Summary Metrics to Assess Alzheimer Disease–Related Hypometabolic Pattern with ^{18}F-FDG PET: Head-to-Head Comparison

Anna Caroli[1,2], **Annapaola Prestia**[1], **Kewei Chen**[3], **Napatkamon Ayutyanont**[3], **Susan M. Landau**[4], **Cindee M. Madison**[4], **Cathleen Haense**[5], **Karl Herholz**[6], **Flavio Nobili**[7], **Eric M. Reiman**[3], **William J. Jagust**[4], and **Giovanni B. Frisoni**[1], and **the EADC-PET Consortium, NEST-DD, and Alzheimer's Disease Neuroimaging Initiative**

[1]LENITEM Laboratory of Epidemiology and Neuroimaging, IRCCS S. Giovanni di Dio-FBF, Brescia, Italy

[2]Medical Imaging Unit, Biomedical Engineering Department, Mario Negri Institute for Pharmacological Research, Bergamo, Italy

[3]Banner Alzheimer's Institute, Phoenix, Arizona

[4]Helen Wills Neuroscience Institute, University of California, Berkeley, California

[5]Hannover Medical School, Clinic for Nuclear Medicine, Hannover, Germany

[6]University of Manchester, United Kingdom

[7]Division of Clinical Neurophysiology, Department of Neurosciences, Ophthalmology and Genetics, University of Genoa, Genoa, Italy

Abstract

In the recently revised diagnostic criteria for Alzheimer disease (AD), the National Institute on Aging and Alzheimer Association suggested that confidence in diagnosing dementia due to AD and mild cognitive impairment (MCI) due to AD could be improved by the use of certain biomarkers, such as ^{18}F-FDG PET evidence of hypometabolism in AD-affected brain regions. Three groups have developed automated data analysis techniques to characterize the AD-related pattern of hypometabolism in a single measurement. In this study, we sought to directly compare the ability of these three ^{18}F-FDG PET data analysis techniques—the PMOD Alzheimer discrimination analysis tool, the hypometabolic convergence index, and a set of meta-analytically derived regions of interest reflecting AD hypometabolism pattern (metaROI)—to distinguish moderate or mild AD dementia patients and MCI patients who subsequently converted to AD dementia from cognitively normal older adults.

Methods—One hundred sixty-six ^{18}F-FDG PET patients from the AD Neuroimaging Initiative, 308 from the Network for Efficiency and Standardization of Dementia Diagnosis, and 176 from the European Alzheimer Disease Consortium PET study were categorized, with masking of group classification, as AD, MCI, or healthy control. For each AD-related ^{18}F-FDG PET index, receiver-operating-characteristic curves were used to characterize and compare subject group classifications.

For correspondence contact: Giovanni B. Frisoni, Laboratory of Epidemiology and Neuroimaging, IRCCS San Giovanni di Dio-FBF, via Pilastroni 4, 25125 Brescia, Italy. gfrisoni@fatebenefratelli.it.

Results—The 3 techniques were roughly comparable in their ability to distinguish each of the clinical groups from cognitively normal older adults with high sensitivity and specificity. Accuracy of classification (in terms of area under the curve) in each clinical group varied more as a function of data-set than by technique. All techniques were differentially sensitive to disease severity, with the classification accuracy for MCI due to AD to moderate AD varying from 0.800 to 0.949 (PMOD Alzheimer tool), from 0.774 to 0.967 (metaROI), and from 0.801 to 0.983 (hypometabolic convergence index).

Conclusion—The 3 tested techniques have the potential to help detect AD in research and clinical settings. Additional efforts are needed to clarify their ability to address particular scientific and clinical questions. Their incremental diagnostic value over other imaging and biologic markers makes them easier to implement by other groups for these purposes.

Keywords

The availability of in vivo biomarkers of Alzheimer disease (AD) neuropathology has recently led to the development of new criteria (1–5) that reconceptualize AD as a disease featuring the combination of brain amyloidosis and neurodegeneration. To effectively translate the revised diagnostic criteria into clinically validated criteria for AD diagnosis, reliable metrics, among the several developed so far, need to be selected, compared, and standardized.

Hippocampal atrophy and cerebrospinal fluid biomarkers have been shown to be valid indicators of AD pathology (6,7), and standardization efforts are ongoing (8,9). The third diagnostic marker proposed in the revised diagnostic criteria is cortical temporoparietal hypometabolism on ^{18}F-FDG PET. Temporoparietal hypometabolism has been shown for many years to be a valid indicator of synaptic dysfunction that accompanies neurodegeneration in AD (10–12) and can be used as a diagnostic marker from the earliest stages of disease (13,14).

In the last few years, several global indices of AD-related hypometabolism, based on different image processing procedures with different levels of complexity, robustness, and automation, have been developed: the so-called PMOD (PMOD Technologies) Alzheimer discrimination analysis tool (PALZ) (11,15), an AD-related hypometabolic convergence index (HCI) (16), and an average metabolism computed on a set of metaanalytically derived regions of interest reflecting an AD hypometabolism pattern (metaROI) (17). All of these indices provide objective measures of metabolic damage specific to AD and are based on voxel-by-voxel analysis of ^{18}F-FDG PET images. PALZ and HCI computation involves the comparison of individual ^{18}F-FDG PET images to a reference database of scans of cognitively normal elderly individuals through the voxelwise t test. The PALZ score is computed as a voxel-by-voxel sum of t scores in a predefined AD-pattern mask, whereas HCI is calculated as the inner product of the individual z map and a predefined AD z map. PALZ and HCI have been shown able to separate patients with clinical AD from healthy older persons (15,16). HCI has been shown able to predict the development of dementia in patients with mild cognitive impairment (MCI), whereas metaROI has been shown sensitive in the detection of longitudinal cognitive and functional changes in AD and MCI patients (17). However, to our knowledge, the diagnostic performances of the 3 global indices have never been directly compared.

The aim of this study was to perform a head-to-head comparison of the diagnostic performance of available global indices of AD-specific hypometabolism at different disease stages. This is a necessary step toward the standardization of ^{18}F-FDG PET biomarkers,

opening the way to effective translation of the revised diagnostic criteria of the National Institute on Aging and Alzheimer Association into daily clinical practice.

MATERIALS AND METHODS

Subjects

The subjects and patients were taken from 3 independent ^{18}F-FDG PET multicenter datasets: AD Neuroimaging Initiative (ADNI) ((18) $n = 166$), the Network for Efficiency and Standardization of Dementia Diagnosis (NEST-DD) ((11,19) $n = 308$), and the European Alzheimer Disease Consortium PET Study (EADC-PET) ((20) $n = 176$). A detailed description of the datasets can be found in section 1.1 of the supplemental data (supplemental materials are available online only at http://jnm.snmjournals.org).

Control Dataset—The control dataset, used in receiver-operating-curve (ROC) analyses to assess the diagnostic performance of ^{18}F-FDG PET global indices on different patient datasets, was chosen so as to be independent from data used to define any metric involved in the study. This is a mandatory factor to avoid circularity. We included controls from NEST-DD (all those not previously used to define the PALZ metric (11,15), $n = 35$) and all controls from the EADC-PET database with available and high-quality baseline ^{18}F-FDG PET data ($n = 113$).

Test Datasets—Test datasets, used to assess and compare the diagnostic performance of different global indices, were all the available independent samples of AD patients taken from the NEST-DD, EADC-PET, and ADNI databases with available baseline ^{18}F-FDG PET data, namely ADNI patients with mild to moderate AD ($n = 96$) and MCI due to AD ($n = 73$), NEST-DD patients with mild to moderate AD ($n = 242$) (15) and MCI due to AD ($n = 31$), and EADC-PET patients with MCI due to AD ($n = 63$). MCI due to AD was defined as a diagnosis of MCI at baseline followed by conversion to AD during follow-up; MCI patients who converted to non-AD dementia were excluded from the study.

To further investigate diagnostic performance at different disease stages, we subdivided AD patients from the NEST-DD and ADNI datasets into mild and moderate (NEST-DD: 82 patients with mild AD and 160 with moderate AD; ADNI: 53 patients with mild AD and 43 with moderate AD) on the basis of the Mini Mental State Examination (MMSE) score (mild AD: MMSE 24; moderate AD: MMSE < 24).

Groups of patients with MCI due to AD were disaggregated into fast and slow converters on the basis of conversion time (fast converters: conversion time 12 mo; slow converters: conversion time > 12 mo) to assess the ability of each index to predict short-term or long-term clinical progression from MCI to probable AD (NEST-DD: 13 fast converters and 18 slow converters; ADNI: 24 fast converters and 49 slow converters; EADC-PET: 17 fast converters and 46 slow converters).

^{18}F-FDG PET Summary Metrics

^{18}F-FDG PET global indices of AD-related hypometabolism used in the study were PALZ (11), HCI (16), and metaROI average (17) (the latter converted to W-scores according to previously published procedures (21,22)). Each global index was computed for all baseline ^{18}F-FDG PET images in the development and test datasets.

All metrics are based on voxel-by-voxel analysis of ^{18}F-FDG PET images and provide a single AD-related hypometabolism measure; however, they are computed using different

processing procedures. A complete description is provided in section 1.3 of the supplemental data.

PALZ score computation requires the commercially available PMOD software, whereas both HCI and metaROI require the Statistical Parametric Mapping (SPM) software package (http://www.fil.ion.ucl.ac.uk/spm/), running on a Matlab environment (The MathWorks, Inc.). In addition, computation of metaROI W-scores requires 5 previously defined metaROIs (left angular, right angular, left temporal, right temporal, and bilateral posterior cingulate binary masks in Montreal Neurological Institute space) (17), which are available on the ADNI Web site (http://adni.loni.ucla.edu/) and a pons and cerebellar vermis ROI in Montreal Neurological Institute space used for scaling. HCI computation requires a predefined normative database, a z score map of AD-related cerebral hypometabolism, and the HCI software package, written in Matlab.

As PMOD supports different image formats (Digital Imaging and Communications in Medicine [DICOM], ECAT, Analyze, Neuroimaging Informatics Technology Initiative [NIFTI], Interfile, and others), the PALZ score can be computed from ^{18}F-FDG PET images in any format without the need to convert them. However, the use of different formats gives rise to slight differences in PALZ scores (slight differences already appearing in images after normalization), in particular between Analyze and NIFTI/DICOM formats. DICOM format is preferred, as it contains patient-, acquisition-, and reconstruction-related descriptive information, allowing the display of images with consistent orientation (left/right, anterior/posterior, inferior/superior) and the automatic extraction of patient's age, needed for calculation; if any other format is used, images must be checked for correct display orientation and eventually reoriented to anatomic position, preliminary to PALZ score computation. On the other hand, both HCI and metaROI, based on SPM subroutines, require Analyze or NIFTI format, and ^{18}F-FDG PET images in other formats need to be converted to either of them. For HCI and metaROI computation, Analyze and NIFTI formats are equivalent, as the use of either leads to the same result. A preliminary orientation check is required for Analyze. For both formats, ^{18}F-FDG PET reorientation to anatomic position (usually needed, as ^{18}F-FDG PET images are acquired in radiologic convention) is automatically performed during SPM normalization.

Once the ^{18}F-FDG PET image is imported into PMOD, the PALZ score computation is fully automated and takes about 2 min per subject. Although PALZ computation could be performed in batch mode, any benefit provided by batch processing is limited because a visual check of image normalization is advisable for each case. Once all subjects of the normative dataset and all patients have been normalized to the default PET template space, and a text file listing patient normalized images has been created, HCI computation is automatically performed by the HCI package. HCI computation requires about 5 min per subject and can be performed in a batch. MetaROI computation, requiring the application of different SPM subroutines (normalization, metaROI average computation, scaling to pons and cerebellar vermis, average computation, and final age correction) is the most time-consuming (about 15 min per subject) and the least automated procedure. MetaROI computation can be performed in a batch. Computational times were estimated using a Core 2 Duo processor T5450 personal computer (Intel) (1.66-GHz, 667-MHz front-side bus, 2-MB level 2 cache, Windows XP operating system [Microsoft]).

The technical requirements, robustness, and automation of different ^{18}F-FDG PET global indices are summarized in Table 1.

ROC Analyses

For each global index of AD-like hypometabolism (PALZ, HCI, and metaROI) and for each test dataset (ADNI AD, ADNI MCI due to AD, NEST-DD AD, NEST-DD MCI due to AD, and EADC-PET MCI due to AD) and subgroup of patients (mild and moderate AD, fast and slow converters) separately, ROC curves were generated together with the control dataset. Areas under the curve (AUCs) and pertinent 95% confidence intervals were computed. To compare the diagnostic performance of different global indices of AD-like hypometabolism for each group and subgroup of patients, AUCs related to different indices were compared using the test of De Long et al. for 2 correlated ROC curves (23), setting the threshold for significance at a *P* value of 0.05.

All statistical analyses were performed using R software (www.r-project.org/), version 2.12.1. ROC analyses were performed using the pROC R package (24).

RESULTS

Normative and Patient Dataset Characterization

The normative datasets included 148 healthy controls from the NEST-DD and EADC-PET datasets (mean age $\pm$ SD, 66 $\pm$ 7 y; 69 men and 79 women). The NEST-DD controls were enrolled in 4 clinical centers (Cologne: $n = 19$, 62 $\pm$ 4 y old, 63% women; Liege: $n = 12$, 64 $\pm$ 5 y old, 58% women; Florence: $n = 2$, 60 $\pm$ 8 y old, no women; Dresden: $n = 2$, 57 $\pm$ 11 y old, 50% women), and the EADC-PET controls were enrolled in 5 different centers (Brescia: $n = 27$, 65 $\pm$ 5 y old, 52% women; Genoa: $n = 36$, 69 $\pm$ 7 y old, 67% women; Marseilles: $n = 10$, 66 $\pm$ 5 y old, 40% women; Munich: $n = 19$, 68 $\pm$ 8 y old, 47% women; Amsterdam: $n = 21$, 66 $\pm$ 7 y old, 38% women). Table 2 shows the main sociodemographic and clinical features and the ^{18}F-FDG PET global indices of controls included in the normative dataset, disaggregated by dataset and enrollment center. The healthy subjects who were enrolled in the different clinical centers had comparable clinical features and ^{18}F-FDG PET global indices, suggesting homogeneity of the multicenter normative dataset.

Ninety-five AD patients from ADNI (76 $\pm$ 7 y old, 41% women) and 242 AD patients from NEST-DD (71 $\pm$ 8 y old, 65% women), subdivided into mild and moderate AD, were included in the study. Seventy-one patients with MCI due to AD from ADNI (75 $\pm$ 7 y old, 39% women), 31 from NEST-DD (71 $\pm$ 6 y old, 61% women), and 63 from EADC-PET (71 $\pm$ 9 y old, 56% women), subdivided into fast and slow converters, were further included. Table 3 shows the main sociodemographic and clinical features and the ^{18}F-FDG PET global indices of each group and subgroup of AD and MCI due to AD. As expected, in both the ADNI and the NEST-DD datasets, AD patients at increasing disease stage (MCI due to AD, mild AD, and moderate AD) showed decreasing MMSE scores and increasing AD-like hypometabolism according to all 3 global metrics. Within patients with MCI due to AD, fast and slow converters showed similar MMSE scores but no univocal trends in AD-like hypometabolism.

^{18}F-FDG PET Summary Metrics in Comparison: Diagnostic Performance

Figures 1–5 compare the diagnostic performance of ^{18}F-FDG PET global indices of AD-like hypometabolism in terms of ROC curves for each group and subgroup of AD (Figs. 1 and 2) and MCI due to AD (Figs. 3–5). Pertinent AUCs are summarized and statistically compared in Supplemental Table 1.

In all ADNI test datasets, HCI showed the highest AUCs; HCI AUC was significantly higher than metaROI and PALZ both in the whole AD group (*P* < 0.005 and *P* < 0.001, respectively) and in the mild and moderate AD subgroups (mild AD: *P* < 0.005 and *P* <

0.005, respectively; moderate AD: $P < 0.005$ and $P < 0.01$, respectively). HCI AUC was significantly higher than PALZ AUC in the whole group of MCI due to AD ($P < 0.05$) and in the fast-converters subgroup ($P < 0.05$). In all NEST-DD test datasets, metaROI showed the highest AUCs; metaROI AUC was significantly higher than PALZ in the whole AD group and in the mild AD subgroup ($P < 0.05$), whereas in the moderate AD subgroup, AUCs pertinent to all global metrics were notably high, with no significant differences among them. MetaROI AUC was significantly higher than both the PALZ AUC and the HCI AUC in the whole group of MCI due to AD ($P < 0.01$ and $P < 0.05$, respectively) and significantly higher than PALZ AUC in the slow-converters subgroup ($P < 0.05$). In EADC-PET, the test dataset for MCI due to AD and the PALZ subgroups showed the highest AUCs; PALZ AUCs were significantly higher than metaROI AUCs both in the whole group of MCI due to AD ($P < 0.005$) and in the fast- and slow-converters subgroups ($P < 0.05$), and PALZ AUCs were significantly higher than HCI in the whole group of MCI due to AD ($P < 0.05$) and in the fast-converters subgroup ($P < 0.05$).

For all indices, AUC increased in AD groups at increasing disease stage (being highest for moderate AD and lowest for MCI due to AD), varying from 0.800 to 0.949 (PALZ), from 0.774 to 0.967 (metaROI), and from 0.801 to 0.983 (HCI). Differences in AUCs between mild and moderate AD were much larger in the NEST-DD than in the ADNI dataset. Diagnostic performance in subgroups with MCI due to AD was not consistent across different datasets, being either higher in fast converters than slow converters (ADNI: all indices; EADC-PET: PALZ) or vice versa (NEST-DD: all indices; EADC-PET: HCI and metaROI).

DISCUSSION

In the current study, we considered three ^{18}F-FDG PET global indices (PALZ, HCI, and metaROI) providing objective measures of AD-related hypometabolism, and we compared them both in technical terms and in terms of diagnostic performance on several independent groups of patients at different stages of AD, taken from the 3 largest ^{18}F-FDG PET datasets currently available (ADNI, NEST-DD, and EADC-PET).

Global metrics show differences in complexity, technical requirements, and automation level. Their diagnostic performance considerably changed according to test dataset and disease stage, pointing out that no global index can be defined as the best-performing. For all indices, diagnostic performance improved with increasing disease severity, whereas in MCI due to AD (fast and slow converters), diagnostic performance was not consistent across different datasets.

In the literature, there are few reports of the diagnostic performance of ^{18}F-FDG PET global metrics in AD patients at different disease stages. PALZ performance was recently assessed in both ADNI and NEST-DD AD patient groups (15): despite using different normative data-sets to assess specificity (either ADNI or NEST-DD, according to AD patient group), the authors found similar ROC curves and AUCs; to our knowledge, PALZ diagnostic performance in MCI due to AD has never been studied. HCI performance was previously assessed in terms of its ability to distinguish between AD patients, MCI patients who converted to AD, stable MCI patients, and controls and to predict rates of progression from MCI to probable AD (16). Because in the current paper we used a modified version of HCI, current findings could not reliably be compared with previous ones. Finally, metaROI index performance was previously assessed in terms of sensitivity to detect longitudinal change in both cognitive and functional measurements within AD and MCI (17); to our knowledge, metaROI diagnostic performance in AD patients at different disease stages has never been studied.

All three ^{18}F-FDG PET global metrics under comparison were developed specifically for the discrimination between AD patients and controls. ^{18}F-FDG PET metrics of AD-like hypometabolism could be used neither for differential diagnosis among various forms of dementia (which could, however, show abnormal scores on any of them) nor for highlighting vascular damage (which should be assessed using different techniques). Thus, patients with dementing diseases other than AD were not considered for the current investigation.

The diagnostic performance of ^{18}F-FDG PET indices was assessed in patients with AD at different stages (ranging from MCI due to AD to moderate AD), whereas MCI patients who did not convert to AD during follow-up were not considered. Although it would have been interesting to compare the ability of ^{18}F-FDG PET indices to identify patients who will never convert to dementia (true-negatives), given that the minimum observation time required to ensure no conversion is 5–6 y (25) we could not exclude the possibility that patients who had not converted during the follow-up time (much shorter than 5 y for most available MCI patients) would have converted in the future, and we would thus have had unreliable results.

The control dataset used in ROC analyses to assess the diagnostic performance of ^{18}F-FDG PET global indices on different AD patient datasets included controls from the NEST-DD and EADC-PET databases. Despite the many strengths of ADNI, in that study the healthy subjects may not be fully representative of the healthy population, as they have been shown (although in quite a small sample) to have a high rate of Pittsburgh compound B positivity (26), probably due to the recruitment modality. On the other hand, achieving a representative normative database is quite difficult, independently of selection modality. Despite the fact that controls from the EADC-PET and NEST-DD datasets have shown homogeneous sociodemographic, clinical, and metabolic features across different enrollment centers, one should be skeptical about the representativeness of the healthy elderly population. The use of the same normative dataset to assess the diagnostic performance of all ^{18}F-FDG PET global metrics under comparison on each test dataset improved the reliability of head-to-head comparisons. Furthermore, the independence of the normative dataset from all datasets used to develop and optimize different metrics made it possible to avoid any circularity, which could have biased the comparison.

Because each algorithm handled age differently, the age of the controls could be a potential confounding factor. Age correction embedded in PALZ and metaROI computation enabled the removal of any variance due to age. As the current implementation of HCI does not take age into account but significant linear dependence was found in the normative dataset, further work should be done to investigate the effect of age on HCI and to properly correct for such an effect under all possible diagnostic conditions.

Some limitations should be considered in the interpretation of the present results. First, as visual rating by expert physicians still represents the gold standard clinical method of assessing AD-like hypometabolism on ^{18}F-FDG PET, the diagnostic performance and accuracy of global metrics should be preliminarily compared with visual rating by independent expert raters. Second, the 3 global metrics included in this head-to-head comparison are not the only automated methods to assess AD-like hypometabolism on ^{18}F-FDG PET images; the 3 metrics should be further compared with other available voxel-based techniques, such as single-case SPM (27) or 3-dimensional stereotactic surface projection and NEUROSTAT-based indices (28). Third, patients with MCI due to AD were disaggregated into fast and slow converters on the basis of conversion time since enrollment; however, because the time of symptom onset is unknown (fast converters could be enrolled after having symptoms for a long time), caution should be used when considering such

subgroupings. Finally, considerations about the user friendliness of the three ^{18}F-FDG PET summary metrics are based on their current implementation; however, they were all implemented for academic use. Additional programming can make them more automated and user-friendly for clinical settings.

CONCLUSION

The current study showed that the 3 tested AD-related ^{18}F-FDG PET global metrics have the potential to help detect AD in research and clinical settings. As different metrics have different technical requirements and levels of automation, the choice among them should be driven by available resources (software and technical skills). Furthermore, as the head-to-head comparison in terms of diagnostic performance revealed that no ^{18}F-FDG PET global index can be defined as the best-performing, the choice among them should rather be based on the specific purpose of use. Additional efforts are needed to clarify the ability of ^{18}F-FDG PET global metrics to address particular scientific and clinical questions (e.g., differential diagnosis of dementia, prediction of subsequent decline over different time points or prediction of neuropathology, reduction of the number of patients needed for a clinical trial using clinical or biomarker endpoints), to determine their incremental diagnostic value over other imaging and biologic markers (e.g., hippocampal atrophy or amyloid load), and to make them easier to implement by other groups for these purposes. The current study is a first step toward several future directions. The potential increase in diagnostic accuracy of the combination of ^{18}F-FDG PET with structural or anatomic imaging and biochemical bio-markers could be investigated, in view of the effective translation of the revised diagnostic criteria of the National Institute on Aging and Alzheimer Association into daily clinical practice. Moreover, ^{18}F-FDG PET global metrics of non–AD-like (e.g., frontotemporal dementia–like) hypometabolism could be designed and developed to help in early and differential diagnosis of non-AD dementias.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Dubois B, Feldman HH, Jacova C, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 2007; 6:734–746. [PubMed: 17616482]

2. Dubois B, Feldman HH, Jacova C, et al. Revising the definition of Alzheimer's disease: a new lexicon. Lancet Neurol. 2010; 9:1118–1127. [PubMed: 20934914]

3. Albert MS, Dekosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011; 7:270–279. [PubMed: 21514249]

4. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011; 7:280–292. [PubMed: 21514248]

5. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011; 7:263–269. [PubMed: 21514250]

6. Frisoni GB, Fox NC, Jack CR Jr, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol. 2010; 6:67–77. [PubMed: 20139996]

7. Blennow K, Hampel H, Weiner M, Zetterberg H. Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. Nat Rev Neurol. 2010; 6:131–144. [PubMed: 20157306]

8. Frisoni GB, Jack CR. Harmonization of magnetic resonance-based manual hippocampal segmentation: a mandatory step for wide clinical use. Alzheimers Dement. 2011; 7:171–174. [PubMed: 21414554]

9. Verwey NA, van der Flier WM, Blennow K, et al. A worldwide multicentre comparison of assays for cerebrospinal fluid biomarkers in Alzheimer's disease. Ann Clin Biochem. 2009; 46:235–240. [PubMed: 19342441]

10. Minoshima S, Giordani B, Berent S, Frey KA, Foster NL, Kuhl DE. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. Ann Neurol. 1997; 42:85–94. [PubMed: 9225689]

11. Herholz K, Salmon E, Perani D, et al. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. Neuroimage. 2002; 17:302–316. [PubMed: 12482085]

12. Jagust W, Reed B, Mungas D, Ellis W, Decarli C. What does fluorodeoxyglucose PET imaging add to a clinical diagnosis of dementia? Neurology. 2007; 69:871–877. [PubMed: 17724289]

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. [PubMed: 15747152]

14. Anchisi D, Borroni B, Franceschi M, et al. Heterogeneity of brain glucose metabolism in mild cognitive impairment and clinical progression to Alzheimer disease. Arch Neurol. 2005; 62:1728–1733. [PubMed: 16286547]

15. Haense C, Herholz K, Jagust WJ, Heiss WD. Performance of FDG-PET for detection of Alzheimer's disease in two independent multicentre samples (NEST-DD and ADNI). Dement Geriatr Cogn Disord. 2009; 28:259–266. [PubMed: 19786778]

16. Chen K, Ayutyanont N, Langbaum JBS, et al. Characterizing Alzheimer's disease using a hypometabolic convergence index. Neuroimage. 2011; 56:52–60. [PubMed: 21276856]

17. Landau SM, Harvey D, Madison CM, et al. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. Neurobiol Aging. 2011; 32:1207–1218. [PubMed: 19660834]

18. Jagust WJ, Bandy D, Chen K, et al. The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core. Alzheimers Dement. 2010; 6:221–229. [PubMed: 20451870]

19. Herholz K. PET studies in dementia. Ann Nucl Med. 2003; 17:79–89. [PubMed: 12790355]

20. Morbelli S, Drzezga A, Perneczky R, et al. Resting-state brain metabolic connectivity in amnestic MCI-AD converters and healthy controls: a joint project of the European Alzheimer's Disease Consortium (EADC). Eur J Nucl Med Mol Imaging. 2010; 37(suppl 2):S259–S260.

21. Boccardi M, Laakso MP, Bresciani L, et al. The MRI pattern of frontal and temporal brain atrophy in fronto-temporal dementia. Neurobiol Aging. 2003; 24:95–103. [PubMed: 12493555]

22. Jack CR Jr, Petersen RC, Xu YC, et al. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology. 1997; 49:786–794. [PubMed: 9305341]

23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988; 44:837–845. [PubMed: 3203132]

24. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011; 12:77. [PubMed: 21414208]

25. Petersen RC, Doody R, Kurz A, et al. Current concepts in mild cognitive impairment. Arch Neurol. 2001; 58:1985–1992. [PubMed: 11735772]

26. Jagust WJ, Landau SM, Shaw LM, et al. Relationships between biomarkers in aging and dementia. Neurology. 2009; 73:1193–1199. [PubMed: 19822868]

27. Signorini M, Paulesu E, Friston K, et al. Rapid assessment of regional cerebral metabolic abnormalities in single subjects with quantitative and nonquantitative [^{18}F]FDG PET: a clinical validation of statistical parametric mapping. Neuroimage. 1999; 9:63–80. [PubMed: 9918728]

28. Minoshima S, Frey KA, Koeppe RA, et al. A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. J Nucl Med. 1995; 36:1238–1248. [PubMed: 7790950]
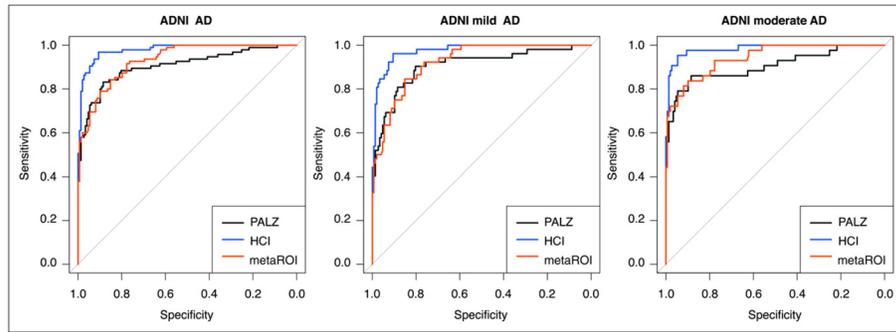
FIGURE 1.
ROC curves showing diagnostic performance of ^{18}F-FDG PET global indices of AD-like hypometabolism (black: PALZ; blue: HCI; red: metaROI) on different groups of AD patients from ADNI.
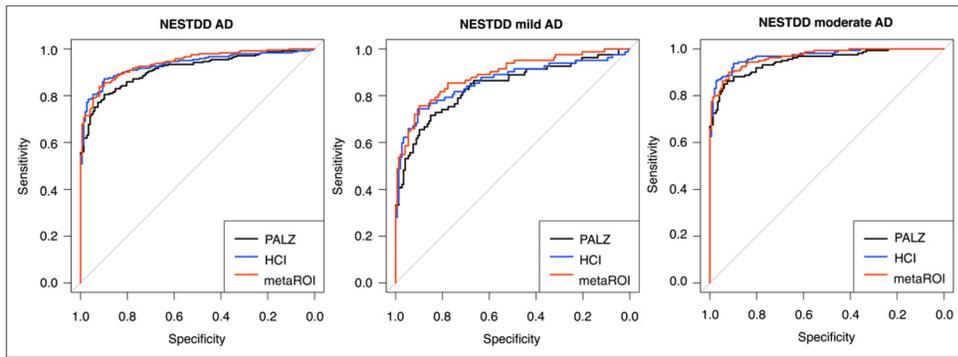
FIGURE 2.
ROC curves showing diagnostic performance of ^{18}F-FDG PET global indices of AD-like hypometabolism (black: PALZ; blue: HCI; red: metaROI) on different groups of AD patients from NEST-DD dataset.
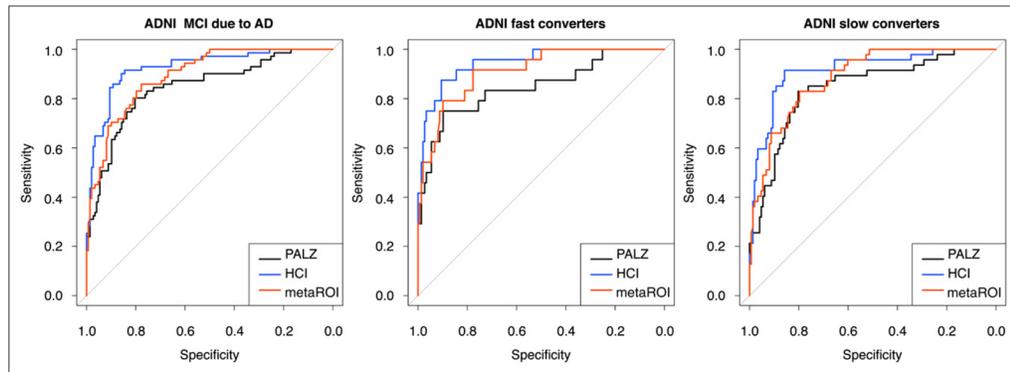
FIGURE 3.
ROC curves showing diagnostic performance of ^{18}F-FDG PET global indices of AD-like hypometabolism (black: PALZ; blue: HCI; red: metaROI) on different groups of MCI due to AD patients from ADNI dataset.
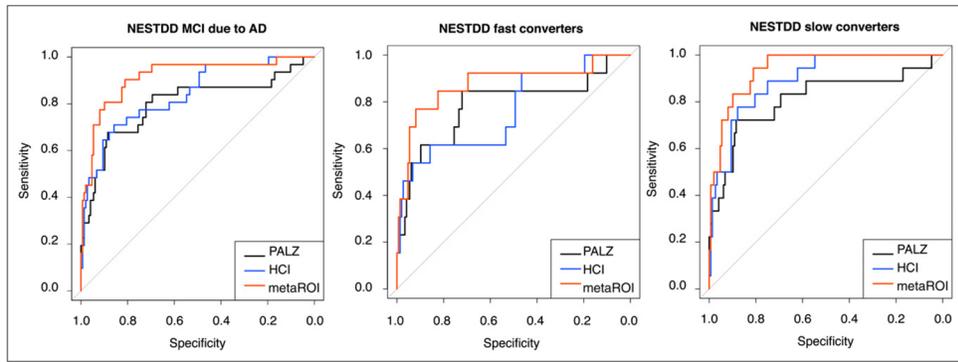
FIGURE 4.
ROC curves showing diagnostic performance of ^{18}F-FDG PET global indices of AD-like hypometabolism (black: PALZ; blue: HCI; red: metaROI) on different groups of MCI due to AD patients from NEST-DD dataset.
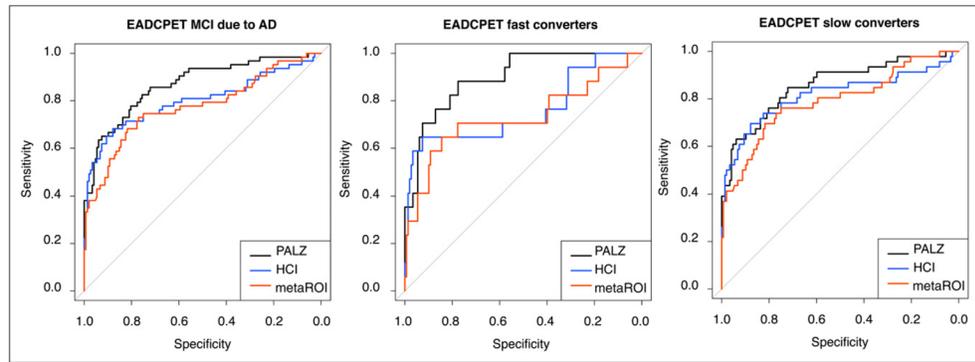
FIGURE 5.
ROC curves showing diagnostic performance of ^{18}F-FDG PET global indices of AD-like hypometabolism (black: PALZ; blue: HCI; red: metaROI) on different groups of MCI due to AD patients from EADC-PET dataset.

TABLE 1

Technical Requirements, Automation, and Robustness of Different ^{18}F-FDG PET Global Indices Involved in This Head-to-Head Comparison

Characteristic	PALZ	metaROI	HCI
Software required	PMOD	Matlab, SPM	Matlab, SPM, HCI package
Additional requirements	—	5 metaROIs, pons/cerebellar vermis ROI	Normative dataset, AD z score map
Automation	High	Medium	High
Supported image formats	DICOM, ECAT, Analyze, NIFTI, and others	Analyze, NIFTI	Analyze, NIFTI
Robustness	Slight change with image format; preliminary reorientation to anatomic position needed (formats other than DICOM)	No change with image format; automatic reorientation to anatomic position	No change with image format, automatic reorientation to anatomic position

SPM = statistical parametric mapping.

TABLE 2

Sociodemographic and Clinical Features, ^{18}F-FDG PET Global Indices of Healthy Subjects Included in the Normative Dataset, and Disaggregated by Dataset and Enrollment Center

Dataset	Center	n	Age (y)	Sex (female)	MMSE	PALZ (t scores)	metaROI (W-scores)	HCI[*]
NEST-DD	Cologne	19	62 ± 4	12 (63%)	29 ± 1	4,432 ± 2,147	0.55 ± 1.32	−253 ± 504
	Liege	12	64 ± 5	7 (58%)	29 ± 1	3,916 ± 6,683	0.22 ± 1.28	−159 ± 450
	Florence	2	60 ± 8	0 (0%)	29 ± 1	7,168 ± 8,480	1.67 ± 1.71	330 ± 899
	Dresden	2	57 ± 11	1 (50%)	29 ± 1	5,258 ± 2,340	−0.53 ± 1.18	−229 ± 157
EADC-PET	Brescia	27	65 ± 5	14 (52%)	29 ± 1	7,368 ± 6,488	0.11 ± 1.89	120 ± 601
	Genoa	36	69 ± 7	24 (67%)	29 ± 1	6,821 ± 3,974	0.24 ± 1.98	246 ± 639
	Marseilles	10	66 ± 5	4 (40%)	30 ± 0	3,160 ± 1,930	−0.54 ± 1.67	−73 ± 432
	Munich	19	68 ± 8	9 (47%)	29 ± 1	3,464 ± 1,608	−0.36 ± 1.32	−265 ± 646
	Amsterdam	21	66 ± 7	8 (38%)	30 ± 1	3,982 ± 3,098	−0.70 ± 1.88	−2 ± 626

[*] HCI, computed as inner product of individual z map and predefined AD z map on all voxels, is expressed in arbitrary units.

Values are mean ± SD for continuous variables, or frequency (followed by percentage in parentheses) for sex.

TABLE 3

Sociodemographic and Clinical Features and ^{18}F-FDG PET Global Indices of Controls, Test Datasets (AD and MCI Due to AD) and Subgroups of Patients (Mild and Moderate AD, Fast and Slow Converters)

Dataset	Subgroup	n	Age (y)	Sex (female)	MMSE	PALZ (t scores)	metaROI (W-scores)	HCI[*]
Controls		148	66 ± 7	79 (53%)	29 ± 1	5,271 ± 4,489	0.0 ± 1.73	−1 ± 607
ADNI AD	All	95	76 ± 7	39 (41%)	23 ± 2	35,954 ± 32,904	3.89 ± 2.12	2,352 ± 1,190
	Mild	52	75 ± 7	21 (40%)	25 ± 1	32,146 ± 31,968	3.51 ± 1.93	2,103 ± 1,018
	Moderate	43	76 ± 7	18 (42%)	21 ± 1	40,559 ± 33,801	4.33 ± 2.28	2,653 ± 1,320
ADNI MCI due to AD	All	71	75 ± 7	28 (39%)	27 ± 2	18,345 ± 16,368	2.86 ± 1.59	1,490 ± 898
	Fast converter	24	75 ± 6	10 (42%)	27 ± 2	20,170 ± 14,988	3.15 ± 1.65	1,791 ± 995
	Slow converter	47	74 ± 7	18 (38%)	27 ± 2	17,413 ± 17,110	2.72 ± 1.55	1,336 ± 813
NEST-DD AD	All	242	71 ± 8	158 (65%)	21 ± 4	39,572 ± 35,069	4.40 ± 2.27	2,145 ± 1,355
	Mild	82	70 ± 9	50 (61%)	25 ± 1	23,686 ± 22,458	3.25 ± 2.04	1,425 ± 1,185
	Moderate	160	71 ± 8	108 (67%)	19 ± 4	47,615 ± 37,535	4.98 ± 2.16	2,513 ± 1,290
NEST-DD MCI due to AD	All	31	71 ± 6	19 (61%)	25 ± 2	17,893 ± 21,838	3.11 ± 1.65	1,093 ± 1,002
	Fast converter	13	71 ± 6	9 (69%)	24 ± 2	20,704 ± 30,862	2.98 ± 2.21	1,036 ± 1,369
	Slow converter	18	71 ± 6	10 (56%)	26 ± 2	15,862 ± 12,604	3.20 ± 1.16	1,133 ± 668
EADC-PET MCI due to AD	All	63	71 ± 9	35 (56%)	27 ± 2	25,213 ± 24,638	2.32 ± 2.51	1,263 ± 1,275
	Fast converter	17	73 ± 10	11 (65%)	27 ± 3	20,764 ± 16,711	1.64 ± 2.30	1,043 ± 1,096
	Slow converter	46	71 ± 9	24 (52%)	27 ± 1	26,857 ± 26,959	2.57 ± 2.56	1,345 ± 1,337

[*] HCI, computed as inner product of individual z map and predefined AD z map on all voxels, is expressed in arbitrary units.

Values are mean ± SD for continuous variables, or frequency (followed by percentage in parentheses) for sex.