

# Measurement of inter- and intra-observer variability in the routine clinical interpretation of brain 18-FDG PET-CT

Nicolas Brucher · Ramin Mandegaran ·  
Thomas Filleron · Thomas Wagner

Received: 9 July 2014 / Accepted: 16 November 2014 / Published online: 22 November 2014  
© The Japanese Society of Nuclear Medicine 2014

## Abstract

**Objective** To objectify and quantify inter- and intra-observer variability of brain 18-FDG PET-CT interpretation in the context of cognitive and functional impairment amongst the elderly.

**Methods** 25 patients underwent brain 18-FDG PET-CT for investigation of dementia/MCI and frail elderly patients. Three observers interpreted studies in two forms: standardised datasets reconstructed by an outside observer and individualised reconstructions. Observers graded regional 18-FDG uptake in 11 brain areas and gave overall impressions on studies as pathological/normal. One observer repeated this process following a 3-month interval. The Kappa statistic was used to calculate inter- and intra-observer agreement on grading of regional 18-FDG uptake and overall impressions of studies as pathological/normal.

**Results** Moderate inter-observer agreement was observed across standardised and individualised dataset reconstructions when 11 regional brain areas were compared cumulatively and overall impressions on studies were given as pathological vs normal. Higher levels of inter-observer

agreement were found when comparing high versus low grading of regional uptake and when reporting standardised reconstructions. Intra-observer agreement between standardised vs individualised dataset reconstructions were moderate-to-fair across 11 brain regions cumulatively. Temporal intra-observer agreement of individualised dataset reconstructions comparing normal vs pathological opinions showed strong agreement ( $\kappa = 0.884$  [95 % CI 0.662; 1.000]).

**Conclusion** Despite a strong agreement in final diagnosis, this study demonstrates a moderate inter- and substantial intra-observer reproducibility in reporting brain 18-FDG PET-CT. Such results suggest that the visual analysis approach is different between nuclear physicians but leads to the same final diagnosis.

**Keywords** Brain FDG PET-CT · Inter-observer variability · Intra-observer variability · Dementia · Mild cognitive impairment

## Introduction

Current prevalence of dementia worldwide is estimated to be in excess of 35 million people, with nearly 70 % of these diagnoses being attributable to Alzheimer's disease [1]. The total estimated worldwide cost of dementia currently stands at US\$604 billion. With global prevalence expected to exceed 115 million people by 2050, the scale of morbidity and cost attributable to dementia is set to increase dramatically [1]. Furthermore, recent estimates suggest almost one quarter of over 70 US population demonstrates mild cognitive impairment (MCI), considered to represent a prodromal phase of dementia where 10 % per year will progress to dementia [2].

---

N. Brucher  
Department of Radiology, Toulouse University Hospital,  
Toulouse, France

R. Mandegaran (✉) · T. Wagner  
Department of Nuclear Medicine, Royal Free London NHS  
Foundation Trust, Pond Street, London NW3 2QG, UK  
e-mail: ramin1986@hotmail.com

T. Filleron  
Medical Statistics, Toulouse Claudius Regaud Cancer Centre,  
Toulouse, France

In an era where drug discovery is targeted towards the development of disease modifying agents, methods of reliably differentiating Alzheimer's from other forms of dementia and establishing a diagnosis early in the disease process are essential for inclusion in clinical trials and gaining the maximum benefit from potential disease modifying therapies [3]. Establishing an etiological diagnosis is challenging, based on suggestive clinical-radiological findings with imaging usually restricted to CT or MRI.

18-FDG PET-CT has become a standard imaging modality in dementia research and increasingly used in dementia diagnosis and follow-up. Studies have demonstrated significantly improved diagnostic accuracy relative to MRI, effectiveness particularly in differentiating between Alzheimer's and frontotemporal dementia (FTD), whilst detection of abnormal regional cerebral glucose metabolism in MCI highlights the potential of 18-FDG PET-CT in establishing early diagnosis and initiating disease modifying treatment [4–8].

Interpretation of brain 18-FDG PET-CT is challenging, requiring understanding of normal variants and different pathological presentations to form an accurate diagnosis. A high intra- and inter-observer reproducibility is essential to ensure consistency in reporting by individuals and across centres, thereby permitting the proper conduct of multi-centre studies and establishing brain 18-FDG PET-CT brain as a reliable diagnostic tool. To date, only Hoffman et al. [9] have examined the intra- and inter-observer variability of brain 18-FDG PET, prior to the advent of PET-CT. They identified high intra-observer and high inter-observer interpretation agreement for bilateral temporoparietal hypometabolism associated with patients considered clinically to have probable Alzheimer's disease [9]. With advances in PET camera technology and the benefits afforded by correlation with CT findings, no study has since examined the intra- and inter-observer variability of cerebral 18-FDG PET-CT.

Our aim was to assess the intra- and inter-observer variability of cerebral 18-FDG PET-CT amongst a population of patients without cognitive deficit or with clinical evidence of dementia or MCI. To our knowledge, this is the first study attempting to systematically quantify these variables in the context of brain 18-FDG PET-CT.

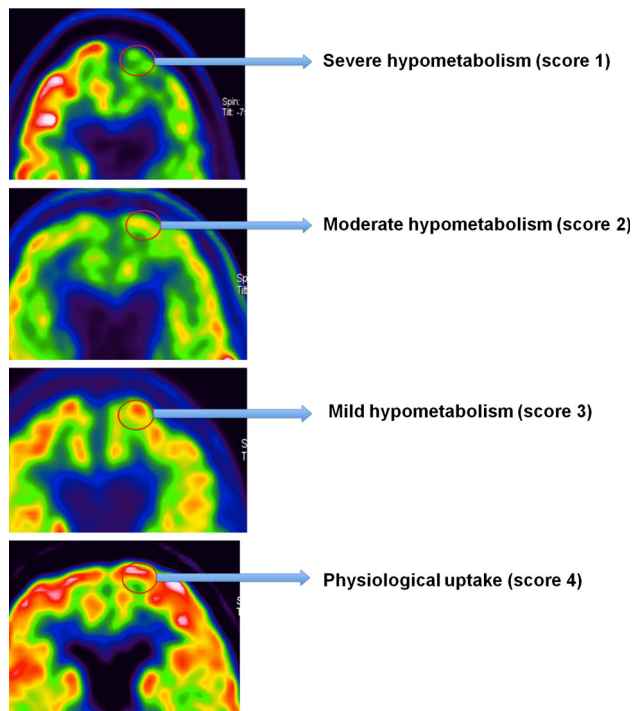
## Methods

This monocentric prospective study included 25 consecutive patients who all underwent brain 18-FDG PET-CT at Toulouse Purpan University Hospital in the context of assessment for dementia, MCI or frail elderly patients

between 14 November 2010 and 15 December 2010 (11 male and 14 female, mean age 72.5 years). Specifically, 13 patients were referred for 18-FDG PET-CT in the context of dementia or MCI assessment: 3 of these patients were referred for the investigation of speech disorders, 3 because of worsening memory impairment, 5 for behavioural disorders and 2 for disorders of praxis. The remaining 12 'frail elderly' patients were considered cognitively normal but referred for baseline studies as part of The Multidomain Alzheimer Prevention Trial (MAPT), all of whom fulfilled the Fried frailty criteria (3 or more of the following: (i) significant unintentional weight loss; (ii) self-reported exhaustion; (iii) objective physical weakness; (iv) slow walking speed; and (v) low physical activity). None of the patients referred had a pre-existing established diagnosis of dementia [10, 11]. The study was organised as an audit of inter-observer variability and therefore, no ethics committee approval was needed.

Image acquisition was performed on a Siemens Biograph 6 camera with an average injected activity of 1.85 MBq/kg (0.5 mCi/kg) in patients fasting for at least 4 h. After 30 min of rest, PET acquisition was performed for 10 min in list mode. We employed the following acquisition and reconstruction parameters: dynamic acquisition of 66 s; a 336 × 336 matrix; iterative reconstruction with OSEM 3D using 3 iterations and 21 subsets, filter all pass. CT parameters were: 130 kV and 50 mAs. After anonymising patients, standardised reconstructions were carried out by an outside observer with 24 axial slices of 3 mm thickness every 5 mm realigned along the bicommissural plane. The colour scale was standardised and normalised to basal ganglia as the site of maximal uptake.

Three experienced nuclear medicine physicians successively interpreted the same dataset in two different forms without the aid of any prior clinical information: the standardised dataset reconstructed by a single outside observer (also an experienced nuclear medicine physician) and the individualised dataset of the same patients (observers were free to realign, normalise and view the data as their preference for the individualised review). All three observers were experienced in interpretation of 18-FDG PET-CT having each reported at least 200 previous studies for similar indications. A numerical scale from 0 to 4 was used to assess degree of brain metabolism (4 = normal physiological uptake, 3 = mild hypometabolism, 2 = moderate hypometabolism, 1 = severe hypometabolism, 0 = no uptake) (Fig. 1). Regional assessment was considered normal or near normal (score 3 and score 4) and abnormal (scores 0, 1 and 2). Observers assigned a numerical score to 11 previously defined areas as follows:



**Fig. 1** Examples of frontal lobe FDG uptake and corresponding visual interpretations of degree of uptake with corresponding assigned uptake scores of 1–4

- Frontal
- Mesial frontal
- Anterior temporal
- Medial temporal
- External temporal
- Parietal
- Occipital
- Posterior cingulate area
- Striatum
- Thalamus
- Cerebellum

Observers then classified examinations as normal or pathological. The above interpretation process was then repeated by one of the observers after a 3-month interval. The Kappa agreement score was used to assess intra-observer agreement both between dataset reconstructions (standardised versus individualised reconstructions) as well as temporally (initial interpretation versus interpretation at 3-month interval). Inter-observer agreement was also assessed using both standardised and individualised datasets ( $P < 0.05$ ).  $\kappa$  magnitude of agreement has been classified by Landis and Koch as 0–0.2 = slight; 0.21–0.40 = fair; 0.41–0.60 = moderate; 0.61–0.80 = substantial; 0.81–1.00 = near perfect agreement [12].

**Table 1** Inter-observer  $\kappa$  values with 95 % CI for all individual and cumulative brain areas as well as global grouping of normal/near normal vs abnormal gradings and overall observer opinions on normal vs pathological reconstructions

	Standardised dataset reconstructions $\kappa$ (95 % CI)	Individualised dataset reconstructions $\kappa$ (95 % CI)
Frontal	0.523 (0.317; 0.711)	0.297 (0.127; 0.467)
Mesial frontal	0.379 (0.239; 0.508)	0.081 (−0.023; 0.203)
Anterior temporal	0.296 (0.149; 0.469)	−0.035 (−0.167; 0.100)
Medial temporal	0.031 (−0.083; 0.167)	−0.136 (−0.228; 0.048)
External temporal	0.330 (0.166; 0.511)	0.262 (0.105; 0.456)
Parietal	0.273 (0.136; 0.434)	0.344 (0.049; 0.591)
Occipital	0.252 (0.082; 0.395)	0.372 (0.224; 0.493)
Posterior cingulate area	0.364 (0.173; 0.526)	0.368 (−0.029; 0.588)
Striatum	0.137 (−0.059; 0.424)	0.144 (−0.018; 0.241)
Thalamus	0.349 (0.058; 0.654)	−0.027 (−0.059; 0.007)
Cerebellum	0.275 (0.123; 0.457)	0.138 (−0.030; 0.334)
Cumulative 11 brain areas	0.346 (0.302; 0.389)	0.212 (0.158; 0.262)
Global grouping 0/1/2 (normal/near normal) vs 3/4 (abnormal)	0.542 (0.462; 0.620)	0.346 (0.261; 0.433)
Overall normal vs pathological	0.307 (0.085; 0.588)	0.333 (0.058; 0.640)

Values shown for both standardised and individualised dataset reconstructions

## Results

### Inter-observer agreement

Inter-observer agreement for standardised dataset reconstructions of all individual brain areas showed a  $\kappa$  value of 0.346 for numerical values and of 0.542 for the normal/near normal vs abnormal assessment (grading 3/4 vs 0/1/2) (Table 1). Inter-observer agreement for individualised dataset reconstructions of quantitative regional uptake demonstrated a  $\kappa$  value of 0.212 for numerical values. Assessment of normal/near normal vs abnormal assessment (grading 3/4 vs 0/1/2) for individualised reconstructions showed a  $\kappa$  value of 0.333. The medial temporal lobe demonstrated  $\kappa$  values of 0.031 (slight agreement) and −0.136 (slight disagreement) for standardised and individualised dataset reconstructions, respectively.

Overall observer opinions on whether scans were normal or pathological demonstrated fair agreement for both standardised and individualised dataset reconstructions. When global grouping of assigned scores was performed (0/1/2 = normal/near normal vs 3/4 = abnormal), moderate

(0.542) and fair (0.346) agreement were demonstrated for standardised and individualised dataset reconstructions, respectively. Inter-observer Kappa value agreement scores are demonstrated in Fig. 2.

#### Intra-observer agreement: standardised vs individualised dataset reconstructions

Observers 1 and 2 demonstrated moderate intra-observer agreement between standardised and individualised dataset reconstructions when all 11 brain areas were considered cumulatively ( $\kappa = 0.438$  and  $0.449$ , respectively) (Table 2). Observer 3 demonstrated moderate intra-observer agreement and  $\kappa$  value of 0.546.

$\kappa$  values were lower in observers 1 and 3 (0.348 and 0.375, respectively) when overall intra-observer opinions on whether a scan was normal or pathological on standardised and individualised dataset reconstructions were compared, whilst agreement was considerably higher in observer 2 ( $\kappa = 0.669$ , i.e. substantial agreement).

#### Intra-observer agreement: temporal agreement

Following a 3-month interval and re-reporting the same standardised dataset reconstructions, overall  $\kappa$  values between the interval reports for the single observer were 0.555 when all 11 brain areas were considered

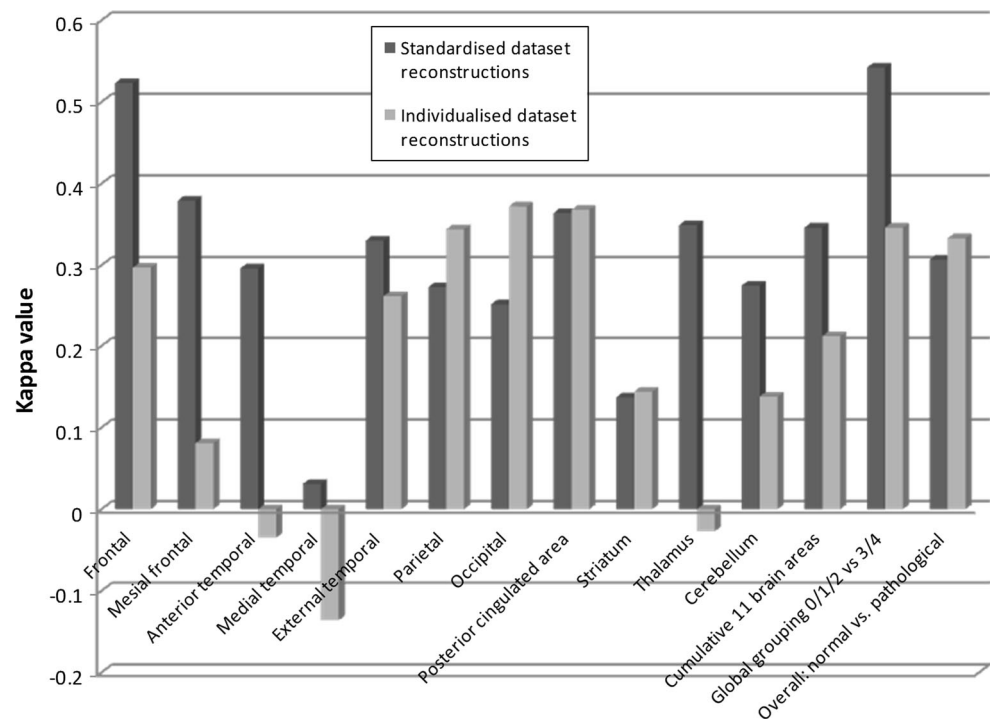
cumulatively (moderate agreement) (Table 3).  $\kappa$  values were 0.4 when interval individualised reconstructions were compared (moderate agreement).

A  $\kappa$  value of 0.348 was found between standardised interval reports when overall normal vs pathological opinions were compared (fair agreement) and of 0.884 when interval individualised reconstructions were compared (near perfect agreement). For normal/near normal vs abnormal assessment (grading  $\frac{3}{4}$  vs 0/1/2),  $\kappa$  values showed substantial agreement: 0.672 for individualised reconstructions and 0.728 for standardised reconstructions.

## Discussion

Our study identified differences in levels of inter-observer agreement when reporting brain FDG PET-CT as assessed with  $\kappa$  values. These differences depended on the type of reconstruction, the comparison of regional quantitative uptake, whether a scan was normal or pathological and whether regional uptake was abnormal or showed normal uptake or mild hypometabolism. Of interest there was significantly better inter-observer agreement when grading-11-defined brain areas were classified as normal/near normal or abnormal (grading  $\frac{3}{4}$  vs 0/1/2) than when analysing quantitative regional uptake. This is likely to represent subjective variation between observers when asked whether uptake in a brain area is normal or has mild, moderate

**Fig. 2** Inter-observer  $\kappa$  values



**Table 2** Intra-observer  $\kappa$  values with 95 % CI across all 3 observers comparing standardised vs individualised dataset reconstructions

	Observer 1 $\kappa$ (95 % CI)	Observer 2 $\kappa$ (95 % CI)	Observer 3 $\kappa$ (95 % CI)
Cumulative 11 brain areas	0.438 (0.368; 0.509)	0.449 (0.369; 0.521)	0.546 (0.464; 0.628)
Overall opinion: normal vs pathological	0.348 (0.089; 0.606)	0.669 (0.376; 0.962)	0.375 (0.034; 0.716)

**Table 3** Intra-observer  $\kappa$  values with 95 % CI following 3-month interval re-reporting of standardised and individualised dataset reconstructions

	Standardised vs standardised reconstruction 3-month interval $\kappa$ (95 % CI)	Individualised vs individualised reconstruction 3-month interval $\kappa$ (95 % CI)
Cumulative 11 brain areas	0.555 (0.484; 0.621)	0.400 (0.311; 0.489)
Global grouping 0/1/2 (normal/near normal) vs 3/4 (abnormal)	0.728 (0.634; 0.821)	0.672 (0.530; 0.814)
Overall normal vs pathological	0.348 (0.089; 0.606)	0.884 (0.662; 1.000)

or severe hypometabolism. It is, however, reassuring that moderate agreement is consistently achieved across brain regions when identifying moderate and severe uptake vs normal or mild hypometabolism.

Overall standardised reconstructions showed higher overall inter-observer agreement. However, given that routine clinical interpretation involves individualised dataset reconstructions, these findings raise the issue of what is the most accurate, reliable and reproducible way of reporting brain FDG PET-CT findings.

Inter-observer agreement improved only slightly amongst individualised reconstructions when observers gave an overall impression of the studies as either normal or pathological ( $\kappa = 0.333$ ). This finding is surprising but our impression is that observers would classify studies with mild hypometabolism as pathological, therefore potentially overcalling studies as abnormal. This is a potential source of inter-observer disagreement as some areas such as the medial temporal cortex are known as difficult to interpret. This is due to its small size, its physiological uptake being lower than the remainder of the cortex (increasing difficulty by the absence of an easily accessible reference standard for this region) and its anatomical orientation. Indeed our regional quantitative uptake inter-observer agreement for the medial temporal cortex demonstrated slight inter-observer disagreement (Medial temporal  $\kappa = -0.136$  for the individualised reconstruction and 0.031 for the standardised reconstruction).

Another very interesting finding was the high intra-observer reproducibility when re-reporting the same studies 3 months later, with moderate agreement for regional quantitative uptake, substantial agreement ( $\kappa$  0.728 and 0.672 for standardised and individualised reconstructions, respectively) for grading 3/4 vs 0/1/2 and excellent agreement for overall impressions of whether studies were normal/pathological ( $\kappa = 0.884$ ). This is particularly relevant as patients on treatment or in research trials are very likely to have repeat scans and the efficacy of routine and experimental drugs is partly assessed by the results of the brain FDG PET-CT scans.

One of the limitations of our study is that we assessed inter- and intra-observer reproducibility without clinical and radiological integration of the findings. It would have been instructive to evaluate the clinical consequences of differences in interpretation of the FDG brain studies in the clinical context, e.g. in differentiating different types of dementia or assess how different results would influence patient management.

One previous study by Hoffman et al. [9] represents the only previous attempt at systematically evaluating inter- and intra-observer variability in the context of brain 18-FDG PET in isolation [9]. This study differed in its methodology in that only individualised reconstructions were compared, whilst the studies were interpreted for (i) bilateral temporoparietal hypometabolism (a finding considered a hallmark of Alzheimer's dementia [4, 13]) and (ii) uptake within temporal, parietal and frontal regions bilaterally. This study described 'excellent intra-observer and inter-observer interpretation agreement' for assessment of hypometabolism globally across bilateral temporoparietal regions with Kappa agreement scores of 0.56 and 0.51, respectively. These levels of  $\kappa$  values are similar to the ones we observed across brain regions when identifying moderate and severe uptake vs normal or mild hypometabolism. However, when considering isolated regional quantitative uptake specifically with regards to an Alzheimer's pattern of disease, we found inter-observer agreement to be somewhat less consistent, given that  $\kappa$  values for the medial temporal lobe—a region most notably implicated in Alzheimer's type neurodegenerative disease on PET and MRI studies [14, 15]—demonstrated slight inter-observer disagreement ( $\kappa = -0.136$  for the

individualised reconstruction and 0.031 for the standardised reconstruction).

At present, visual scan assessment performed by clinical experts such as that performed in our study remains the definitive reference for interpretation of brain 18-FDG PET-CT in routine clinical practice [16]. Literature to date, though primarily derived from dementia studies using brain 18-FDG PET in isolation, has demonstrated good correspondence with clinical outcomes as well as definitive histopathological diagnosis when relying on such visual scan assessments [16–22]. Several groups have attempted to enhance the accuracy and objectivity of brain 18-FDG PET through development of largely automated processing and assessment techniques to quantitatively evaluate scans. As described in the review by Herholz et al. [23], these predominantly automated processes incorporate several common components such as intensity scaling to reference brain regions with means and standard deviations of tracer uptake provided from normal reference samples, providing automated spatial normalisation to a template, as well as quantitative processes to summarise the extent of deviation from reference normal uptake values [23–26]. The recent multi-centre study by Caroli et al. [27] also demonstrated that these objective-automated analysis techniques were capable of identifying patients with MCI that subsequently developed Alzheimer's during follow-up with high sensitivities and specificities of between 70 and 90 %.

Our study emphasises the point that visual interpretation of brain 18-FDG PET-CT in isolation is susceptible to subjective variation, especially for the quantification of regional uptake, but is a robust technique demonstrating moderate inter-observer agreement and substantial intra-observer agreement. We found higher levels of inter-observer agreement when using standardised reconstructions and when using a high vs low grading (normal/near normal vs abnormal, i.e. grading  $\frac{3}{4}$  vs 0/1/2). Furthermore, we found such levels of agreement whilst observers were blind to any clinical information. In routine clinical practice, working closely with clinicians and within multidisciplinary teams will improve the quality of brain FDG PET-CT reports and correlation with clinical-radiological findings is essential for patient management. Advanced training in interpretation of brain 18-FDG PET-CT with regular audit is also very important to ensure high standards of reproducible reporting.

## Conclusion

Our study demonstrated moderate inter-observer and substantial intra-observer reproducibility in reporting brain 18-FDG PET-CT. Automated analysis techniques in routine clinical practice, advanced training and regular audits

of practice are likely to improve overall patient management and their exact role needs to be defined. Interestingly, a strong agreement in final diagnosis was observed in the centre. Such results suggest that the visual analysis approach is different between nuclear physicians but leads to the same final diagnosis. Another interesting way of looking at the data would have been to describe the patterns of uptake to fit pathologies, i.e. Alzheimer's disease or non-Alzheimer's disease. This could be done in a future study looking into the agreement of observers on the final diagnosis.

**Acknowledgments** The authors would like to acknowledge A Hitzel, A Julian and P Payoux, Department of Nuclear Medicine, Toulouse University Hospital.

**Conflict of interest** No potential conflicts of interest were disclosed.

**Informed consent** All human studies have been approved by the appropriate ethics committee and have, therefore, been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and all subsequent revisions.

## References

1. Wimo A, Prince M. World Alzheimer report 2010: the global economic impact of Dementia. Alzheimer's Disease International. Available online: <http://www.alz.co.uk/research/files/WorldAlzheimerReport2010.pdf> (2010). Accessed 27 Dec 2013.
2. Plassman BL, Langa KM, Fisher GG, Heeringa SG, Weir DR, Ofstedal MB, et al. Prevalence of cognitive impairment without dementia in the United States. *Ann Intern Med*. 2008;148:427–34.
3. Salomone S, Caraci F, Leggio GM, Fedotova J, Drago F. New pharmacological strategies for treatment of Alzheimer's disease: focus on disease modifying drugs. *Br J Clin Pharmacol*. 2012;73:504–17.
4. Herholz K, Carter SF, Jones M. Positron emission tomography imaging in dementia. *Br J Radiol*. 2007;80:S160–7.
5. Foster ML, Heidebrink JL, Clark CM, Jagust WJ, Arnold SE, Barbas NR, et al. FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease. *Brain*. 2007;130:2616–35.
6. Higdon R, Foster NL, Koeppe RA, DeCarli CS, Jagust WJ, Clark CM, et al. A comparison of classification methods for differentiating fronto-temporal dementia from Alzheimer's disease using FDG-PET imaging. *Stat Med*. 2004;30:315–26.
7. Shaffer JL, Petrella JR, Sheldon FC, Choudhury KR, Calhoun VD, Coleman RE, et al. Predicting cognitive decline in subjects at risk for Alzheimer disease by using combined cerebrospinal fluid, MR imaging, and PET biomarkers. *Radiology*. 2013;266:583–91.
8. Prestia A, Caroli A, Van Der Flier WM, Ossenkoppele R, Van Berckel B, Barkhof F, et al. Prediction of dementia in MCI patients based on core diagnostic markers for Alzheimer disease. *Neurology*. 2013;80:1048–56.
9. Hoffman JM, Hanson MW, Welsh KA, Earl N, Paine S, Delong D, et al. Interpretation variability of 18FDG-positron emission tomography studies in dementia. *Invest Radiol*. 1996;31:316–22.
10. Vellas B, Carrie I, Gillette-Guyonnet S, Touchon J, Dantoin T, Dartigues JF, et al. MAPT study: a multidomain approach for

- preventing Alzheimer's disease: design and baseline data. *J Prev Alzheimers Dis.* 2014;1:13–22.
11. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci.* 2001;56:M146–56.
  12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–74.
  13. Mosconi L. Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease. FDG-PET studies in MCI and AD. *Eur J Nucl Med Mol Imaging.* 2005;32:486–510.
  14. Ishkii K, Kitagaki H, Kono M, Mori E. Decreased medial temporal lobe oxygen metabolism in Alzheimer's disease shown by PET. *J Nucl Med.* 1996;37:1159–65.
  15. Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermeersch P, et al. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *J Neurol Neurosurg Psychiatry.* 1992;55:967–72.
  16. Bohnen NI, Djang DS, Herholz K, Anzai Y, Minoshima S. Effectiveness and safety of 18F-FDG PET in the evaluation of dementia: a review of the recent literature. *J Nucl Med.* 2012;53:59–71.
  17. Panegyres PK, Rogers JM, McCarthy M, Campbell A, Wu JS. Fluorodeoxyglucose-positron emission tomography in the differential diagnosis of early onset dementia: a prospective, community based study. *BMC Neurol.* 2009;9:41–9.
  18. Döbert N, Pantel J, Frölich L, Hamscho N, Menzel C, Grünwald F. Diagnostic value of FDG-PET and HMPAO-SPET in patients with mild dementia and mild cognitive impairment: metabolic index and perfusion index. *Dement Geriatr Cogn Disord.* 2005;20:63–70.
  19. Silverman DH, Small GW, Chang CY, Lu CS, Kung De Aburto MA, Chen W, et al. Positron emission tomography in evaluation of dementia: regional brain metabolism and long-term outcome. *JAMA.* 2001;286:2120–7.
  20. Foster ML, Heidebrink JL, Clark CM, Jagust WJ, Arnold SE, Barbas NR, et al. FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease. *Brain.* 2007;130:2616–35.
  21. Jagust W, Reed B, Mungas D, Ellis W, Decarli C. What does fluorodeoxyglucose PET imaging add to a clinical diagnosis of dementia? *Neurology.* 2007;69:871–7.
  22. Hyman BT, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Carrillo MC, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement.* 2012;8:1–13.
  23. Herholz K. Use of FDG PET as an imaging biomarker in clinical trials of Alzheimer's disease. *Biomark Med.* 2012;6:431–9.
  24. Herholz K, Salmon E, Perani D, Baron JC, Holthoff V, Frölich L, et al. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *Neuroimage.* 2002;17:302–16.
  25. Chen K, Ayutyanont N, Langbaum JB, Fleisher AS, Reschke C, Lee W, et al. Characterising Alzheimer's disease using hypometabolic convergence index. *Neuroimage.* 2011;56:52–60.
  26. Kakimoto A, Kamekawa Y, Ito S, Yoshikawa E, Okada H, Nishizawa S, et al. New computer-aided diagnosis of dementia using positron emission tomography: brain regional sensitivity-mapping method. *PLoS One.* 2011;6(9):e25033.
  27. Caroli A, Prestia A, Chen A, Ayutyanont N, Landau SM, Madison CM, et al. Summary metrics to assess Alzheimer disease-related hypometabolic pattern with 18F-FDG PET: head-to-head comparison. *J Nucl Med.* 2012;53:592–600.