasma input models: a reversible single-tissue model (1T2k), an irreversible two-tissue model (2T3k), and a reversible two-tissue model (2T4k), all three both with and without additional blood volume parameter (V_B) [18].

For all six tracers, the preferred models across TACs were obtained using various model selection criteria. Equations below show the implementation of the various model selection criteria. Akaike Information Criterion (AIC) [[6]; Equation (Eq. 1), AIC unbiased (AIC_C) [[7]; Eq. 2], Model Selection Criterion (MSC) [[8]; Eq. 3], Schwartz Criterion (SC) [[9]; Eq. 4], and the F-test [[10]; Eq. 5].

$$AIC = n \ln \left(\sum_i w_i (y(t_i) - \hat{y}(t_i))^2 / n \right) + 2p \quad (1)$$

$$AIC_C = n \ln \left(\sum_i w_i (y(t_i) - \hat{y}(t_i))^2 / n \right) + 2p + \frac{2p(p+1)}{n-p-1} \quad (2)$$

$$MSC = \ln \frac{\sum_i w_i [y(t_i) - \bar{y}(t_i)]^2}{\sum_i w_i [y(t_i) - \hat{y}(t_i)]^2} - 2p/n \quad (3)$$

$$SC = n \ln \left(\sum_i w_i [y(t_i) - \hat{y}(t_i)]^2 / n \right) + p \ln(n) \quad (4)$$

$$F = \frac{(Q_1 - Q_2) / (p_2 - p_1)}{Q_2 / (n - p_2)} \quad (5)$$

where n represents the number of observations, p the number of parameters, w_i the weight applied to residual of i th acquisition, y are the measured values (PET scan), \hat{y} are the predicted (fitted) values, \bar{y} is the mean of the measured values, and t_i time of i th acquisition. For the F-test, Q_1 represents the sum of squares for the model with p_1 parameters, and Q_2 the sum of squares of model with p_2 parameters, where $p_1 < p_2$ is required. w_i ($w_i = 1/\sigma_i^2$) was estimated using the following equation [19]:

$\sigma^2 = \alpha \cdot \text{dcf} \cdot \text{dcf} \frac{T}{L}$, where σ^2 represents the variances for each frame and is calculated based on the whole scanner true counts (T); dcf is the decay correction factor, L represents frame length, and α is the proportionality constant signifying the variance level.

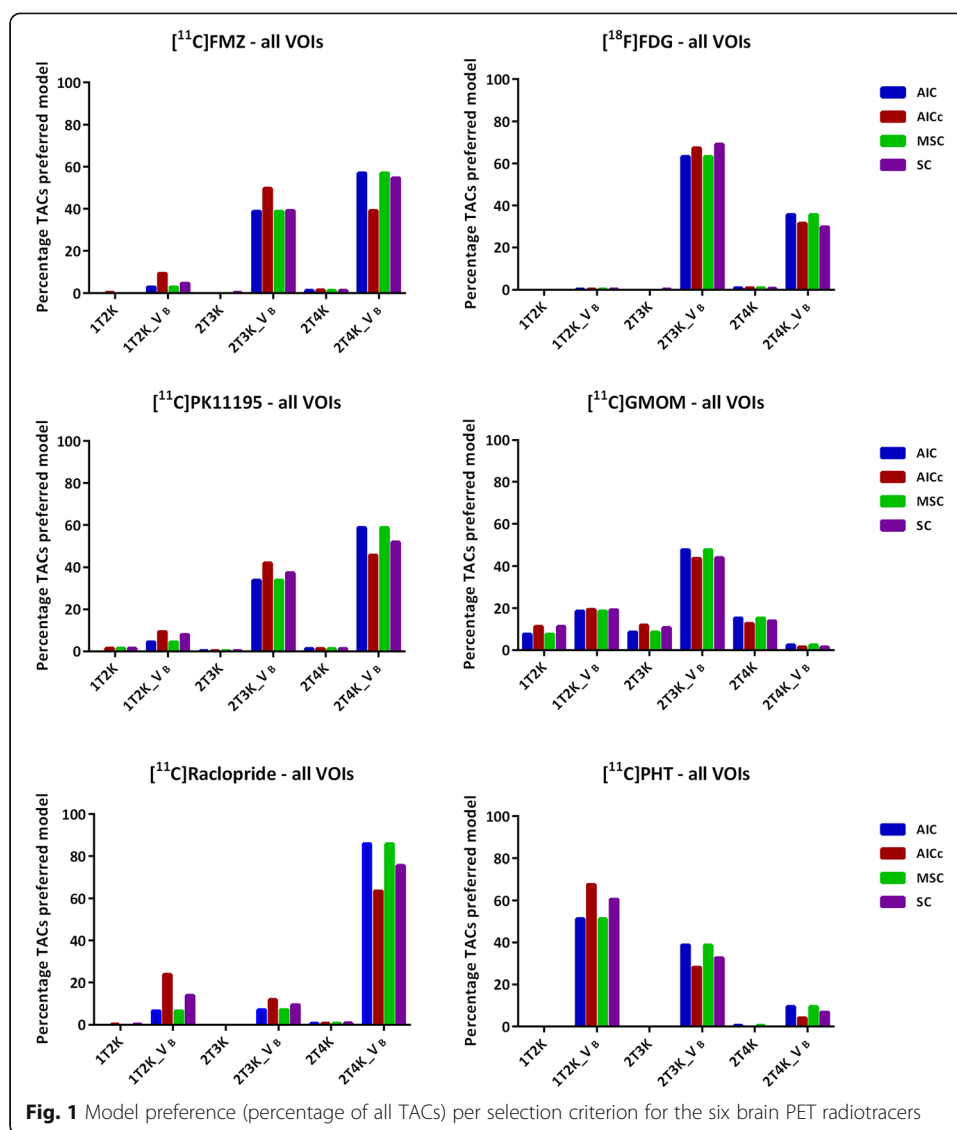
For, AIC, AIC_C, and SC, a lower value implies a better fit. Therefore, the lowest value obtained from the model selection criterion indicates the preferred kinetic model, i.e. the model that provides the best fit to a TAC with the smallest number of model parameters. For MSC, a higher value implies a better fit and hence the highest value will be indicative of the most appropriate model. F-test directly compares two models and returns a F-statistic with $p_2 - p_1$ and $n - p_2$ degrees of freedom. If this F-statistic is larger than the tabulated value at a specified p value, the complex model has a better fit (with a significance of $p < 0.05$).

Agreement of the model selection criteria was evaluated by calculating the percentage preference for a kinetic model across all TACs. This was performed for all subjects and for each model selection criteria and tracer. The F-test was considered as the reference for model comparison as it is a frequently used hypothesis test [20], [21]. This F-test, however, is difficult to apply when comparing multiple models, subjects, TACs, and tracer studies. For example, in the present study, this would have resulted in 61,200 comparisons (30 comparisons per tracer per subject per VOI). Therefore, a more pragmatic approach was followed by first identifying the most likely plausible model based on known tracer kinetics, visual inspection of the fits, and AIC results, which can be calculated very efficiently. Next, the F-test was applied to determine if this model was the preferred one (per TAC) when compared to all other models. Finally, regional agreement/disagreement of AIC_C, SC, and MSC to AIC and eventually F-test was evaluated. To examine if any existent disagreement between model selection criteria was a general effect or driven by a single subject, the frequency of model preferences per subject were also evaluated for each model selection criteria. Finally, in order to examine the effect of VOI size (assuming that smaller VOIs result in TACs with higher noise levels) on the model selection criteria, similar analyses were performed for VOIs with sizes smaller and larger than 5 mL separately.

Results

For all six radiotracers, strong agreement was observed between the different model selection criteria when examining the frequency of model preferences across all TACs (Fig. 1). Only one small deviation was observed for FMZ: AIC_C preferred 2T3k_V_B, whereas the other model selection criteria preferred 2T4k_V_B. The reduced regional agreement for FMZ when comparing AIC with AIC_C turned out to be an 18% difference in model preference. When examining within subject preference, it was shown that this difference was mainly driven by one subject. For this subject, 2T3k_V_B was preferred according AIC_C and F-test where AIC, MSC, and SC all preferred 2T4k_V_B for this subject. For the other remaining four FMZ studies, all five criteria agreed on 2T4k_V_B being the preferred model for FMZ.

Even though all model selection criteria agreed on the preferred model across TACs (with exception of one FMZ study), slight disagreements between criteria did exist. The F-test confirmed the AIC preferred models within single subjects for 87% of all cases. AIC and MSC gave exactly the same model preferences for all the tracers studied. Regional disagreement of only 13 and 7% for AIC_C and SC, respectively, with AIC was

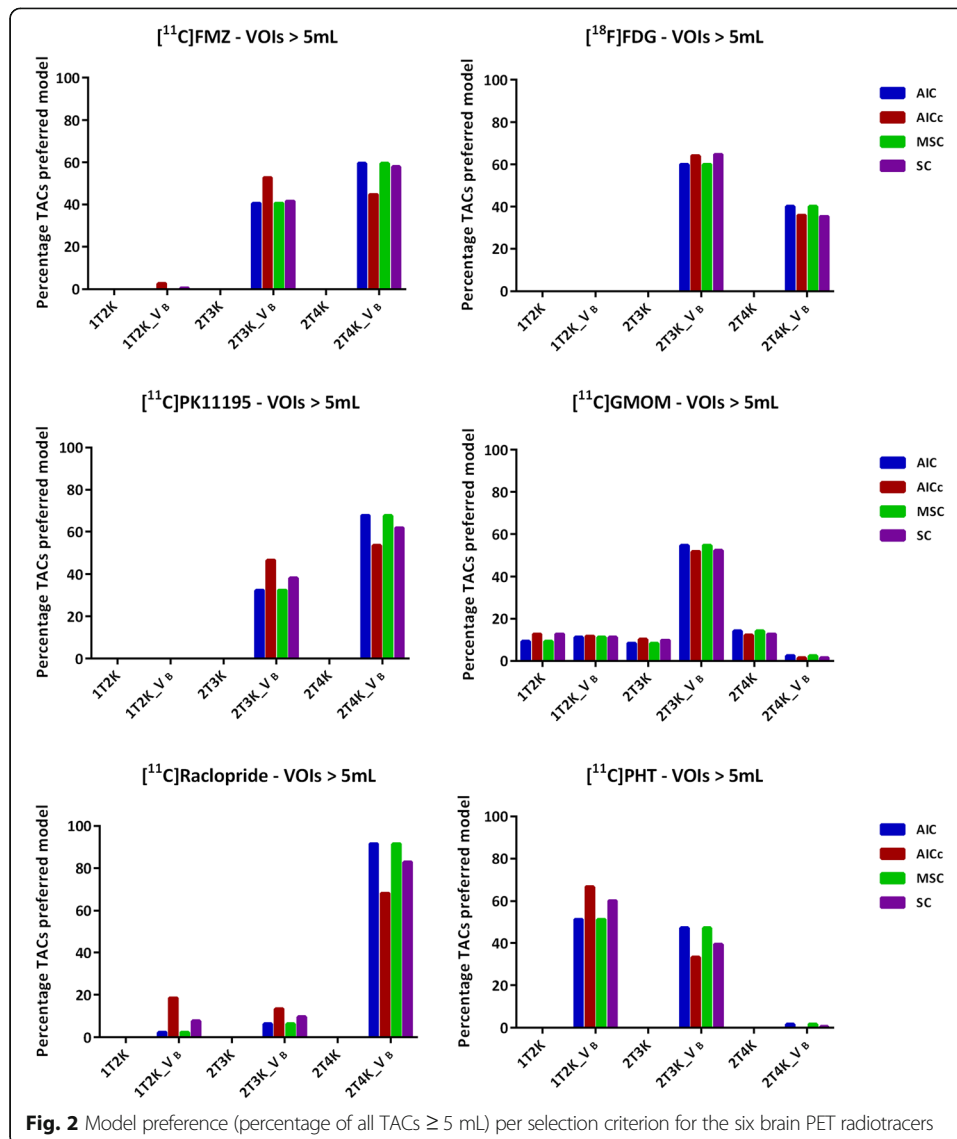


observed for $[^{11}\text{C}]\text{PK11195}$. In case of $[^{11}\text{C}]\text{Raclopride}$, regional disagreement of 22% when comparing AIC with AIC_C and 10% between AIC and SC was observed. For $[^{18}\text{F}]\text{FDG}$, relatively low disagreement of 4 and 6% for AIC_C and SC, respectively, was observed. A relatively low (4%) disagreement of AIC with both AIC_C and SC (4%) was obtained for $[^{11}\text{C}]\text{GMOM}$. A regional disagreement of 16% for AIC_C and 9% for SC with AIC was seen for $[^{11}\text{C}]\text{PHT}$. Despite regional differences, and between subject deviations, all model selection criteria showed strong agreement in model preference.

Effect of noise on the model preference can be observed in Figs. 2 and 3. For all the tracers, a variation in the model preference is observed with varying VOI size, which is primarily due to noise. However, the VOI size (either smaller or larger than 5 mL), in general, did not have a notable effect on the agreement of the model selection criteria (Figs. 2 and 3).

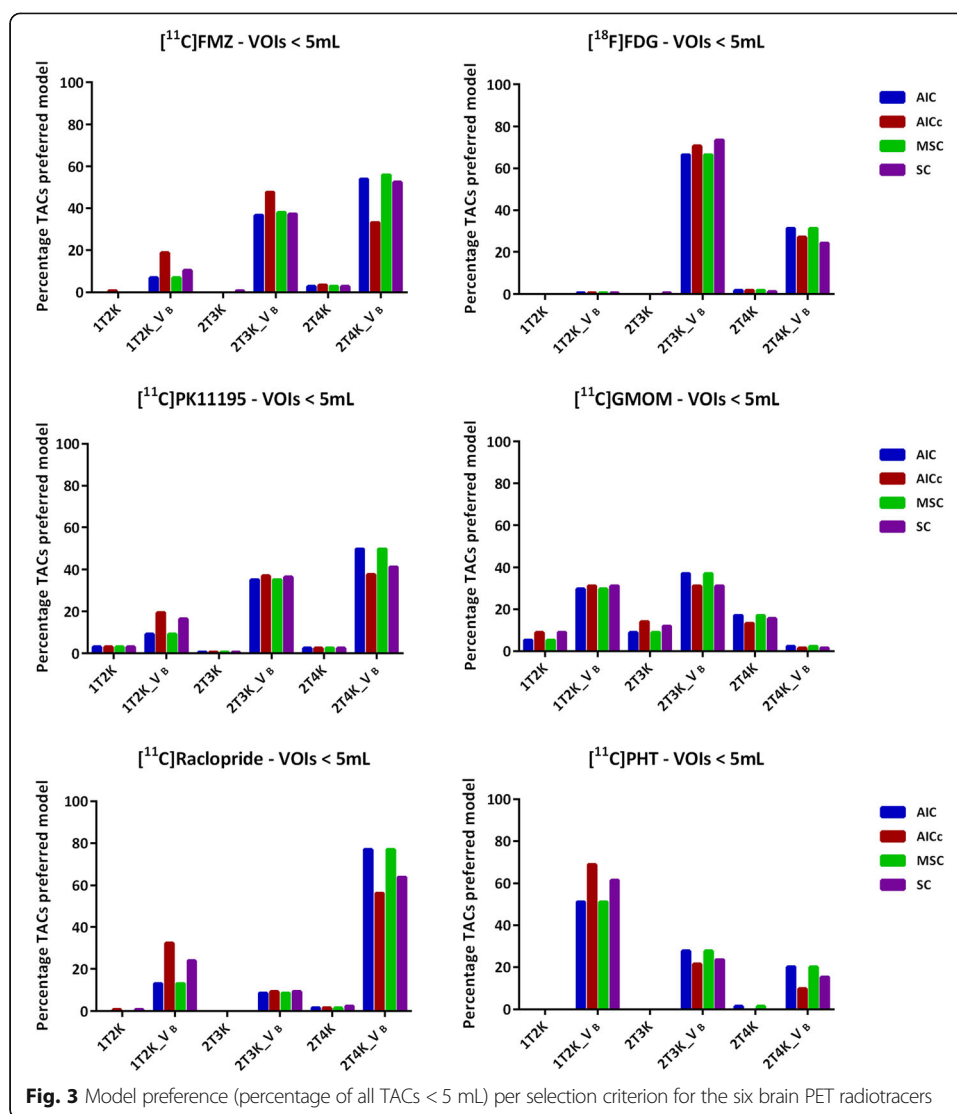
Discussion

The aim of the present study was to assess the overall agreement and/or disagreement in model preference of five commonly used model selection criteria. The model



selection criteria evaluated are all based on similar assumptions, possibly explaining that the choice for a certain model selection criterion hardly affected identification of the preferred model. For a mathematical evaluation of the underlying assumptions of model selection criteria and their considerations, see [22], [23], [24].

In short, small differences between criteria exist with respect to the penalty applied for the number of parameters and observations (frames). AIC, AIC_C, SC, and MSC all include an extra penalty term for the number of parameters, trying to account for the possibility of over-fitting the model. The penalty used in AIC_C seems to be stronger than that used in SC and MSC when compared to AIC. AIC becomes strongly biased when the ratio of model parameters to observations (acquisition frames) increases, for which AIC_C tries to correct. When the number of observations increases, AIC_C converges to AIC. Due to the stronger penalty, AIC_C seems to prefer the models with less parameters when the ratio is bigger. For [¹¹C]FMZ and [¹¹C]PHT, both with < 20 observations, AIC_C indeed preferred the model with less parameters slightly more frequently compared with the other criteria. For [¹¹C]PHT, this was reflected in a



regional disagreement between AIC and AIC_C of 16% of the TACs. For FMZ, this resulted in AIC_C identifying a different kinetic model when examining all VOIs across all subjects as illustrated in Fig. 1. However, this difference in model preference was mainly driven by a single subject where AIC_C disagreed strongly with the other model selection criteria, which was confirmed by the F-test. For the remaining four subjects, AIC_C and F-test preferred the same model for FMZ as did the other criteria.

For [¹¹C]PK11195, [¹¹C]raclopride, [¹⁸F]FDG, and [¹¹C]FMZ, the F-test did not agree with AIC for only four individual (subject) cases (one subject from each tracer study). For [¹¹C]PHT and [¹¹C]GMOM, the F-test agreed with AIC in all cases. The overall agreement for all tracers and subjects was 87% between AIC and F-test. However, on a (tracer) study level, this did not affect the model preference for that tracer. The F-test is designed to directly compare two models and is especially suited for this purpose. AIC and MSC yielded identical results. Compared with AIC, MSC includes normalization and is therefore independent of scaling of the data points. The normalization incorporated in MSC, however, seems to have little effect for the data described in this study.

Even though the model selection criteria showed strong overlap, it is not always straightforward which model provides the best fit for the data. For example, ~30% of all TACs clearly identified 2T4k_ V_B for [^{18}F]FDG, whereas the kinetics of all other TACs were described by 2T3k_ V_B .

A total of 68 gray matter VOIs were included per subject, implying that the analysis comprised of both receptor-rich and receptor-devoid VOIs (which varies depending on the target of interest/tracer) and thus the observed regional differences in model preferences (Figs. 1, 2, and 3). In addition to the receptor density, another reason for regional differences in the same subject for the same tracer could be the size of VOI. Smaller VOI tends to have lower counts and thereby higher noise, which makes it difficult to define the optimal pharmacokinetic model. Figure 2 illustrates the impact of noise on the model selection when using different model selection criteria. Noise could be a probable cause for the observed differences between Figs. 2 and 3.

Preferred model can be affected by several factors such as in vivo kinetics of the VOI (underlying receptor density), subject status, input function (particularly parent fraction and metabolites), tracer free fraction, motion, and even the scan duration. It should be noted that a description of tracer kinetics and the identification of the preferred kinetic model using model selection criteria is only one of the many evaluations needed to identify the optimal model. For example, in order to determine the optimal kinetic model, several other datasets and studies are required such as test-retest data, data for both healthy controls and diseased subjects, data correlating model results with pathology, and/or outcome (disease duration, cognitive scores, or survival). Also, the intended application, e.g. differential diagnosis or response assessment, may or may not allow for the use of simplified models or methods. Moreover, formulating a hypothesis on the preferred kinetic model based on physiological properties of the tracer, e.g. derived from preclinical studies, is advised.

The number of subjects included for each tracer might be a limitation for this study; however, typically first in man PET studies are limited in subject number and often restricted to healthy subjects. In these circumstances, use of model selection criteria provides a first indication of one or more suitable candidate kinetic models which should be further developed and evaluated. Even though both healthy controls and patients were included for the analysis, no impact on the conclusions is expected, as over a wide range of kinetics with quite consistent results in the performance of the various criteria was observed. Therefore, possible changes in tracer kinetics in healthy versus diseased subjects would likely have little impact on the performance of these criteria, but should be verified in individual cases.

Conclusions

All model selection criteria tested resulted in similar conclusions with only minor, non-relevant differences in overall observed model preference.

Additional file

Additional file 1: Figure S1. Typical whole brain gray matter TACs of all the six brain PET radiotracers. (TIFF 43 kb)

Abbreviations

[^{11}C]PHT: [^{11}C]Phenytol; [^{18}F]FDG: [^{18}F]Fluorodeoxyglucose; 1T2k: Reversible single-tissue model; 2T3k: Irreversible two-tissue model; 2T4k: Reversible two-tissue model; AIC: Akaike Information Criterion; dcf: Decay correction factor; FMZ: Flumazenil; GMOM: N-(2-Chloro-5-thiomethylphenyl)-N'-(3-methoxy-phenyl)-N'-methylguanidine; L: Frame

length; MSC: Model selection criterion; N : Number of observations; P : Number of parameters; PK11195: (1-[2-Chlorophenyl]-N-methyl-N-[1-methyl-propyl]-3-isoquinoline carboxamide); Q_1 : Sum of squares for the model with p_1 parameters; Q_2 : Sum of squares of model with p_2 parameters; SC: Schwartz Criterion; T: Whole scanner true counts; TACs: Time activity curves; t_i : Time of i th acquisition; V_B : Blood volume parameter; VOI: Volumes of interest; w_i : Weight applied to residual of i th acquisition; \bar{y} : Mean of the measured values; y : Measured values (PET scan); \hat{y} : Predicted (fitted) values; α : Proportionality constant signifying the variance level; σ^2 : Variances for each frame

Acknowledgements

The [^{11}C]GMOM study was supported by the Center for Translational Molecular Medicine (LeARN 02N-101) and European Union's Seventh Framework Programme (FP7/2007-2013), grant agreement no. HEALTH-F2-2011-278850 (INMiND). The [^{11}C]PK11195 study was supported by the EC (FP5-project NCI-MCI, QLK6-CT-2000-00502), The Netherlands Organisation for Scientific Research (VIDI grant 016.066.309), and by The Netherlands Brain Foundation (Grant no. 9F01.21). The [^{11}C]Raclopride study was financially supported by the EU Seventh Framework Programme EURIPIDES (FP7/2007-2013 under grant agreement no. 201380) and The Netherlands Organisation for Scientific Research (NWO), VIDI grant 016.066.309. The [^{11}C]PHT study was supported by the European Union Seventh Framework Program EURIPIDES (FP7/ 2007-2013 under grant agreement no. 201380). Additional support was provided by a scholarship from the Malaysian Ministry of Education and University Sains Malaysia. The [^{11}C]FMZ study was supported in part by a grant provided by ZonMw 907-00-012 (Dutch Organization for Health Research and Development) and Lundbeck BV, Amsterdam, The Netherlands.

Availability of data and materials

Sharing of data was not approved and not consented by the patients at the time of data collection and can therefore not be shared.

Authors' contributions

SSVG contributed to the project design, data analysis and interpretation, drafting, and final approval of the paper. SMA contributed to the data analysis and interpretation, drafting, and final approval of the paper. MY contributed to the data interpretation and review and final approval of the paper. ADW contributed to the tracer development, data acquisition, review, and final approval of the paper. AAL contributed to the data interpretation, review, and final approval of the paper. BNMB contributed to the data acquisition, review, and final approval of the paper. RB contributed to the project design, data analysis and interpretation, drafting, and final approval of the paper.

Authors' information

SSVG is a postdoctoral researcher at VU University Medical Center. SMA is a postdoctoral researcher at the University of Amsterdam. MY is a clinical physicist at VU University Medical Center. ADW is a professor and head of the Department of Chemistry at VU University Medical Center. AAL is a professor and head of research at VU University Medical Center. BNMB is a professor of molecular brain imaging at VU University Medical Center. RB is a professor and head of the Department of Physics at VU University Medical Center.

Ethics approval and consent to participate

The original studies had all been approved by the Medical Ethics Review Committee of the VU University Medical Center. All subjects had provided written informed consent after complete explanation of the study procedures.

Consent for publication

All subjects had provided written informed consent after complete explanation of the study procedures.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Radiology and Nuclear Medicine, VU University Medical Center, P.O. 7057, 1007, MB, Amsterdam, The Netherlands. ²Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.

Received: 6 April 2017 Accepted: 23 November 2017

Published online: 06 December 2017

References

- Mansor S, Boellaard R, Froklage FE, et al. Quantification of dynamic ^{11}C -Phenitoin PET studies. *J Nucl Med*. 2015; 56(9):1372–7.
- Hackett SL, Liu D, Chalkidou A, Marsden P, Landau D, Fenwick JD. Estimation of input functions from dynamic [^{18}F]FLT PET studies of the head and neck with correction for partial volume effects. *EJNMMI Res*. 2013;3(1):84.
- Zanotti-Fregonara P, Leroy C, Roumenov D, Trichard C, Martinot JL, Bottlaender M. Kinetic analysis of [^{11}C]befloxadone in the human brain, a selective radio ligand to image monoamine oxidase a. *EJNMMI Res*. 2013; 3(1):78.
- Ito H, Nyberg S, Halldin C, Lundkvist C, Farde L. PET imaging of central 5-HT_{2A} receptors with carbon-11-MDL 100,907. *J Nucl Med*. 1998;39(1):208–14.

5. Golla SS, Boellaard R, Oikonen V, et al. Quantification of [18F]DPA-714 binding in the human brain: initial studies in healthy controls and Alzheimer's disease patients. *J Cereb Blood Flow Metab.* 2015;35(5):766–72.
6. Akaike H. A new look at the statistical model identification. 19, 716–723. 1–1-1974. doi:10.1109/TAC.1974.1100705.
7. Sugiura N. Further analysis of the data by Akaike's information criterion and the finite corrections. *CommStatist [A 7].* 1978;7:13–26.
8. MicroMath SLMU. Statistics output: model selection criterion. 2016.
9. Schwarz G. Estimating the dimension of a model. *Ann Statist.* 1978;6(2):461–4.
10. Cunningham VJ. Non-linear regression techniques in data analysis. *Med Inform (Lond).* 1985;10(2):137–42.
11. Klumpers UM, Veltman DJ, Boellaard R, et al. Comparison of plasma input and reference tissue models for analysing [(11C)flumazenil studies. *J Cereb Blood Flow Metab.* 2008;28(3):579–87.
12. Schuitemaker A, van Berckel BN, Kropholler MA, et al. Evaluation of methods for generating parametric (R)-[11C]PK11195 binding images. *J Cereb Blood Flow Metab.* 2007;27(9):1603–15.
13. van Velden FH, Mansor SM, van Assema DM, et al. Comparison of HRRT and HR+ scanners for quantitative (R)-[11C]verapamil, [11C]raclopride and [11C]flumazenil brain studies. *Mol Imaging Biol.* 2015;17(1):129–39.
14. Boellaard R, van Lingen A, Lammertsma AA. Experimental and clinical evaluation of iterative reconstruction (OSEM) in dynamic PET: quantitative characteristics and effects on kinetic modeling. *J Nucl Med.* 2001;42(5):808–17.
15. van der Doef TF, Golla SS, Klein PJ, et al. Quantification of the novel N-methyl-d-aspartate receptor ligand [11C]GMOM in man. *J Cereb Blood Flow Metab.* 2016;36(6):1111–21.
16. Svarer C, Madsen K, Hasselbalch SG, et al. MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. *NeuroImage.* 2005;24(4):969–79.
17. Hammers A, Allom R, Koepp MJ, et al. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp.* 2003;19(4):224–47.
18. Gunn RN, Gunn SR, Cunningham VJ. Positron emission tomography compartmental models. *J Cereb Blood Flow Metab.* 2001;21(6):635–52.
19. Yaqub M, Boellaard R, Kropholler MA, Lammertsma AA. Optimization algorithms and weighting factors for analysis of dynamic PET studies. *Phys Med Biol.* 2006;51(17):4217–32.
20. Kletting P, Glatting G. Model selection for time-activity curves: the corrected Akaike information criterion and the F-test. *Z Med Phys.* 2009;19(3):200–6.
21. Glatting G, Kletting P, Reske SN, Hohl K, Ring C. Choosing the optimal fit function: comparison of the Akaike information criterion and the F-test. *Med Phys.* 2007;34(11):4285–92.
22. Burnham K, Anderson D. Model selection and multimodel inference: a practical information-theoretic approach. New York: Springer-Verlag; 2002.
23. Schoniger A, Wohling T, Samaniego L, Nowak W. Model selection on solid ground: rigorous comparison of nine ways to evaluate Bayesian model evidence. *Water Resour Res.* 2014;50(12):9484–513.
24. Motulsky H, Christopoulos A. Fitting models to biological data using linear and nonlinear regression. A practical guide to curve fitting. San Diego: GraphPad Software Inc.; 2004.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
