European Journal of Nuclear Medicine and Molecular Imaging
Validation of FDG-PET datasets of normal controls for the extraction of SPM-based brain metabolism maps
--Manuscript Draft--

Abstract:	Purpose: An appropriate healthy control dataset is mandatory to achieve good performance in voxel-wise analyses. We aimed at evaluating [18F]FDG PET brain datasets of healthy controls (HC), based on publicly available data, for the extraction of voxel-based brain metabolism maps at the single-subject level. Methods: Selection of HC images was based on visual rating, after Cook's distance and Jack-knife analyses, to exclude artefacts and/or outliers. The performance of these HC datasets (ADNI-HC and AIMN-HC) to extract hypometabolism patterns in single patients was tested in comparison with the standard reference HC dataset (HSR-HC) by means of Dice score analysis. We evaluated the performance and comparability of the different HC datasets in the assessment of single-subject SPM-based hypometabolism in three independent cohorts of patients, namely ADD, bvFTD and DLB. Results: Two-steps Cook's distance analysis and the subsequent Jack-knife analysis resulted in the selection of n=125 subjects from the AIMN-HC dataset and n=75 subjects from the ADNI-HC dataset. The average concordance between SPM hypometabolism t-maps in the three patient cohorts, as obtained with the new datasets and compared to the HSR-HC standard reference dataset, was 0.87 for the AIMN-HC

	dataset and 0.83 for the ADNI-HC dataset. Pattern expression analysis revealed high overall accuracy (>80%) of the SPM t-map classification according to different statistical thresholds and sample sizes. Conclusions: The applied procedures ensure validity of these HC datasets for the single-subject estimation of brain metabolism using voxel-wise comparisons. These well-selected HC datasets are ready-to-use in research and clinical settings.
Response to Reviewers:	Details about the raised technical aspects were provided in the manuscript and highlighted. A careful English editing was also performed, in accordance with the Referee's suggestion. We thank the Editor and the two Referees for helping us to improve our manuscript.

To Prof. Arturo Chiti, MD

Editor-in-Chief of *European Journal of Nuclear Medicine and Molecular Imaging*

Milan, 17 December 2020

Dear Editor,

Details about the raised technical aspects were provided in the manuscript and highlighted.

A careful English editing was also performed, in accordance with the Referee's suggestion.

We thank the Editor and the two Referees for helping us to improve our manuscript.

All authors have read and approved the final version of the manuscript. The authors certify that the manuscript represents valid work and that neither this manuscript nor one with substantially similar content has been published under their authorship or is being considered for publication elsewhere, and that there is no ghostwriting by anyone not named on the author list. The authors report no disclosures relevant to the manuscript.

On behalf of all co-authors,

Daniela Perani

Prof. Daniela Perani, MD
Vita-Salute San Raffaele University
Division of Neuroscience San Raffaele Scientific Institute
Neurology Nuclear Medicine Unit, San Raffaele Hospital
Via Olgettina 60, Milan 20132, Italy

Validation of FDG-PET datasets of normal controls for the extraction

of SPM-based brain metabolism maps

Silvia Paola Caminiti[1,2,¥], Arianna Sala[1,2,¥], Luca Presotto[3,¥],

Andrea Chincarini[4], Stelvio Sestini[5], Daniela Perani[1,2,3,#],

for the Alzheimer's Disease Neuroimaging Initiative (ADNI) * and

for the Associazione Italiana Medicina Nucleare (AIMN) datasets

1) Vita-Salute San Raffaele University, Milan, Italy
2) In vivo human molecular and structural neuroimaging Unit, Division of Neuroscience, IRCCS San Raffaele Scientific Institute, Milan, Italy
3) Nuclear Medicine Unit, San Raffaele Hospital, Milan, Italy
4) Istituto Nazionale di Fisica Nucleare, Genova, Italy
5) Nuovo Ospedale di Prato (NOP) S. Stefano, Prato, Italy

The AIMN Neurology Study-Group collaborators: Orazio Schillaci and Valentina Berti, Maria Lucia Calcagni, Angelina Cistaro, Silvia Morbelli, Flavio Nobili, Sabina Pappatà, Duccio Volterrani, Clara Luigia Gobbo.

[¥] Authors equally contributing to the manuscript.

*Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

[#] Correspondence:

Professor Daniela Perani, MD

Via Olgettina 60, 20132 Milan, Italy

phone: +39-02-26432224 or 26432223

email: perani.daniela@hsr.it

Abstract

Purpose: An appropriate healthy control dataset is mandatory to achieve good performance in voxel-wise analyses. We aimed at evaluating [18F]FDG PET brain datasets of healthy controls (HC), based on publicly available data, for the extraction of voxel-based brain metabolism maps at the single-subject level.

Methods: Selection of HC images was based on visual rating, after Cook's distance and Jack-knife analyses, to exclude artefacts and/or outliers. The performance of these HC datasets (ADNI-HC and AIMN-HC) to extract hypometabolism patterns in single patients was tested in comparison with the standard reference HC dataset (HSR-HC) by means of Dice score analysis. We evaluated the performance and comparability of the different HC datasets in the assessment of single-subject SPM-based hypometabolism in three independent cohorts of patients, namely ADD, bvFTD and DLB.

Results: Two-steps Cook's distance analysis and the subsequent Jack-knife analysis resulted in the selection of n=125 subjects from the AIMN-HC dataset and n=75 subjects from the ADNI-HC dataset. The average concordance between SPM hypometabolism t-maps in the three patient cohorts, as obtained with the new datasets and compared to the HSR-HC standard reference dataset, was 0.87 for the AIMN-HC dataset and 0.83 for the ADNI-HC dataset. Pattern expression analysis revealed high overall accuracy (>80%) of the SPM t-map classification according to different statistical thresholds and sample sizes.

Conclusions: The applied procedures ensure validity of these HC datasets for the single-subject estimation of brain metabolism using voxel-wise comparisons. These well-selected HC datasets are ready-to-use in research and clinical settings.

Keywords

Fluorodeoxyglucose; PET; brain hypometabolism; healthy control dataset; voxel-wise analysis; SPM; dementia; neurodegeneration.

Introduction

[18F]Fluorodeoxyglucose-Positron Emission Tomography ([18F]FDG PET) is the most widely used tool for the *in vivo* measurement of brain metabolism, a crucial biomarker in dementia as a proxy for synaptic dysfunction and neurodegeneration [1]. Disease-specific hypometabolism patterns, obtained with [18F]FDG PET imaging, provide support to the early and differential diagnosis of several neurodegenerative conditions [2,3].

Metrics for [18F]FDG PET imaging hold substantial value to estimate the extent of brain metabolic derangement in neurodegenerative conditions (ND), providing distinctive metabolic patterns that increase diagnostic accuracy, even in the prodromal disease phase [2,4]. Multiple studies have reported indeed high sensitivity and specificity in studies adopting appropriate metrics for [18F]FDG PET;. overall, good quality of evidence exists in early and differential diagnosis of ND dementia. Crucially, highly accurate, reliable, and reproducible measures of neurodegeneration are essential for inclusion in therapeutic trials and for treatment evaluation and monitoring.

While the use of [18F]FDG PET is recommended by several clinical/research diagnostic criteria [5–11], in 2015, a Cochrane review questioned the diagnostic and prognostic accuracy of [18F]FDG PET in early prodromal phases of dementia [12]. The imminent reply of the European Association of Nuclear Medicine (EANM) [13] objected such conclusion, on the grounds that it was mostly based on studies with evaluation of [18F]FDG PET images limited to the visual inspection of radiotracer distribution, thus neglecting the objective measures as a major bias. The implementation of standardized [18F]FDG PET readouts and operator-independent maps were strongly advocated by both the United States Society of Nuclear Medicine and the Molecular Imaging and the European Association of Nuclear Medicine [13,14].

Different advanced parametric tools, each including a HC database for statistical comparison, have been introduced so far. These tools can be subdivided into those that were developed for commercial/clinical use, which are the majority, e.g. CortexID (GE), Neurocloud-PET (Qubiotech Health Intelligence) (www.neurocloud.es), MIMNeuro (MIM Software), NeuroQ™ (Syntermed), Syngo.PET (Siemens) and Brass (Hermes Medical Solutions, PNEURO/PALZ (PMOD Technologies), and those that were developed for research purposes, e.g. NeuroSTAT and PETQuant (Cortechs Lab). In this context, indices providing a single measurement of the expression of hypometabolism patterns, such as the Hypometabolic Convergence Index, or the ADNI metaROI, are also available [15,16]. Of note, even when objective measurements are applied, based on the available tools, results may vary depending on the tool characteristics. One must also be aware of the lack of validation and standardization, for example with respect to the selected intensity normalization methods, or to differences amongst scanners and centers in the acquisition procedures [17]. In general, head-to-head comparisons are still required to harmonize quantification procedures; so far, the complexity of the current quantification tools, especially those developed for research uses, represents a limitation for the routine use by nuclear physicians in clinical settings [18].

Crucially, the selection criteria applied for the inclusion of normal subjects within each software are often not described in detail, nor published in the literature. The selection of the healthy controls (HC), i.e. classified as cognitively healthy at the time of the scan, in these tools, is often based on the availability of [18F]FDG PET only, and does not take into account eventual patterns of neurodegeneration at the [18F]FDG PET scan, nor the presence of brain pathology at CT or MRI, and crucially the cognitive and clinical trajectories of each included subject. As an example, in the Hypometabolic Convergence Index, 47 healthy controls from the ADNI database were used for comparison, of which four had converted to MCI after the baseline scan, one at month 6 (who turned to Alzheimer's disease dementia at month 36) and three at month 24 [15]. An appropriate healthy control dataset is mandatory to obtain a good performance in voxel-wise analyses. Criteria for clinical normality are usually based on the absence of cognitive impairments, neurological or psychiatric diseases and the use of structured interviews. Furthermore, for better quality results, HC selected for comparisons should be followed up longitudinally to confirm long lasting cognitive stability. It has been shown that using a control group of subjects followed longitudinally for 4 years (instead of mixed databases) strongly increases the accuracy (1.4- to 2-fold) of disease detection in an automated [18F]FDG PET analysis [19].

In this work, we aimed to provide a validation of two [18F]FDG PET databases of publicly available, clinically well-characterized samples of HC to the research and clinical communities. These normative datasets can be used by applying several statistical approaches, such as t-test in the SPM toolbox or others.

Currently, SPM does not provide a normal dataset to perform brain [18F]FDG PET voxel-wise statistical analyses. Here, we showed an application of SPM procedures that are freely downloadable from the web. Notably, while the earliest versions needed a MATLAB® platform, the most recent standalone version does not. Our group has previously validated this semi-quantitative optimized method based on SPM [20,21]. It includes a spatial normalization based on a [18F]FDG PET dementia-specific template (freely available at http://inlab.ibfm.cnr.it/inlab/PET_template.php) [22], that improves the detection of subtle metabolic abnormalities associated with specific cognitive impairment and the recognition of different patterns of neurodegeneration – particularly in early disease stages. The procedure is based on a large healthy subjects dataset of [18F]FDG PET images for statistical comparison (n=112) belonging to a San Raffaele Hospital internal dataset [20,22]. This method showed very high accuracy for differential diagnosis of ND dementia, atypical parkinsonian disorders, and crucially for prognosis in the risk assessment of progression in the prodromal phase of mild cognitive impairment [2,3]. Here, after implementing an analytical pipeline to build up the appropriate HC datasets, we investigated the robustness of the SPM t-maps obtained with the two new well-selected HC datasets, both based on publicly available data provided by the Italian national association of nuclear medicine (AIMN) and from the Alzheimer's Disease Neuroimaging Initiative (ADNI), respectively. By using these datasets, SPM hypometabolism t-maps were extracted from three different clinical groups, namely patients affected by Alzheimer's disease dementia (ADD), dementia with Lewy body (DLB) and the behavioural variant of frontotemporal dementia (bvFTD). Notably, we also investigated the impact of the size of the control dataset on the resulting single-subject SPM t-maps, at different statistical thresholds, so as to provide guidance to other centres.

4

We would like to promote the use of semi-quantitative objective measures of brain metabolism by means of statistical comparisons with these normative datasets for an accurate assessments of brain metabolism in research and clinical settings.

Materials and Methods

HC datasets

We considered two HC [18F]FDG PET datasets, one obtained from the AIMN website (aimn.it) and one from the ADNI website (adni.loni.usc.edu).

The AIMN is a voluntary non-profit association, whose purpose is to promote the application and development of the medical and biological use of the physical properties of the atomic nucleus. The AIMN represents the Italian scientific reference for Nuclear Medicine and Molecular Imaging activities. Qualified members of AIMN can submit an online request to obtain [18F]FDG PET images from HC subjects acquired by different Italian Nuclear Medicine Units . We downloaded n=187 [18F]FDG PET scans, i.e. all the images available on the AIMN database (https://www.aimn.it/site/page/gds/gds-5). This database comprised subjects aged between 20 and 84 years (mean: 61.88±13.69). These subjects were selected because they were characterized by absence of global cognitive impairment, as assessed by a MMSE score ≥ 28, and were cognitively normal after an average 4-year clinical follow-up.

The ADNI is an American public-private partnership launched in 2003 led by Principal Investigator Michael W. Weiner, MD to collect longitudinal data on MCI and ADD patients as well as on healthy subjects. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early ADD (see www.adni-info.org for more information). All data generated by the ADNI study investigators are entered in the data repository hosted at the Laboratory of Neuroimaging (LONI) at the University of Southern California, the LONI Image & Data Archive (IDA). Qualified researchers worldwide can submit an online data access request and begin using ADNI data including cognitive/neuropsychological, image, biofluid and genetic data sets within a few days from request submission. The ADNI data repository contains more than 500 healthy control subjects acquired with [18F]FDG PET. Among them, we selected n=101 subjects (age range: 56-94 years; mean: 73.49±6.97). These subjects were selected because they were characterized by absence of global cognitive impairment, as assessed by a MMSE score ≥ 28, and were cognitively normal after an average 4-year clinical follow-up. Moreover, subjects showing amyloid PET positivity were excluded from the analysis. Amyloid PET positivity was established based on quantitative assessment of global signal uptake value ratio (SUVr), as provided by ADNI. Details on such procedures are described in the ADNI analytical pipelines (http://adni.loni.usc.edu/).

Presence of significant amyloid-β burden (amyloid-positivity) was established based on quantitative assessment of global SUVr in either [11C]PiB-PET or [18F]Florbetapir-PET scans, as provided by ADNI.

Presence of significant amyloid-β burden (amyloid-positivity) was established based on a validated cut-off of 1.5 SUVr for [11C]PiB-PET scans and of 1.11 SUVr for [18F]Florbetapir-PET scans, as derived from large studies in healthy controls. Further details are described in the ADNI analytical pipelines (http://adni.loni.usc.edu/).

Overall, these selection procedures allowed us to exclude, with higherconfidence, HC subjects who might be on a trajectory towards mild cognitive impairment or dementia.

HSR-HC reference dataset – A group of 112 healthy control [18F]FDG PET scans was here included as reference standard. The HSR-HC dataset was previously developed and validated by our group for the extraction of hypometabolism maps in individual neurological patients, or groups (see section "*SPM t-maps extraction in dementia conditions*"). This represents the internal HC database at San Raffaele Hospital (HSR) Nuclear Medicine Unit, which is largely used for research purposes and in support to clinical routine [20,21,23–27].

Patient groups datasets

From a large clinical cohort referred to the Departments of Neurology and Nuclear Medicine Unit of San Raffaele Hospital (Milan, Italy), we retrospectively selected the patients who underwent [18F]FDG PET scans and received a diagnosis of probable dementia, namely 79 cases diagnosed with probable ADD [11], 81 with probable DLB [5], and 56 fulfilling current clinical criteria for bvFTD [10]. All patients had a confirmed neurological diagnosis after at least 5-year follow-up. These groups of patients were included in order to compare the results of the SPM procedure using the different HC datasets for the extraction of single-subject hypometabolism maps.

All subjects and/or authorized representatives provided written informed consent, following detailed explanation of each experimental procedure. ADNI subjects gave written informed consent at the time of enrolment for data collection and completed questionnaires approved by each participating sites Institutional Review Board. The protocols conformed to the ethical standards of the Declaration of Helsinki for protection of human subjects.

[18F]FDG PET Scanning Procedures and Pre-processing

AIMN and HSR scanning procedures- An [18F]FDG PET brain study was performed in all subjects (healthy controls and patients) according to the conventional neurological acquisition protocols and to the European Association of Nuclear Medicine (EANM) guidelines [28]. Before radiopharmaceutical injection, subjects were fasted for at least six hours to ensure that the measured blood glucose level was <120 mg/dL. Subjects underwent a 3D PET scan (time interval between injection and scan start ranged from 30 to 45 min; scan duration ranged from 10 to 15 min depending on the PET scanner characteristics) after the injection of

[18F]FDG (185–250 MBq: usually, 5–8 mCi via a venous cannula). Images were reconstructed using an ordered subset expectation maximization algorithm.

[18F]FDG PET images belonging to AIMN subjects were obtained using either a PET/CT system GE Discovery STE, GE Discovery 710, or Siemens Biograph 16 scanner. [18F]FDG PET images belonging to HSR subjects were obtained using a either a GE Discovery ST, GE Discovery STE, Siemens Biograph Hi-rez or Siemens ECAT EXACT HR+ PET scanner.

ADNI PET Scanning procedures - A list of PET and PET/CT scanners used in ADNI centres is available elsewhere (http://adni.loni.usc.edu). ADNI acquisition procedures are comparable to those described above, with acquisitions starting 30 minutes post injection, and six 5-minute frame being acquired for the next 30 minutes. As for the [18F]FDG PET images included, only the last three 5-minute frames were retrieved from the ADNI database and combined to obtain a single 15-min static image. In such way, we ensured uniform acquisition procedures for all [18F]FDG PET images, independently of the acquisition site. Assuming the image reconstruction algorithm is unbiased, summing dynamic frames gives identical results as performing a static acquisition of the same time length. It also allows to correct for inter-frame motion. Even in case of statistical reconstruction algorithms, small bias has been reported exclusively in "cold structures" and for very short frames (less than 10 seconds) [29].

ADNI acquisition procedures are detailed elsewhere (http://adni.loni.usc.edu/data-samples/data-types/pet/). [18F]FDG-PET images belonging to ADNI subjects were obtained using ECAT HR+, General Electric Discovery LS, General Electric Discovery ST, General Electric Discovery STE, Siemens/ECAT HRRT, Siemens Biograph Hi-Rez, Philiphs Gemini TF, Siemens mCT, ECAT Biograph.

Pre-processing – A visual quality check of the images was performed to identify potential artefacts (e.g. acquisition issues, excessive patient motion) and issues related to technical characteristics, such as the use of compatible reconstruction algorithms. Then, [18F]FDG PET images of patients and controls were normalized to the optimized [18F]FDG PET template [21], using SPM12 (https://www.fil.ion.ucl.ac.uk/spm/). They were then scaled to the global mean of the activity within the brain [30] and finally smoothed with an isotropic 3D Gaussian kernel (8 mm FWHM), accordingly with the validated pipeline proposed for our single-subject SPM-based analysis [20,22]. This kind of smoothing is required for the random field theory to be applicable. It is also effective in reducing the number of multiple comparisons to be performed [31].

HC datasets processing

Study design – The validation of the HC groups was performed following two main steps. In the first step, we performed a quantitative outliers detection by means of *"Cook's distance analysis"* [22]. In the second step, we optimized outlier detection by classifying SPM t-maps extracted by means of a jack-knife approach, where every normalized HC scan was evaluated with respect to the remaining sample of HC (AIMN-HC or ADNI-HC) via a two-sample t-test in SPM12. The resulting SPM t-maps were visually inspected by two neuroimaging experts, in order to confirm or discard the outliers that had been quantitatively detected in the first step.

Cook's distance analysis - In order to detect outliers and to remove them from the two HC datasets (AIMN and ADNI), we measured the Cook's distance for each subject.

Cook's distance for subject *i* is defined according to the following formula:

$$D_i = \frac{\sum_j (\hat{y}_j - \hat{y}_{j(i)})^2}{ps^2}$$

Where $\hat{y}_i$ is the predicted model response for subject j, and $\hat{y_i}_{(j)}$ the estimated response when the model is estimated without subject i, p the number of covariates and s the mean squared error of the regression model. In this case, $\hat{y}_i$ are images and their squared difference is computed as the Euclidian distance.

Thus, we obtained a Cook's distance value for each analysed subject. The Cook's distance values were averaged over all the voxels that are included in the SPM analysis, and then plotted in order to define the critical D value. We considered the "elbow" on the arm of the distribution, as representing the best critical D value. A Cook's distance below the critical D value is expected not to have any large impact on sample distribution. Thus, the subject can be considered representative of the group and can be included in the HC sample. Instead, a Cook's distance above the critical D value indicates that the subject has a disproportionate impact on the estimated general linear model.

We computed Cook's distance twice for each HC group (AIMN or ADNI HC groups): first, in the whole HC group, in order to exclude the most egregious outliers; second, we re-computed Cook's distance by using only the subjects that survived to the first step. We obtained two new HC groups, which were than further analysed in the following step.

Jack-knife approach – The SPM statistical voxel-wise procedure consists in a t-test, in which a single individual is compared with a dataset of HC, entering age as a covariate. This statistical comparison provides t-scores for each brain voxel [20]. In this case, every normalized [18F]FDG PET images was evaluated with respect to the remaining sample via a two-sample t-test so that a SPM t-map is obtained for each HC subject (jack-knife approach). SPM t-maps were generated from this statistical comparison in order to identify eventual brain areas of significant hypometabolism ($p<0.05$). Only clusters containing more than 100 voxels were considered to be significant [20]. Two neuroimaging experts visually inspected all the SPM t-maps in order to confirm the scan negativity, namely the absence of hypometabolism patterns compatible with neurodegenerative processes.

All the above analyses allowed to obtain HC groups without outliers.

Validation of HC datasets for t-maps extraction in dementia conditions

SPM single-subject standardized procedure - We tested for the validity of the two new selected HC datasets (AIMN-HC and ADNI-HC) in extracting single-subject hypometabolism t-maps in three different group of patients, namely ADD, DLB and bvFTD. To do so, we ran our standardized SPM procedure [20] on the three group of patients three times, the first time by using our validated internal HSR-HC dataset as reference of

comparison, the second time with the AIMN-HC dataset and the third time with the ADNI-HC dataset. The p-value of the single-subject hypometabolism maps was set at p<0.05 uncorrected, with a cluster-forming threshold of n = 100 voxels [20]. We selected images regardless of the scanner used for acquisition, since, as shown in a previous study [32], the SPM single-subject procedure provides comparable results even when images are acquired with different scanners. First, a voxel-wise map of commonalities was computed in order to evaluate the degree of overlap between the single-subject hypometabolism pattern in each dementia group, using the two new HC groups (AIMN-HC and ADNI-HC). The commonality value of each voxel in these commonality maps was computed as the proportion of patients in which that specific voxel was found to be hypometabolic. The resulting commonality maps were first visually inspected to evaluate their consistency with the topography of known hypometabolism patterns reported in the literature. Second, the t-maps obtained with the HSR-HC dataset (HSR SPM t-maps), representing the reference standard, were directly compared to t-maps obtained with the AIMN-HC and ADNI-HC datasets. To do so, after the single subject SPM procedure was run for each patient against the three HC dataset (HSR-HC, AIMN-HC and ADNI-HC), we assessed concordance between SPM t-maps obtained with the two HC datasets by means of Dice analysis. A Dice score for two binary images A and B is defined as:

$$DICE = \frac{2*(A \cap B)}{A + B}$$

Where with $A \cap B$ we mean the number of elements present in the intersection between the two masks (i.e. voxels that are commonly hypometabolic in the two masks), and with $A + B$ we mean the sum of all the elements in A and B (i.e. voxels that are hypometabolic in mask A and mask B). Dice coefficient takes the value of 1 if A and B assume the same logical value in every voxel (high concordance), and a value of 0 if they always disagree (null concordance). Basically, Dice scores represent the amount of spatial overlap of the identified brain hypometabolic regions in each single patient. Resulting Dice coefficients were then averaged over all patients to provide a summary measure of concordance in each cohort, when using AIMN-HC and ADNI-HC datasets for comparison. Further, a voxel-wise Dice map was computed over all patients, to measure how often both analyses agreed in each single voxel.

We also performed Chi-squared tests to compare the accuracy values obtained by using the two different datasets.

Sample size and statistical thresholds effects – We also assessed the comparability of the results of the SPM standardized single-subject procedure for each HC dataset, by changing the sample size of the HC dataset and the statistical thresholds applied for the SPM t-map extraction. Thus, we performed again our SPM standardized single-subject procedure considering for comparison a randomly subsampled pool of HC from the AIMN database (n=10, n=30, n=75, n=125), and for ADNI database (n=10, n=30, n=75). In the same analysis, we assessed how the HC group sample sizes and the use of different statistical thresholds and corrections (i.e. p<0.05 with Family-wise error correction, p<0.01 and p<0.05 uncorrected for multiple comparisons) may influence the expression of the hypometabolism patterns in the three groups of patients

9

(ADD, DLB and bvFTD). We created three different binary masks, one representing the AD-like pattern, one representing the DLB-like pattern and one representing the bvFTD-like pattern. The masks were obtained by averaging the t-maps derived from SPM single-subject analysis for each clinical group and converting them into binary images. Then, we assessed the pattern expression for each patient by means of a voxel-wise comparison between binary masks and patient's SPM t-maps. The pattern expression (PE) was defined with the following formula:

$$PE = \frac{N_{SPEC}}{Nmask}$$

Where N_{SPEC} represents the number of voxels found to be hypometabolic within the prototypical mask and N_{mask} represents the number of voxels of the prototypical mask.

A rater-independent evaluation was obtained by classifying each pattern as AD, bvFTD or DLB-like according to the maximum level of similarity to the standardized binary mask prototypical for each condition (i.e. the highest PE). Then, the rater-independent evaluation (AD-like, DLB-like, bvFTD-like pattern) was compared with the clinical diagnosis at follow-up (ADD, DLB or bvFTD diagnosis) in order to assess the classification accuracy of the SPM maps for each clinical condition.

Results

Selection of HC datasets

The initial assessment excluded n=20 HC from the original n=187 subjects within the AIMN database. The excluded images had a resolution recovery algorithm for reconstruction incorporating correction of partial volume effect, which makes these images incomparable with the rest of the HC database [33]. Further, quantification with resolution recovery, on top of not being comparable with standard reconstruction, provides divergent results among vendors [34].

In the first step, the Cook's distance analysis identified n=23 outliers out of 167 subjects in the AIMN database and n=11 outliers out of 101 subjects in the ADNI database (D=0.0051 as critical value). The second-step analysis identified n=32 outliers out of 167 subjects in theAIMN database and n=12 outliers out of 101 subjects in theADNI database (setting D=0.0084 as critical value). Overall, the two steps voxel-wise quantitative outliers' detection analyses resulted in the selection of n=135 for the AIMN and n=89 for the ADNI datasets, representative of the same normal distribution (Figure 1).

Amongst the images included by means of Cook's analysis, Jack-knife analysis and neuroimaging experts subsequently excluded those images showing hypometabolism clusters that, even if of limited extension, were deemed as pathological. After this final analysis, n=125 HC subjects for AIMN and n=75 HC subjects for ADNI were evaluated as having normal [18F]FDG PET images. These subjects were included in the final HC datasets. Demographical characteristics of the two final HC datasets are reported in Table 1.

In order to provide the scientific community with the 18F]FDG PET selected images, the list of codes (RID) of the HC subjects selected from ADNI are provided in Supplementary Table 1.

Evaluation of HC datasets for t-maps extraction in dementia conditions

The SPM single-subject procedure obtained by using the selected AIMN-HC and ADNI-HC groups for comparison yielded patterns of brain hypometabolism in the three groups of patients consistent with those reported in the literature [6,11,27,35–37]. Figure 2 and 3 (left panel) show the commonalities of SPM t-maps for the three clinical groups, obtained by using the AIMN dataset and the ADNI dataset, respectively. Specifically, the commonalities of AD hypometabolism t-maps involved temporoparietal association cortices, the precuneus and posterior cingulate cortex, the DLB hypometabolism t-maps involved the lateral and medial occipital cortex, temporo-parietal and frontal cortex, bilaterally, whereas the bvFTD hypometabolism t-maps showed the involvement of the dorsolateral frontal cortex, the insula and temporal regions (see Fig. 2-3).

When we compared the degree of overlap between hypometabolism patterns obtained with the new HC datasets (AIMN and ADNI) and the reference standard HSR-HC dataset, we obtained an average concordance of 0.87 in the AIMN-HSR comparison and 0.83 in the ADNI-HSR comparison. In the ADD clinical group, the AIMN-HSR concordance was 0.88 and 0.83 for the ADNI-HSR comparisons. For the DLB clinical group, we obtained an average concordance of 0.86 in the AIMN-HSR comparison and of 0.82 in the ADNI-HSR comparison. For the bvFTD clinical group, we obtained an average concordance of 0.89 in the AIMN-HSR comparison and of 0.86 in the ADNI-HSR comparison. In Figure 2 and 3 (right panel), we show the voxel-wise maps of Dice scores, representing the agreement in resulting hypometabolism maps in the two analyses with different HC datasets. In the hallmark regions of hypometabolism, the agreement between SPM t-maps obtained with AIMN-HC and HSR-HC datasets was higher than 0.94 (ADD=0.96, DLB=0.95 and bvFTD=0.94) and equal to 0.93 in the ADNI-HSR comparison (ADD=0.93, DLB=0.94 and bvFTD=0.93). This indicates, at the voxel level, that the SPM statistical method obtained using different control databases produce hypometabolism SPM t-maps with very high levels of spatial concordance.

Pattern expression analysis revealed that the number of voxels found to be statistically significant within the disease typical group-level maps increased proportionally with the number of HC subjects included in the single-subject SPM comparison (Supplementary Table 2-4 and Figure 5). Of note, the most stringent statistical threshold, namely $p < 0.05$ with family-wise error correction, when using a small HC sample size, i.e. n=10 HC subjects, prevented us from obtaining any statistically significant result in all three groups of patients (PE=0). The highest pattern expression (PE=0.65) was achieved when n=125 subjects were used for comparison and $p < 0.05$ uncorrected was applied as statistical threshold (Supplementary Table 2-4).

An optimal accuracy (>80%) of the SPM t-map classification with respect to the clinical diagnosis was achieved when a large sample size (>30) was used, regardless of the statistical threshold applied (see Figure 4 and Tables 2-3 for details). Of note, the selection of more stringent statistical thresholds, i.e. FWE-correction,

allowed to obtain hypometabolism voxel in hallmark regions, such as the hypometabolism in parietal cortex for AD, but prevented a full-blown pattern expression. This translated into a lower accuracy in the discrimination between conditions that present some overlap in their typical hypometabolism patterns, such as ADD and DLB. Accordingly, using a high statistical threshold, the DLB-like pattern reached a classification accuracy of $\approx$60-70%.

The classification accuracy obtained with the two HC datasets did not show any statistically significant difference, namely performance of AIMN and ADNI HC datasets were equivalent ($\chi^2_{AD\ 0.05\ unc}$=2.07, $\chi^2_{DLB\ 0.05\ unc}$=6.8, $\chi^2_{BvFTD\ 0.05\ unc}$=1.8; $\chi^2_{AD\ 0.01\ unc}$=0.9 $\chi^2_{DLB\ 0.01\ unc}$=4.3 $\chi^2_{BvFTD\ 0.01\ unc}$=2.1 ; $\chi^2_{AD\ 0.05\ FWE}$=0.2 $\chi^2_{DLB\ 0.05\ FWE}$=4.7 $\chi^2_{BvFTD\ 0.05\ FWE}$=3.1)

Discussion

Patients with different dementias may show partially overlapping clinical presentations, which prevents a proper clinical classification. The last decades have progressively witnessed a shift from a purely clinical diagnosis to a biomarker-supported diagnosis, and molecular neuroimaging techniques such as PET have played a leading role in the dementia research diagnostic work-up [6,10,11,38–40]. The ability to discriminate between different neurodegenerative conditions, together with the ability to detect a disease process even before the occurrence of clinical manifestations, has huge implications for diagnosis and prognosis in imaging research, clinical trials and therapeutic approaches. In the case of cerebral glucose metabolism, a large amount of literature has provided clear evidence of specific, disease-related [18F]FDG PET patterns [20,27,35,36,41–43], which are significantly accurate and useful. Sensitive and specific [18F]FDG PET analysis tools for the detection of brain hypometabolism at single-subject level are crucial, particularly for the detection of specific, early brain metabolic changes. However, only methods employing parametric and voxel-based analysis techniques can provide unbiased, statistically defined measures of brain abnormality across the whole brain [44].

The SPM single-subject standardized procedure presented in this study, allows the identification of disease-specific brain hypometabolism patterns in single individuals with high statistical power [20]. Still, the ability to identify patterns of neurodegeneration with high sensitivity and specificity depends on specific operational standards. Among them, the use of internationally recognized control datasets (e.g. US ADNI, NEST-DD or the European Alzheimer's Disease PET Consortium—EADC-PET) that include individuals who were acquired with standardized procedure and are clinically characterized so as to exclude cognitive deterioration, is recommended [44].

In this study, we provided a step-by-step selection and validation of two healthy control [18F]FDG PET datasets that have been collected by two internationally recognized scientific associations (AIMN and ADNI). Standardized acquisition procedures for [18F]FDG PET imaging were applied with a high standard, with the purpose to increase the diagnostic impact of this technique in clinical practice, and to ensure consistency of

data collection for longitudinal studies. Of note, these images were acquired with different PET scanners; however, the application of standardized acquisition procedures followed by appropriate post-processing provided reliable results. In a previous study we demonstrated that the use of HC data acquired with different PET scanners shows no influence on the performance of standardized SPM procedures [45].

Here, by following several quality-control steps, we selected two subsets of HC subjects that provide comparable and very accurate results in the detection of brain metabolic patterns of neurodegeneration.

First, we selected the subjects characterized by absence of global cognitive impairment, and we also considered cognitive stability after a 4-year clinical follow-up. Since early abnormalities in [18F]FDG PET scans are likely to affect regions characterized by very early pathological changes [46], inclusion of subjects in preclinical phases in a cohort of HC would hamper statistical power, making detection of hypometabolism, especially in regions of crucial relevance for diagnostic and prognostic purposes, more biased. Moreover, for the ADNI-HC database, we excluded subjects showing cerebral amyloid retention at amyloid-PET.

It must be noted that, while information on amyloid pathology and clinical follow-up was available only in the ADNI-HC dataset, we put into place a stringent analytical pipeline in both the AIMN-HC and ADNI-HC datasets, so as to ensure selection of only those HC subjects, with a brain [18F]FDG PET scan that clearly excluded ongoing dysfunctional changes. All these procedures hence allowed us to select only HC subjects with a negative [18F]FDG PET scan. To this regard, it is worth noting that our rigorous analytical pipeline allowed us to select AIMN-HC and ADNI-HC datasets providing highly concordant results in the extraction of brain hypometabolism patterns as compared to the validated reference standard (close to 100% concordance in the most relevant hypometabolism hallmarks) and both yielding optimal (>80%) accuracy in discriminating among the three major dementing conditions.

Age-related decreases of brain metabolism could characterize HC subjects. However, hypometabolism in HC does not follow the topographical distribution of the hypometabolism patterns related to neurodegenerative conditions. Age-related hypometabolism is more distributed, it may involve medial frontal cortex, and anterior cingulate cortex, but not with the severity and the extension in posterior brain regions or the dorsolateral frontal cortex as observed in patients with ND [47]. In order to select only those HC subjects with "normal" brain metabolism, we performed a quantitative data-driven outlier's detection analysis, by means of Jack-Knife approach and Cook's distance analysis (Figure 1). The two quantitative procedures allowed us to exclude HC subjects with levels of brain metabolism deviating from the normal distribution of the overall sample.

Overall, the selection of HC subjects following different quality control steps allowed to obtain, in the comparison with three independent cohorts of patients, patterns of brain hypometabolism with high statistical power, yielding hypometabolism topographies that are clearly consistent with the literature (Figure 2-3 and Table 3-4). Accurate discrimination between dementing conditions was achieved both when considering the complete HC datasets and also when limited HC subsamples, selected in a randomized way, were extracted

for comparison. This result further confirms the high similarity and internal consistency of the HC [18F]FDG PET images selected through the two-step analytical pipeline.

It was previously reported that when random field theory is used in neuroimaging to estimate the number of multiple comparisons, as it is performed in SPM, the less degrees of freedom are used in the test, the more conservative this correction becomes [48]. This phenomenon is especially strong when less than 30 HC are present, as clearly shown in the present PE analysis (Supplementary Table 2 and Figure 3). This highlights the importance of using a proper number of HC for constructing a normal dataset. A high dataset numerosity ensures good statistical power to detect all the hypometabolic clusters of crucial diagnostic relevance, also when more conservative statistical thresholds are applied. A previous study showed that diagnostic performance of [18F]FDG PET in discriminating ADD patients from HC, as measured by receiver operating characteristic curve, was relatively poor when the sample size of the HC dataset was small (i.e. <10), possibly because of high statistical noise due to small sample size [49]. They affirmed that only large normal datasets (e.g. > 60) would yield lower statistical noise and more stable diagnostic performance. Nevertheless, the authors acknowledged that such a very large dataset is not easily available in many clinical settings. Accordingly, Gallivanone and colleagues showed that the use of a HC dataset >50 is critical when the [18F]FDG uptake of brain regions crucial for diagnosis of AD, is unclear at visual inspection of the tracer radioactivity distribution [50].

Of note, our data also highlighted that the use of a FWE strategy should be carefully considered, when the number of healthy controls available for comparison is less than 10. Note that the proposed standardized procedure using large and well-selected HC datasets, allows to obtain reliable patterns of brain hypometabolism in single individuals also when less conservative statistical thresholds are applied. The use of less conservative thresholds (e.g. 0.05 uncorrected) and/or larger HC sample size (>30) provides the maximum pattern expression (PE=0.65) without producing false negative values, i.e. voxel being significantly hypometabolic but not belonging to the prototypical hypometabolism pattern (Figure 4).

The extraction of reliable and informative hypometabolism pattern is particularly relevant in supporting differential diagnosis in conditions characterized by partially overlapping hypometabolism patterns, such as ADD and DLB. A full-blown pattern expression allows to obtain more information, that may be crucial to obtain a correct diagnosis for the clinician. In this regard, in a previous study, we demonstrated that the presence of an occipital lobe hypometabolism yields great ability to discriminate between DLB and ADD [36]. The use of a standardized approach, allowing the extraction of early brain dysfunctional changes at the single-subject level, predating a full-blow expression of dementia, improves diagnosis in prodromal disease phases, such as in mild cognitive impairment, avoiding multiple clinical and instrumental examinations over months and years [51].

In conclusion, in this study we describe the performance of two large, well-selected HC datasets from a European and American database, respectively. These HC datasets, selected through a stringent two-step analytical pipeline, perform optimally and equally well for the purpose of single-subject brain metabolism

assessment. The HC datasets described in the current study are ready-to-use, and the analytical pipeline here described might be helpful to other research centres, promoting the quality and reliability of brain metabolism estimation for the wide clinical and research community.

Information Sharing Statement

We will make available the HC datasets to the whole research and clinical community. All ADNI data are shared without embargo through the LONI Image and Data Archive (IDA), a secure research data repository. Interested scientists may obtain access to ADNI imaging, clinical, genomic, and biomarker data for the purposes of scientific investigation, teaching, or planning clinical research studies. Details about the ADNI-HC subjects selected for this study are reported in the Supplementary Material, Table 1. For those researchers who intend to use the ADNI-HC imaging database for their studies, it is recommended to follow the full pre- and post-processing procedures presented in the current paper in order to obtain comparable results.

The AIMN-HC dataset will be soon available on the SPM official website http://www.fil.ion.ucl.ac.uk/spm/ through the following URL: https://dorian.ge.infn.it/.

Compliance with ethical standards

The authors declare that they have no conflict of interest.

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Table 1. Descriptive statistics of the selected healthy controls in the HSR, AIMN and ADNI datasets.

	HSR-HC	AIMN-HC	ADNI-HC
N	112	125	75
Sex (M/F), N	53/59	58/67	33/42
Age, mean±SD	64.68±9.34	65.78±11.33	72.25±5.25
Education, mean±SD	-	10.5±4.29	16.60±2.48
MMSE, mean±SD	-	28.75±1.69	28.97±1.40

Table 2. Descriptive statistics of the patient groups

	Probable ADD	Probable DLB	bvFTD
N	79	81	56
Sex (M/F), N	32/47	27/29	43/38
Age, mean±SD	67.58±7.51	72.10±7.60	69.41±7.67
Education, mean±SD	11.83±6.96	9.44±4.52	9.00±5.66
MMSE, mean±SD	18.80±3.65	17.54±4.47	19.77±6.67

Table 3. Statistical accuracy of the SPM t- hypometabolism maps to correctly classify clinical diagnosis for different HC-AIMN sample sizes and statistical thresholds.

	p<0.05, Uncorrected				p<0.01, Uncorrected				p<0.05, FWE-corrected			
	n=10	n=30	n=75	n=125	n=10	n=30	n=75	n=125	n=10	n=30	n=75	n=125
ADD	87%	89%	90%	92%	83%	88%	87%	88%	0%	95%	90%	88%
DLB	81%	87%	85%	85%	78%	83%	81%	81%	0%	62%	69%	68%
bvFTD	93%	94%	93%	93%	93%	93%	93%	92%	0%	85%	88%	88%
Total	87%	90%	89%	90%	85%	88%	87%	87%	0%	80%	82%	81%

Abbreviations: ADD, Alzheimer's disease dementia; DLB, dementia with Lewy body; bvFTD, behavioural variant of frontotemporal dementia; FWE, family-wise error correction.
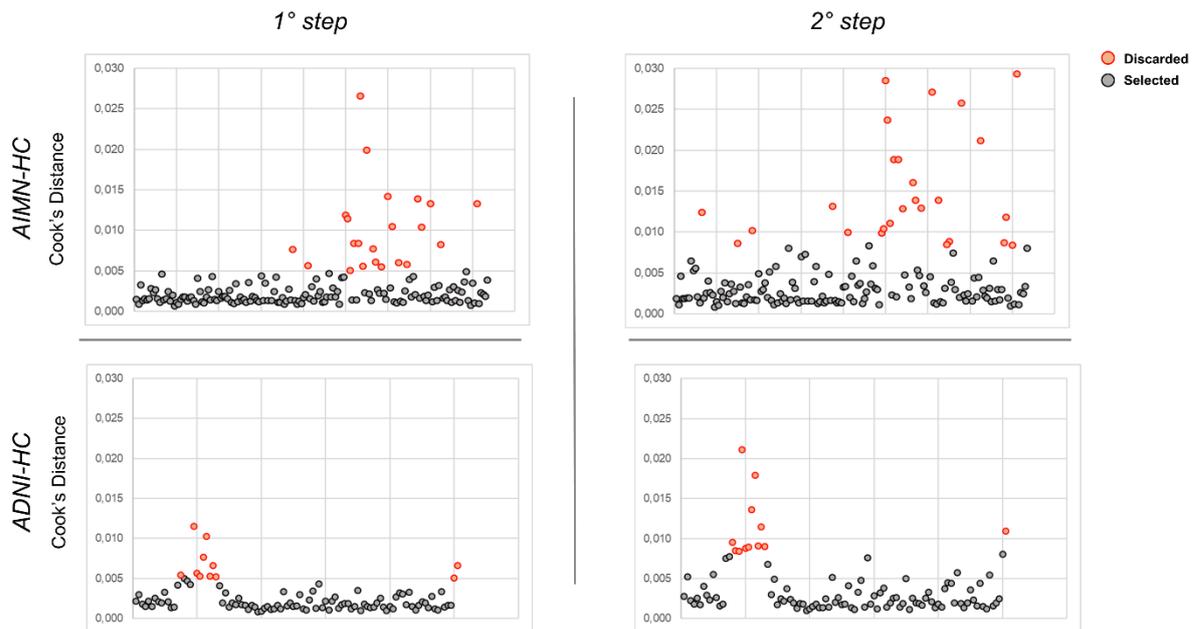
Table 4. Statistical accuracy of the SPM t- hypometabolism maps to correctly classify clinical diagnosis for different HC-ADNI sample sizes and statistical thresholds.

	p<0.05, uncorrected			p<0.01, uncorrected			p<0.05, FWE-corrected		
	n=10	n=30	n=75	n=10	n=30	n=75	n=10	n=30	n=75
ADD	95%	95%	95%	92%	94%	90%	0%	93%	87%
DLB	86%	88%	88%	85%	87%	85%	0%	63%	63%
bvFTD	92%	92%	93%	91%	91%	94%	0%	88%	89%
Total	91%	92%	92%	89%	91%	90%	0%	81%	80%

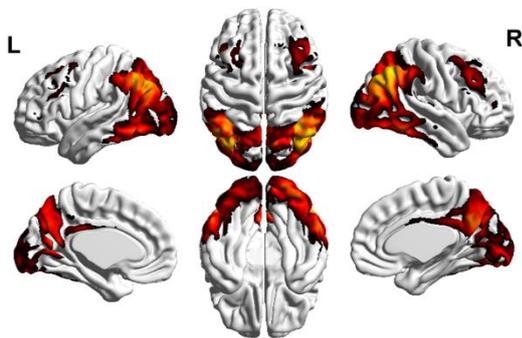**Figure 1. Scatter plot of the two steps Cook's distance analysis**.

Outliers (red dots) are represented by [18F]FDG PET images with larger Cook's distance values than the critical D value. The analysis was performed separately for both AIMN-HC and ADNI-HC datasets (upper and bottom panel respectively).
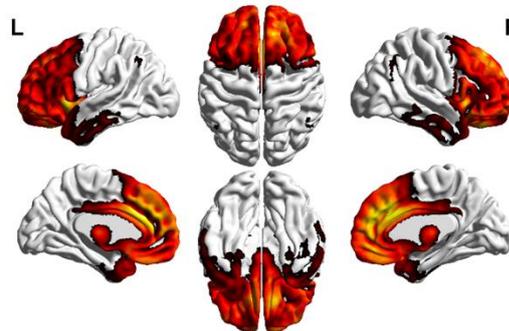
COMMONALITIES

CONCORDANCE

Alzheimer's disease dementia

Alzheimer's disease dementia

Dementia with Lewy bodies
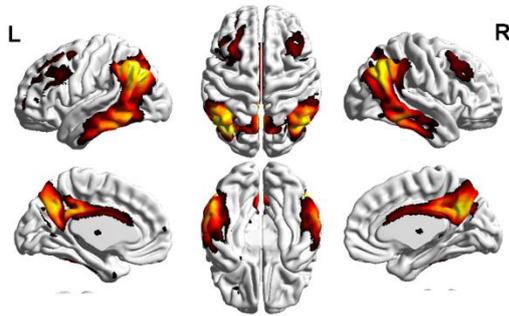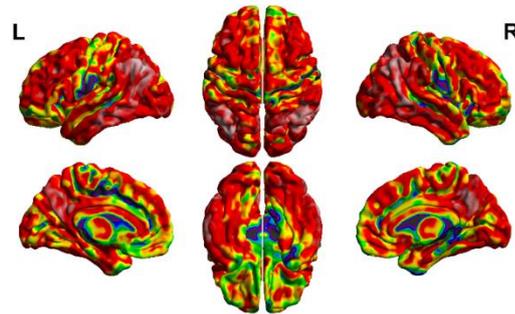
Dementia with Lewy bodies

Behavioral variant of frontotemporal dementia

Behavioral variant of frontotemporal dementia

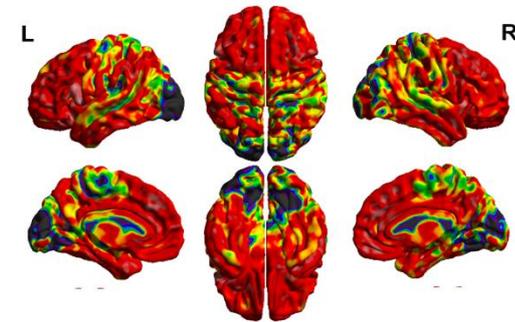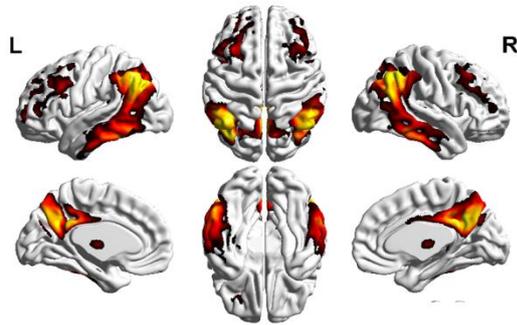0.30 *Commonality* 1

0 *DICE coefficient* 1
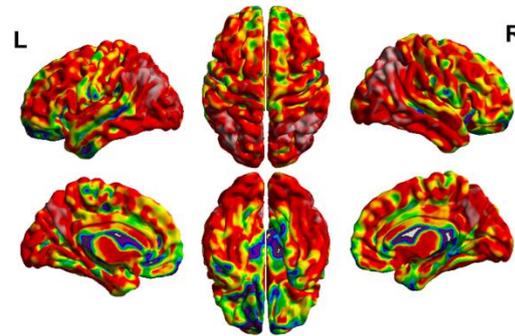
Figure 2. Voxel-wise maps of hypometabolism commonalities and concordance (AIMN-HC dataset).

Left column: commonalities of the SPM t-maps in ADD, DLB and bvFTD groups. The value of each voxel represents the proportion of patients, in each group, presenting with hypometabolism in that voxel, as estimated from the comparison with the AIMN-HC dataset. Commonality maps are fairly consistent with the known patterns of hypometabolism in each dementing condition. Only voxels with a commonality>0.30 are shown, for visualization purposes.

Right column: concordance between SPM t-maps in ADD, DLB and bvFTD groups. The value of each voxel represents the average amount of spatial overlap between single-subjects hypometabolism maps, as estimated from the comparison with the AIMN-HC dataset, vs. single-subjects hypometabolism maps estimated from the comparison with the reference HSR-HC dataset.

Figure 3. Voxel-wise maps of hypometabolism commonalities and concordance (ADNI-HC dataset).

Left column: commonalities of the SPM t-maps in ADD, DLB and bvFTD groups. The value of each voxel represents the proportion of patients, in each group, presenting with hypometabolism in that voxel, as estimated from the comparison with the ADNI-HC dataset. Commonality maps are fairly consistent with the known patterns of hypometabolism in each dementing condition. Only voxels with a commonality>0.30 are shown, for visualization purposes.
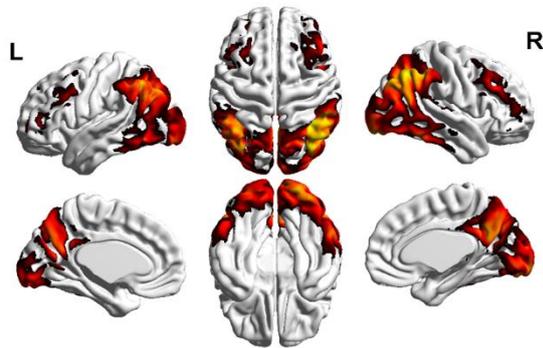
Right column: concordance between SPM t-maps in ADD, DLB and bvFTD groups. The value of each voxel represents the average amount of spatial overlap between single-subject hypometabolism maps, as estimated from the comparison with the ADNI-HC dataset, vs. single-subjects hypometabolism maps estimated from the comparison with the reference HSR-HC dataset.
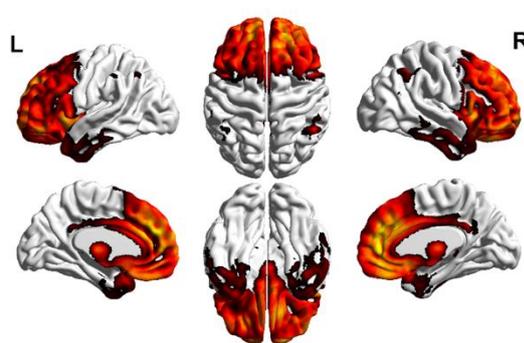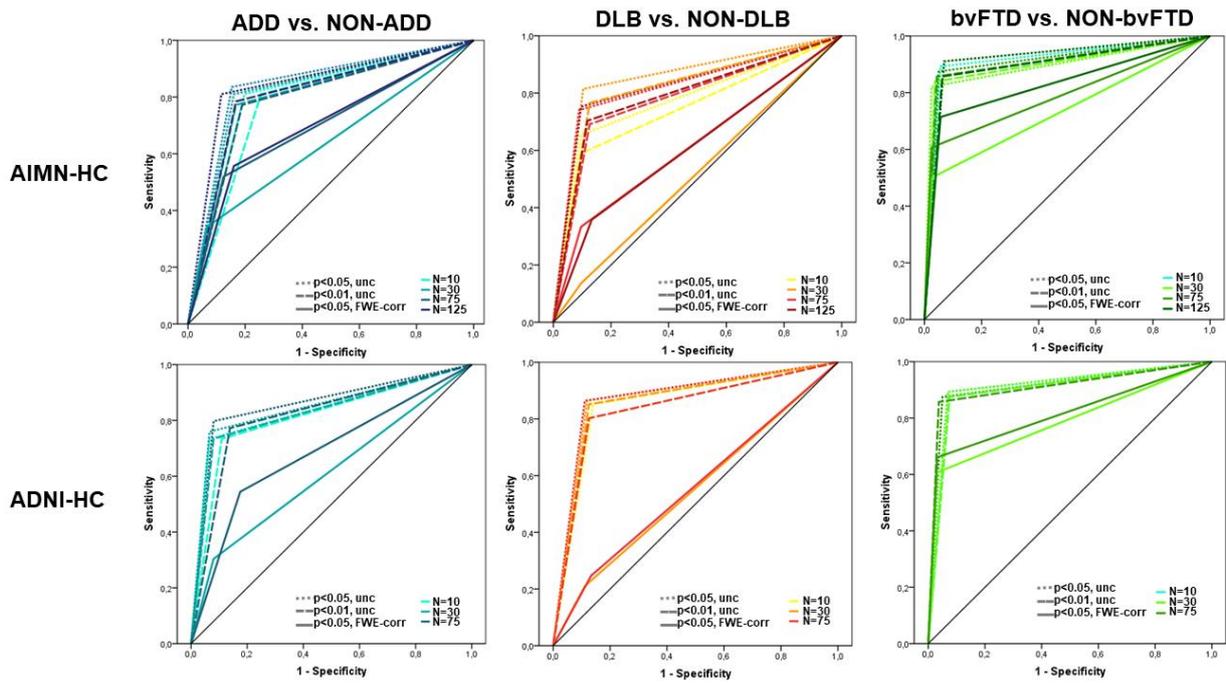
Figure 4. Receiver Operating Characteristic (ROC) curves. ROC curves comparing diagnostic performance of SPM t-maps using normal database with different sample sizes (10, 30, 75, 125) and different statistical thresholds and corrections (p<0.05 uncorrected, p<0.01 uncorrected and p<0.05 FWE-corrected) for correctly classifying patients with ADD, DLB and bvFTD.
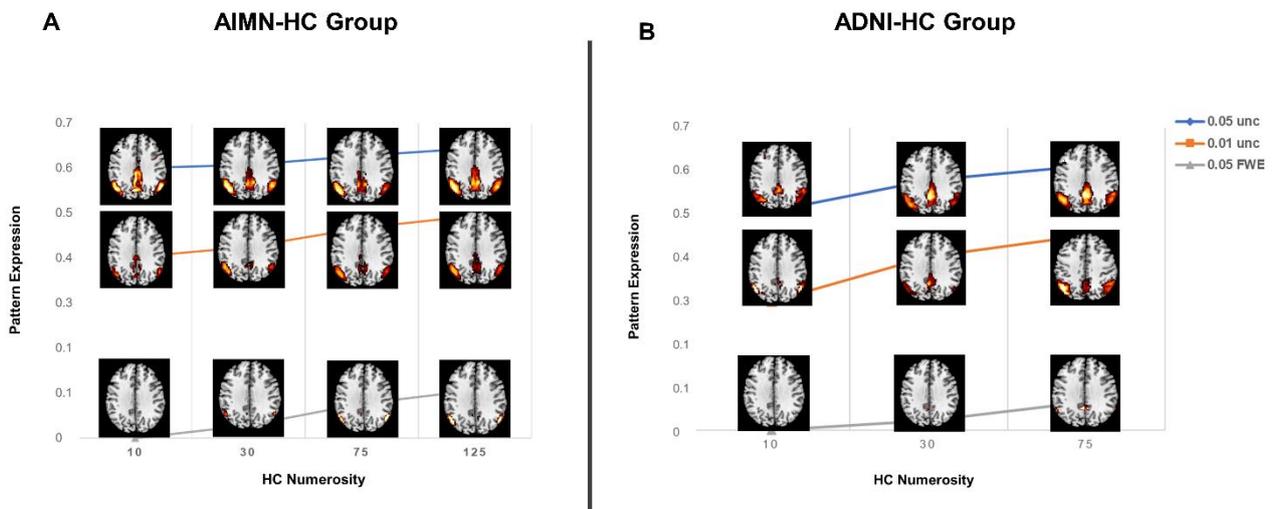
Figure 5. Pattern expression in the ADD group when different statistical thresholds and HC sample sizes are applied for the extraction of SPM single-subject hypometabolism maps. Panel **A** and **B** shows results obtained when considering the AIMN-HC dataset and the ADNI-HC datasets for the SPM single-subject analysis, respectively. Brain renderings of group-level hypometabolism maps at different thresholds and with different HC sample sizes are shown for illustrative purposes.

References

1. Kato T, Inui Y, Nakamura A, Ito K. Brain fluorodeoxyglucose (FDG) PET in dementia. Ageing Res Rev. Elsevier B.V.; 2016;30:73–84.

2. Leonardo I, Arianna S, Caminiti SP, Daniela P. The emerging role of PET imaging in dementia. F1000Research. Faculty of 1000 Ltd.; 2017;6.

3. Perani D, Caminiti SP, Carli G, Tondo G. PET Neuroimaging in Dementia Conditions. PET SPECT Neurol. Springer; 2020;211–82.

4. Perani D, Schillaci O, Padovani A, Nobili F, Leonardo I, Anthony P, et al. A survey of FDG-and amyloid-PET imaging in dementia and GRADE analysis. Biomed Res Int. Hindawi; 2014;2014.

5. McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology. 2005;65:1863–72.

6. Gorno-Tempini M., Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011;76:1006–14.

7. Armstrong MJ, Litvan I, Lang AE, Bak TH, Bhatia KP, Borroni B, et al. Criteria for the diagnosis of corticobasal degeneration. Neurology. 2013;80:496–503.

8. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, 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. Alzheimer's Dement. Elsevier Ltd; 2011;7:280–92.

9. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, 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. Elsevier Ltd; 2011;7:270–9.

10. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain. 2011;134:2456–77.

11. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, 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. Alzheimer's Dement. Elsevier Ltd; 2011;7:263–9.

12. Smailagic N, Vacante M, Hyde C, Martin S, Ukoumunne O, Sachpekidis C. 18 F- FDG PET for the early diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev. John Wiley & Sons, Ltd; 2015;

13. Morbelli S, Garibotto V, Van De Giessen E, Arbizu J, Chételat G, Drezgza A, et al. A Cochrane review on brain [18 F] FDG PET in dementia: limitations and future perspectives. Springer; 2015.

14. Waxman AD, Herholz K, Lewis DH, Herscovitch P, Minoshima S, Mountz JM, et al. Society of Nuclear Medicine procedure guideline for FDG PET brain imaging. Soc Nucl Med (Version 10). 2009;

15. Chen K, Ayutyanont N, Langbaum JBS, Fleisher AS, Reschke C, Lee W, et al. Characterizing Alzheimer's disease using a hypometabolic convergence index. Neuroimage. Elsevier; 2011;56:52–60.

16. Landau S, Jagust W. UC Berkeley FDG MetaROI Methods. Alzheimer's Dis Neuroimaging Initiat. 2011;

17. López-González FJ, Silva-Rodríguez J, Paredes-Pacheco J, Niñerola-Baizán A, Efthimiou N, Martín-Martín C, et al. Intensity normalization methods in brain FDG-PET quantification. Neuroimage. Elsevier; 2020;222:117229.

18. Nobili F, Festari C, Altomare D, Agosta F, Orini S, Van Laere K, et al. Automated assessment of FDG-PET for differential diagnosis in patients with neurodegenerative disorders. Eur J Nucl Med Mol Imaging. Springer; 2018;45:1557–66.

19. Mosconi L, Tsui WH, Pupi A, De Santi S, Drzezga A, Minoshima S, et al. 18F-FDG PET database of longitudinally confirmed healthy elderly individuals improves detection of mild cognitive impairment and Alzheimer's disease. J Nucl Med. Soc Nuclear Med; 2007;48:1129–34.

20. Perani D, Della Rosa PA, Cerami C, Gallivanone F, Fallanca F, Giovanna E, et al. Validation of an optimized SPM procedure for FDG-PET in dementia diagnosis in a clinical setting. NeuroImage Clin. Elsevier B.V.; 2014;6:445–54.

21. Della Rosa PA, Cerami C, Gallivanone F, Prestia A, Caroli A, Castiglioni I, et al. A standardized [18 F]-FDG-PET template for spatial normalization in statistical parametric mapping of dementia. Neuroinformatics. Springer; 2014;12:575–93.

22. Della Rosa PA, Cerami C, Gallivanone F, Prestia A, Caroli A, Castiglioni I, et al. A Standardized [18F]-FDG-PET Template for Spatial Normalization in Statistical Parametric Mapping of Dementia. Neuroinformatics. Springer; 2014;12:575–93.

23. Caminiti SP, Alongi P, Majno L, Volontè MA, Cerami C, Gianolli L, et al. Evaluation of an optimized [18F] fluoro- deoxy- glucose positron emission tomography voxel- wise method to early support differential diagnosis in atypical Parkinsonian disorders. Eur J Neurol. Wiley Online Library; 2017;24:687-e26.

24. Caminiti SP, Sala A, Iaccarino L, Beretta L, Pilotto A, Gianolli L, et al. Brain glucose metabolism in Lewy body dementia: Implications for diagnostic criteria. Alzheimer's Res Ther. 2019;11.

25. Iaccarino L, Chiotis K, Alongi P, Almkvist O, Wall A, Cerami C, et al. A cross-validation of FDG-and amyloid-PET biomarkers in mild cognitive impairment for the risk prediction to dementia due to Alzheimer's disease in a clinical setting. J Alzheimer's Dis. IOS Press; 2017;59:603–14.

26. Cerami C, Dodich A, Greco L, Iannaccone S, Magnani G, Marcone A, et al. The role of single-subject brain metabolic patterns in the early differential diagnosis of primary progressive aphasias and in prediction of progression to dementia. J Alzheimer's Dis. IOS Press; 2017;55:183–97.

27. Sala A, Caprioglio C, Santangelo R, Vanoli EG, Iannaccone S, Magnani G, et al. Brain metabolic signatures across the Alzheimer's disease spectrum. Eur J Nucl Med Mol Imaging. Springer; 2020;47:256–69.

28. Varrone A, Asenbaum S, Vander Borght T, Booij J, Nobili F, Någren K, et al. EANM procedure guidelines for PET brain imaging using [18F]FDG, version 2. Eur J Nucl Med Mol Imaging. 2009;36:2103–10.

29. Jian Y, Planeta B, Carson RE. Evaluation of bias and variance in low-count OSEM list mode reconstruction. Phys Med Biol. IOP Publishing; 2014;60:15.

30. Buchert R, Wilke F, Chakrabarti B, Martin B, Brenner W, Mester J, et al. Adjusted scaling of FDG positron emission tomography images for statistical evaluation in patients with suspected Alzheimer's disease. J Neuroimaging. Wiley Online Library; 2005;15:348–55.

31. Nichols TE. Multiple testing corrections, nonparametric methods, and random field theory. Neuroimage. Elsevier; 2012;62:811–5.

32. Presotto L, Ballarini T, Caminiti SP, Bettinardi V, Gianolli L, Perani D. Validation of ^{18}F–FDG-PET Single-Subject Optimized SPM Procedure with Different PET Scanners. Neuroinformatics. 2017;15.

33. Rahmim A, Qi J, Sossi V. Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls. Med Phys. Wiley Online Library; 2013;40.

34. Kaalep A, Sera T, Rijnsdorp S, Yaqub M, Talsma A, Lodge MA, et al. Feasibility of state of the art PET/CT systems performance harmonisation. Eur J Nucl Med Mol Imaging. Springer; 2018;45:1344–61.

35. Cerami C, Dodich A, Lettieri G, Iannaccone S, Magnani G, Marcone A, et al. Different FDG-PET metabolic patterns at single-subject level in the behavioral variant of fronto-temporal dementia. Elsevier Ltd; 2016;83:101–12.

36. Caminiti SP, Sala A, Iaccarino L, Beretta L, Pilotto A, Gianolli L, et al. Brain glucose metabolism in Lewy body dementia: implications for diagnostic criteria. Alzheimers Res Ther. BioMed Central; 2019;11:20.

37. McKeith IG, Boeve BF, DIckson DW, Halliday G, Taylor JP, Weintraub D, et al. Diagnosis and management of dementia with Lewy bodies. Neurology. 2017;89:88–100.

38. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, 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. Alzheimer's Dement. Elsevier; 2011;7:270–9.

39. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, 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. Alzheimer's Dement. Elsevier; 2011;7:280–92.

40. Armstrong MJ, Litvan I, Lang AE, Bak TH, Bhatia KP, Borroni B, et al. Criteria for the diagnosis of corticobasal degeneration. Neurology. AAN Enterprises; 2013;80:496–503.

41. Teune LK, Bartels AL, de Jong BM, Willemsen ATM, Eshuis SA, de Vries JJ, et al. Typical cerebral metabolic patterns in neurodegenerative brain diseases. Mov Disord. Wiley Online Library; 2010;25:2395–404.

42. Caminiti SP, Alongi P, Majno L, Volontè MA, Cerami C, Gianolli L, et al. Evaluation of an optimized [^{18}F]fluoro-deoxy-glucose positron emission tomography voxel-wise method to early support differential diagnosis in atypical Parkinsonian disorders. Eur J Neurol. 2017;24.

43. Mosconi L, Tsui WH, Herholz K, Pupi A, Drzezga A, Lucignani G, et al. Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer's disease, and other dementias. Eur J Nucl Med Mol Imaging. Soc Nuclear Med; 2008;49:390–8.

44. Perani D, Iaccarino L, Bettinardi V. The need for "objective measurements" in FDG and amyloid PET neuroimaging. Clin Transl Imaging. Springer; 2014;2:331–42.

45. Presotto L, Ballarini T, Caminiti SP, Bettinardi V, Gianolli L, Perani D. Validation of 18 F–FDG-PET Single-Subject Optimized SPM Procedure with Different PET Scanners. Neuroinformatics. Springer; 2017;15:151–63.

46. Gordon BA, Blazey TM, Su Y, Hari-Raj A, Dincer A, Flores S, et al. Spatial patterns of neuroimaging biomarker change in individuals from families with autosomal dominant Alzheimer's disease: a longitudinal study. Lancet Neurol. Elsevier; 2018;17:241–50.

47. Mosconi L. Glucose metabolism in normal aging and Alzheimer's disease: methodological and physiological considerations for PET studies. Clin Transl imaging. Springer; 2013;1:217–33.

48. Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res. Sage Publications Sage CA: Thousand Oaks, CA; 2003;12:419–46.

49. Chen W-P, Samuraki M, Yanase D, Shima K, Takeda N, Ono K, et al. Effect of sample size for normal database on diagnostic performance of brain FDG PET for the detection of Alzheimer's disease using automated image analysis. Nucl Med Commun. LWW; 2008;29:270–6.

50. Gallivanone F. The impact of different 18FDG PET healthy subject scans for comparison with single patient in SPM analysis. Q J Nucl Med Mol imaging. Minerva medica; 2014;61:115–32.

51. Cerami C, Della Rosa PA, Magnani G, Santangelo R, Marcone A, Cappa SF, et al. Brain metabolic maps in Mild Cognitive Impairment predict heterogeneity of progression to dementia. NeuroImage Clin. Elsevier; 2015;7:187–94.

2. Leonardo I, Arianna S, Caminiti SP, Daniela P. The emerging role of PET imaging in dementia. F1000Research. Faculty of 1000 Ltd.; 2017;6.

3. Perani D, Caminiti SP, Carli G, Tondo G. PET Neuroimaging in Dementia Conditions. PET SPECT Neurol. Springer; 2020;211–82.

4. Perani D, Schillaci O, Padovani A, Nobili F, Leonardo I, Anthony P, et al. A survey of FDG-and amyloid-PET imaging in dementia and GRADE analysis. Biomed Res Int. Hindawi; 2014;2014.

5. McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology. 2005;65:1863–72.

6. Gorno-Tempini M., Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011;76:1006–14.

7. Armstrong MJ, Litvan I, Lang AE, Bak TH, Bhatia KP, Borroni B, et al. Criteria for the diagnosis of corticobasal degeneration. Neurology. 2013;80:496–503.

8. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, 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. Alzheimer's Dement. Elsevier Ltd; 2011;7:280–92.

9. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, 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. Elsevier Ltd; 2011;7:270–9.

10. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain. 2011;134:2456–77.

11. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, 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. Alzheimer's Dement. Elsevier Ltd; 2011;7:263–9.

12. Smailagic N, Vacante M, Hyde C, Martin S, Ukoumunne O, Sachpekidis C. 18 F- FDG PET for the early diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev. John Wiley & Sons, Ltd; 2015;

13. Morbelli S, Garibotto V, Van De Giessen E, Arbizu J, Chételat G, Drezgza A, et al. A Cochrane review on brain [18 F] FDG PET in dementia: limitations and future perspectives. Springer; 2015.

14. Waxman AD, Herholz K, Lewis DH, Herscovitch P, Minoshima S, Mountz JM, et al. Society of Nuclear Medicine procedure guideline for FDG PET brain imaging. Soc Nucl Med (Version 10). 2009;

15. Chen K, Ayutyanont N, Langbaum JBS, Fleisher AS, Reschke C, Lee W, et al. Characterizing Alzheimer's disease using a hypometabolic convergence index. Neuroimage. Elsevier; 2011;56:52–60.

16. Landau S, Jagust W. UC Berkeley FDG MetaROI Methods. Alzheimer's Dis Neuroimaging Initiat. 2011;

17. López-González FJ, Silva-Rodríguez J, Paredes-Pacheco J, Niñerola-Baizán A, Efthimiou N, Martín-Martín C, et al. Intensity normalization methods in brain FDG-PET quantification. Neuroimage. Elsevier; 2020;222:117229.

18. Nobili F, Festari C, Altomare D, Agosta F, Orini S, Van Laere K, et al. Automated assessment of FDG-PET for differential diagnosis in patients with neurodegenerative disorders. Eur J Nucl Med Mol Imaging. Springer; 2018;45:1557–66.

19. Mosconi L, Tsui WH, Pupi A, De Santi S, Drzezga A, Minoshima S, et al. 18F-FDG PET database of longitudinally confirmed healthy elderly individuals improves detection of mild cognitive impairment and Alzheimer's disease. J Nucl Med. Soc Nuclear Med; 2007;48:1129–34.

20. Perani D, Della Rosa PA, Cerami C, Gallivanone F, Fallanca F, Giovanna E, et al. Validation of an optimized SPM procedure for FDG-PET in dementia diagnosis in a clinical setting. NeuroImage Clin. Elsevier B.V.; 2014;6:445–54.

21. Della Rosa PA, Cerami C, Gallivanone F, Prestia A, Caroli A, Castiglioni I, et al. A standardized [18 F]-FDG-PET template for spatial normalization in statistical parametric mapping of dementia. Neuroinformatics. Springer; 2014;12:575–93.

22. Della Rosa PA, Cerami C, Gallivanone F, Prestia A, Caroli A, Castiglioni I, et al. A Standardized [18F]-FDG-PET Template for Spatial Normalization in Statistical Parametric Mapping of Dementia. Neuroinformatics. Springer; 2014;12:575–93.

23. Caminiti SP, Alongi P, Majno L, Volontè MA, Cerami C, Gianolli L, et al. Evaluation of an optimized [18F] fluoro- deoxy- glucose positron emission tomography voxel- wise method to early support differential diagnosis in atypical Parkinsonian disorders. Eur J Neurol. Wiley Online Library; 2017;24:687-e26.

24. Caminiti SP, Sala A, Iaccarino L, Beretta L, Pilotto A, Gianolli L, et al. Brain glucose metabolism in Lewy body dementia: Implications for diagnostic criteria. Alzheimer's Res Ther. 2019;11.

25. Iaccarino L, Chiotis K, Alongi P, Almkvist O, Wall A, Cerami C, et al. A cross-validation of FDG-and amyloid-PET biomarkers in mild cognitive impairment for the risk prediction to dementia due to Alzheimer's disease in a clinical setting. J Alzheimer's Dis. IOS Press; 2017;59:603–14.

26. Cerami C, Dodich A, Greco L, Iannaccone S, Magnani G, Marcone A, et al. The role of single-subject brain metabolic patterns in the early differential diagnosis of primary progressive aphasias and in prediction of progression to dementia. J Alzheimer's Dis. IOS Press; 2017;55:183–97.

27. Sala A, Caprioglio C, Santangelo R, Vanoli EG, Iannaccone S, Magnani G, et al. Brain metabolic signatures across the Alzheimer's disease spectrum. Eur J Nucl Med Mol Imaging. Springer; 2020;47:256–69.

28. Varrone A, Asenbaum S, Vander Borght T, Booij J, Nobili F, Någren K, et al. EANM procedure guidelines for PET brain imaging using [18F]FDG, version 2. Eur J Nucl Med Mol Imaging. 2009;36:2103–10.

29. Buchert R, Wilke F, Chakrabarti B, Martin B, Brenner W, Mester J, et al. Adjusted scaling of FDG positron emission tomography images for statistical evaluation in patients with suspected Alzheimer's disease. J Neuroimaging. Wiley Online Library; 2005;15:348–55.

30. Presotto L, Ballarini T, Caminiti SP, Bettinardi V, Gianolli L, Perani D. Validation of ^{18}F–FDG-PET Single-Subject Optimized SPM Procedure with Different PET Scanners. Neuroinformatics. 2017;15.

31. Rahmim A, Qi J, Sossi V. Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls. Med Phys. Wiley Online Library; 2013;40.

32. Kaalep A, Sera T, Rijnsdorp S, Yaqub M, Talsma A, Lodge MA, et al. Feasibility of state of the art PET/CT systems performance harmonisation. Eur J Nucl Med Mol Imaging. Springer; 2018;45:1344–61.

33. Cerami C, Dodich A, Lettieri G, Iannaccone S, Magnani G, Marcone A, et al. Different FDG-PET metabolic patterns at single-subject level in the behavioral variant of fronto-temporal dementia. Elsevier Ltd; 2016;83:101–12.

34. Caminiti SP, Sala A, Iaccarino L, Beretta L, Pilotto A, Gianolli L, et al. Brain glucose metabolism in Lewy body dementia: implications for diagnostic criteria. Alzheimers Res Ther. BioMed Central; 2019;11:20.

35. McKeith IG, Boeve BF, DIckson DW, Halliday G, Taylor JP, Weintraub D, et al. Diagnosis and management of dementia with Lewy bodies. Neurology. 2017;89:88–100.

36. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, 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. Alzheimer's Dement. Elsevier; 2011;7:270–9.

37. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, 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. Alzheimer's Dement. Elsevier; 2011;7:280–92.

38. Armstrong MJ, Litvan I, Lang AE, Bak TH, Bhatia KP, Borroni B, et al. Criteria for the diagnosis of corticobasal degeneration. Neurology. AAN Enterprises; 2013;80:496–503.

39. Teune LK, Bartels AL, de Jong BM, Willemsen ATM, Eshuis SA, de Vries JJ, et al. Typical cerebral metabolic patterns in neurodegenerative brain diseases. Mov Disord. Wiley Online Library; 2010;25:2395–404.

40. Caminiti SP, Alongi P, Majno L, Volontè MA, Cerami C, Gianolli L, et al. Evaluation of an optimized [^{18}F]fluoro-deoxy-glucose positron emission tomography voxel-wise method to early support differential diagnosis in atypical Parkinsonian disorders. Eur J Neurol. 2017;24.

41. Mosconi L, Tsui WH, Herholz K, Pupi A, Drzezga A, Lucignani G, et al. Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer's disease, and other dementias. Eur J Nucl Med Mol Imaging. Soc Nuclear Med; 2008;49:390–8.

42. Perani D, Iaccarino L, Bettinardi V. The need for "objective measurements" in FDG and amyloid PET neuroimaging. Clin Transl Imaging. Springer; 2014;2:331–42.

43. Presotto L, Ballarini T, Caminiti SP, Bettinardi V, Gianolli L, Perani D. Validation of 18 F–FDG-PET Single-Subject Optimized SPM Procedure with Different PET Scanners. Neuroinformatics. Springer; 2017;15:151–63.

44. Gordon BA, Blazey TM, Su Y, Hari-Raj A, Dincer A, Flores S, et al. Spatial patterns of neuroimaging biomarker change in individuals from families with autosomal dominant Alzheimer's disease: a longitudinal study. Lancet Neurol. Elsevier; 2018;17:241–50.

45. Mosconi L. Glucose metabolism in normal aging and Alzheimer's disease: methodological and physiological considerations for PET studies. Clin Transl imaging. Springer; 2013;1:217–33.

46. Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res. Sage Publications Sage CA: Thousand Oaks, CA; 2003;12:419–46.

47. Chen W-P, Samuraki M, Yanase D, Shima K, Takeda N, Ono K, et al. Effect of sample size for normal database on diagnostic performance of brain FDG PET for the detection of Alzheimer's disease using automated image analysis. Nucl Med Commun. LWW; 2008;29:270–6.

48. Gallivanone F. The impact of different 18FDG PET healthy subject scans for comparison with single patient in SPM analysis. Q J Nucl Med Mol imaging. Minerva medica; 2014;61:115–32.

49. Cerami C, Della Rosa PA, Magnani G, Santangelo R, Marcone A, Cappa SF, et al. Brain metabolic maps in Mild Cognitive Impairment predict heterogeneity of progression to dementia. NeuroImage Clin. Elsevier; 2015;7:187–94.

Figure 4
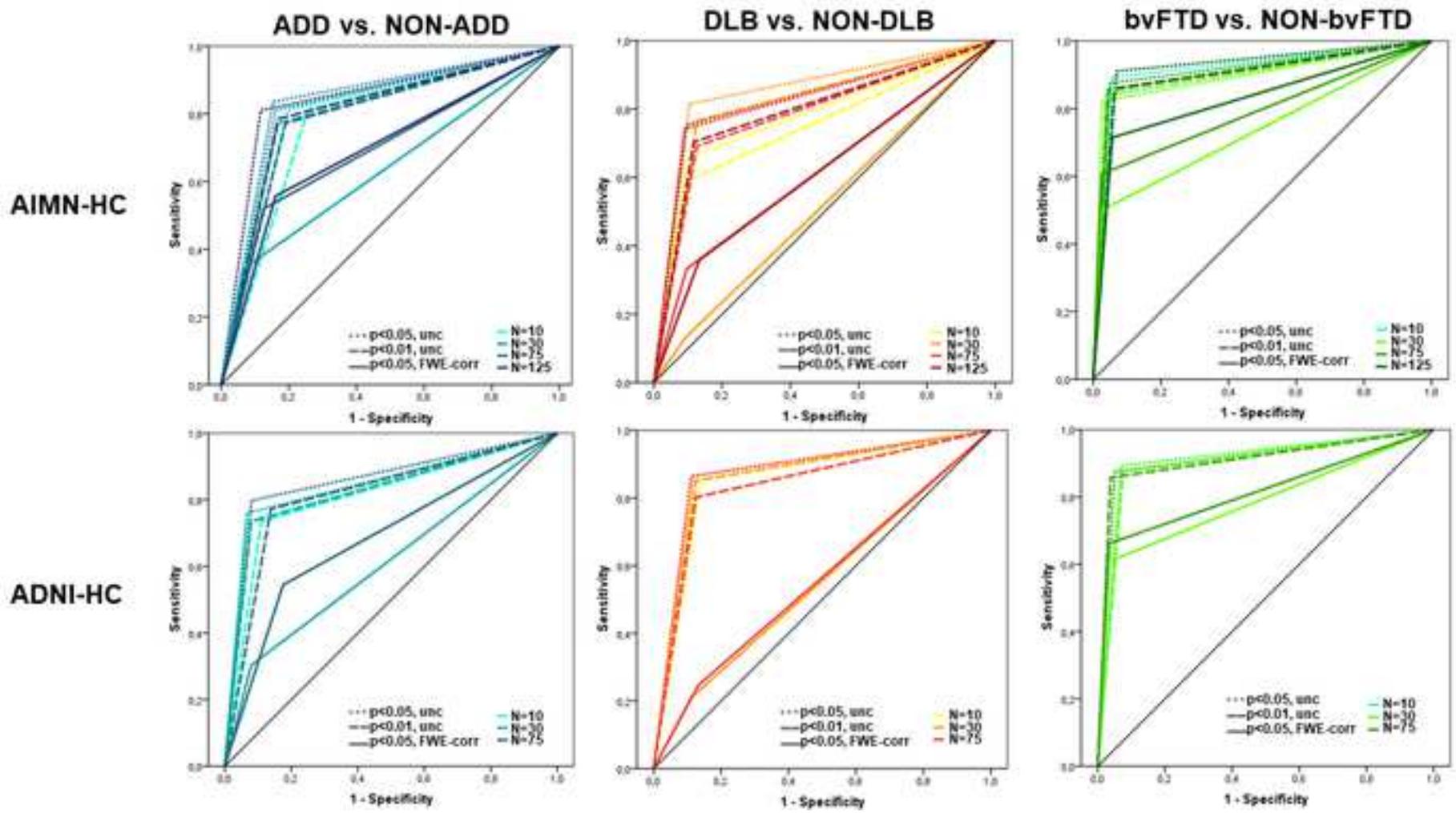
Click here to access/download

Supplementary Material

Supplementary Materials_141120 (1).docx