## Original Paper

# Education-Adjusted Normality Thresholds for FDG-PET in the Diagnosis of Alzheimer Disease

Ismini C. Mainta[a, b]    Sara Trombella[b, c]    Silvia Morbelli[d]    Giovanni B. Frisoni[c, e]
Valentina Garibotto[a, b]    Alzheimer Disease Neuroimaging Initiative (ADNI)

[a]Nuclear Medicine and Molecular Imaging Division, Geneva University Hospitals, Geneva, Switzerland; [b]Laboratory of Neuroimaging and Innovative Molecular Tracers, Geneva University, Geneva, Switzerland; [c]LANVIE, Geneva University, Geneva, Switzerland; [d]Nuclear Medicine Unit, IRCCS San Martino – IST, University of Genoa, Genoa, Italy; [e]Memory Center, Geneva University Hospitals, Geneva, Switzerland

**Abstract**

*Background:* A corollary of the reserve hypothesis is that what is regarded as pathological cortical metabolism in patients might vary according to education. *Objective:* The aim of this study is to assess the incremental diagnostic value of education-adjusted over unadjusted thresholds on the diagnostic accuracy of FDG-PET as a biomarker for Alzheimer disease (AD). *Methods:* We compared cortical metabolism in 90 healthy controls and 181 AD patients from the Alzheimer Disease Neuroimaging Initiative (ADNI) database. The AUC of the ROC curve did not differ significantly between the whole group and the higher-education patients or the lower-education subjects. *Results:* The threshold of wMetaROI values providing 80% sensitivity was lower in higher-education patients and higher in the lower-education patients, compared to the standard threshold derived over the whole AD collective, without, however, significant changes in sensitivity and specificity. *Conclusion:* These data show that education, as a proxy of reserve, is not a major confounder in the diagnostic accuracy of FDG-PET in AD and the adoption of education-adjusted thresholds is not required in daily practice.

© 2018 S. Karger AG, Basel

## Introduction

The concept of reserve in Alzheimer disease (AD) has been introduced to explain the dissociation between the extent of the underlying AD pathology and the clinical expression of dementia [1, 2]. Reserve is conceptualized either in the form of neuronal/brain reserve or in the form of cognitive reserve, or compensation, thanks to neurocomputational flexibility and/or intellectual competence [3].

Multiple proxies, such as measures of intelligence, years of schooling, occupational nature and complexity,

Ismini C. Mainta
Nuclear Medicine Department, Geneva University Medical Center, Faculty of Medicine
University of Geneva, Rue Gabrielle-Perret-Gentil 4
CH–1211 Geneva (Switzerland)
E-Mail isminicharis.mainta @ hcuge.ch

and intellectual activities during leisure time, have been used to assess the degree of reserve [3–7]. Years of formal education, being easily quantifiable, is the proxy most commonly used.

FDG-PET (PET with 18F-fluorodeoxyglucose) is a sensitive and specific imaging biomarker for the diagnosis of AD and the best predictor of progression to dementia in mild cognitive impairment [8–11].

Values for sensitivity and specificity are quite variable, depending on disease severity as well as on the specific metric used, but the overall performance of the test is good, with a median sensitivity and specificity of 80 and 72.5%, respectively [9, 10].

Subjects with a higher reserve who are able to cope with a higher degree of pathology exhibit a more pronounced decrease in metabolism of the regions typically affected compared to patients with lower reserve for a comparable level of clinical impairment [4, 6, 7]. However, the extent of this difference and its potential impact on the diagnostic use of the test is not currently known.

The aim of this study was to evaluate the impact of years of education, as a proxy of reserve, on the diagnostic accuracy of FDG-PET as a biomarker for AD and the threshold of FDG hypometabolism that should be considered abnormal in populations with different levels of education.

## Materials and Methods

### Subjects

This study included 90 healthy controls (HC) and 181 AD patients selected from the Alzheimer's Disease Neuroimaging (ADNI) database (adni.loni.usc.edu). The ADNI is a public-private partnership launched in 2003 and led by Michael W. Weiner, focused on testing the value of various imaging and biological markers for the diagnosis of mild cognitive impairment and early AD. For up-to-date information, see www.adni-info.org.

We included all of the baseline PET scans of patients with a diagnosis of HC or AD available from the ADNI1 study in the preprocessed form (coregistered, averaged, standardized image and voxel size, uniform resolution), with a standard procedure described at http://adni.loni.usc.edu/methods/pet-analysis/pre-processing/. PET data and data on the clinical diagnosis, global functioning (measured using the Mini Mental State Examination [MMSE] score) years of education, age, and FDG-PET scans were extracted in December 2016. We excluded HC for which a subsequent follow-up evaluation showed progression to mild cognitive impairment or AD.

### FDG-PET Processing

In order to characterize each subject, a summary metric of hypometabolism at the sites typically affected by AD, i.e., the MetaROI set, was used [12]. Briefly, this index is computed with extrac-

**Table 1.** Characteristics of the patients and HC

| Characteristic | HC (n = 90) | AD (n = 181) | p value[1] |
|---|---|---|---|
| Age, years | 75.4±4.7 (62 to 87) | 75.6±7 (55 to 88) | 0.82 |
| Male/female ratio | 55/35 | 110/71 | |
| Education, years | 15.8±3.2 (7 to 20) | 15.3±3.1 (4 to 20) | 0.18 |
| MMSE score | 28.9±1.2 (25 to 30) | 25.1±2.6 (18 to 30) | <0.001 |
| wMetaROI score | 1.87±1 (–0.5 to 4.6) | 0.4±1.19 (–4.0 to 3.2) | <0.001 |

Values are presented as means ± SD (range). MMSE, Mini Mental State Examination; HC, healthy controls; AD, Alzheimer disease. [1] Two-tailed *t* test.

tion of the mean activity in a set of 5 predefined ROI (bilateral angular gyrus, posterior cingulum, and bilateral inferior temporal gyrus), normalized to the mean activity of the pons and the cerebellar vermis (as defined in the WFUpickatlas SPM toolbox) and then corrected for age, calculating the wMetaROI scores, as previously tested [13]. The resulting wMetaROI score for each subject was then used for the subsequent analyses.

In addition, adjustment of the wMetaROI score to the respective MMSE was conducted, adopting the same procedure used for age, in order to control for the possible effect of a different level of cognitive impairment between patients.

### Statistical Analyses

For descriptive statistics of the characteristics of the population, we used a *t* test.

The receiver operating characteristic (ROC) curve approach was used to compare the diagnostic accuracy of FDG-PET obtained collectively and in population subgroups defined on the basis of educational attainment.

We first dichotomized the AD population into 2 subgroups using the median value as a cut-off as follows: <16 years for patients with a lower reserve (n = 75) and ≥16 years for patients with a higher reserve (n = 106). Lastly, we repeated the same analysis by taking into consideration only the extremes of educational attainment, i.e., the 1st and 4th quartiles of the distribution, meaning ≤13 years for a lower reserve (n = 52) and ≥18 years for a higher reserve (n = 55). The area under the curve (AUC) values of the ROC curves were compared using a *t* test. The analysis was repeated after MMSE adjustment of the wMetaROI scores.

For the whole population as well as for subgroups defined on the basis of education, we calculated the cut-off that maximized specificity at a minimal sensitivity of 80%.

We tested the use of a standard versus an education-adjusted threshold, calculating the number of false negatives (FN) and false positives (FP) associated with the change in threshold and comparing sensitivities and specificities (McNemar test).

We also tested the use of education-adjusted thresholds in the whole population to estimate the impact of an index derived from a highly or less educated population on a heterogeneous sample using the McNemar test.

We performed a power analysis using PASS 2005 software (NCSS, PASS and GESS; NCSS, Kaysville, UT, USA) to estimate the sample size necessary to identify a difference of 0.1 in the AUC of the ROC curves. The 0.1 difference between AUC was chosen

**Table 2.** Descriptive statistics of LE (<16 years) and HE (≥16 years) patients

| Characteristic | LE <16 years ($n = 75$) | $p$ value[1] (compared to HC) | HE ≥16 years ($n = 106$) | $p$ value[1] (compared to HC) | $p$ value[1] (HE vs. LE) |
|---|---|---|---|---|---|
| Age, years | 76.3±6.6 (58 to 88) | 0.32 | 75.1±7.3 (55 to 88) | 0.74 | 0.26 |
| Male/female ratio | 38/37 | | 72/34 | | |
| Education, years | 12.2±2.1 (4 to 15) | | 17.4±1.5 (16 to 20) | | |
| MMSE score | 24.4±2.5 (20 to 30) | <0.001 | 25.5±2.6 (18 to 30) | <0.001 | <0.01 |
| wMetaROI score | 0.4±1.1 (−2.3 to 2.2) | <0.001 | 0.37±1.3 (−4.0 to 3.2) | <0.001 | 0.7 |

Values are presented as means ± SD (range). MMSE, Mini Mental State Examination; LE, lower education; HE, higher education; HC, healthy controls. [1] Two-tailed $t$ test.

**Table 3.** Descriptive statistics of LE (≤13 years) and HE (≥18 years) patients

| Characteristics | LE ≤13 years ($n = 52$) | $p$ value[1] (compared to HC) | HE ≥18 years ($n = 55$) | $p$ value[1] (compared to HC) | $p$ value[1] (LE vs. HE) |
|---|---|---|---|---|---|
| Age, years | 77.2±6.6 (59 to 88) | 0.06 | 74.8±7.6 (55 to 88) | 0.53 | 0.08 |
| Male/female ratio | 23/29 | | 37/18 | | |
| Education, years | 11.4±1.9 (4 to 13) | | 18.6±0.9 (18 to 20) | | |
| MMSE score | 24.4±2.5 (20 to 30) | <0.001 | 25.7±2.6 (20 to 30) | <0.001 | 0.01 |
| wMetaROI score | 0.6±1 (−1.8 to 2.2) | <0.001 | 0.27±1.1 (−2.8 to 2.1) | <0.001 | 0.09 |

Values are presented as means ± SD (range). MMSE, Mini Mental State Examination; LE, lower education; HE, higher education; HC, healthy controls. [1] Two-tailed $t$ test.

with reference to well-established ranges for the AUC of a diagnostic test as follows: 1–0.9, excellent; 0.9–0.8, good; and 0.8–0.7, fair. A population of 56 cases against 90 controls would allow achievement of 80% power to detect this difference at a significance level of 0.05 (one-tailed).

## Results

The characteristics of the study sample and lower-education (LE) and higher-education (HE) subgroups are presented in Tables 1 and 2.

As expected, AD subjects had a lower MMSE score ($p < 0.001$, two-tailed $t$ test) and wMetaROI ($p < 0.001$) compared to HC, with no statistically significant difference in years of education ($p = 0.18$) or age (Table 1).

When dichotomizing the population with respect to median education, subjects with an LE had a lower MMSE score ($p < 0.01$; Table 2). This was also true when comparing the two extremes, i.e., LE and HE subjects ($p = 0.01$; Table 3).

There was a trend, albeit a nonsignificant one ($p = 0.09$), toward a lower wMetaROI score in HE patients compared to LE patients despite a higher MMSE, as expected from a higher reserve capacity (Table 3).

The AUC of the ROC curve of wMetaROI values to discriminate the whole AD population from HC was 0.832. The cut-off value of the wMetaROI score providing a minimum of 80% sensitivity was 1.38 (sensitivity 80% and specificity 68%).

The AUC of the ROC curve was 0.831 in AD patients with a lower reserve (<16 years) compared to controls and 0.833 in AD patients with a higher reserve (≥16 years; Fig. 1). A $t$ test comparison of the AUC of the 2 ROC curves showed that the difference was not significant ($p = 0.96$). The cut-off value of the wMetaROI score was 1.38 in subjects with a lower reserve (sensitivity 80% and specificity 68%) and 1.37 in subjects with a higher reserve (sensitivity 80% and specificity 68%).

Taking into consideration only the extremes of educational attainment, the AUC of the ROC curve was 0.816
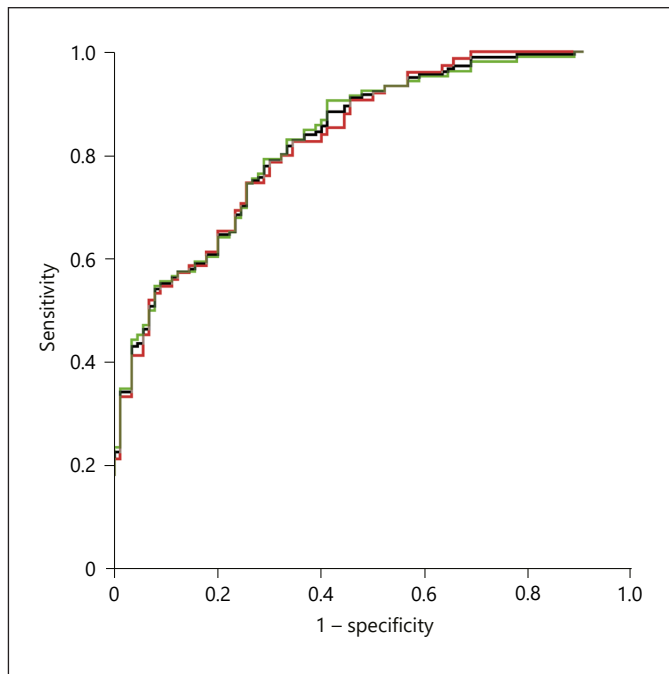
**Fig. 1.** Receiver operating characteristic curves of 18F-FDG-PET in the whole AD population (black) against AD patients with lower (<16 years) (red) and higher reserves (≥16 years) (green).
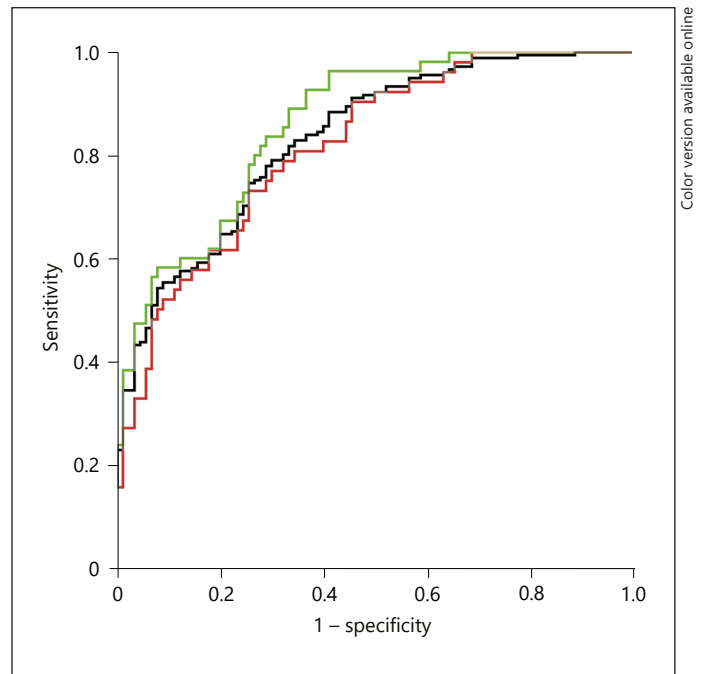
**Fig. 2.** Receiver operating characteristic curves of 18F-FDG-PET in the whole AD collective (black) against the subgroups of extreme educational attainment, i.e., ≤13 years for patients with a lower reserve (red) and ≥18 years for patients with a higher reserve (green).

when analyzing only LE AD patients compared to controls, and 0.861 in HE AD patients (Fig. 2). A *t* test comparison of the AUC of the 2 ROC curves showed that this difference was not significant ($p = 0.39$). The cut-off value of the wMetaROI score was 1.5 for LE subjects (sensitivity 81% and specificity 66%) and 1.2 for HE subjects (sensitivity 80% and specificity 73%).

No significant difference was found between the AUC of the ROC curves for the whole AD collectively and the LE subgroup ($p = 0.73$) or between the AUC for the whole AD population and the HE subgroup ($p = 0.49$).

When applying the standard cut-off calculated for the whole AD population (i.e., 1.38) in the LE subgroup instead of the education-adjusted threshold (i.e., 1.5), there was no significant difference in sensitivity ($p = 1$) or specificity ($p = 1$). The use of an education-adjusted threshold avoided 1 FN but implied 1 additional FP. When applying the standard cut-off in the HE subgroup, instead of the education-adjusted cut-off (i.e., 1.2), there was no significant change in sensitivity ($p = 0.25$) or specificity ($p = 0.063$); the use of an education-adjusted cut-off implied 3 FN and avoided 5 FP findings.

Finally, there was a trend for a higher sensitivity ($p = 0.063$) and no significant difference in specificity ($p = 1$) when using the education-adjusted threshold for LE for the whole population; this change in threshold induced 1 FP and avoided 5 FN.

When using the cut-off obtained for the HE subgroup in the whole population, there was a significant decrease in sensitivity ($p = 0.004$) with a limited gain in specificity which did not reach significance ($p = 0.063$), implying 9 FN and correcting 5 FP.

Table 4 summarizes these results. Table 5 reports the results for the wMetaROI scores corrected for the effect of MMSE, confirming that the diagnostic performance of the test does not differ on the basis of education level.

## Discussion

We specifically tested whether the diagnostic accuracy is different in subgroups with higher and lower reserve capacities, as measured by years of education, and the added value in terms of sensitivity and specificity of edu-

**Table 4.** Summary of wMetaROI scores

|  | AUC | $p$ value (two-tailed) | Threshold | Sensitivity, % | Specificity, % |
|---|---|---|---|---|---|
| Whole AD | 0.832 |  | 1.38 | 80 | 68 |
| LE AD <16 years | 0.831 | 0.96347 | 1.38 | 80 | 68 |
| HE AD ≥16 years | 0.833 |  | 1.37 | 80 | 68 |
| LE AD ≤13 years | 0.816 | 0.390672 | 1.5 | 81 | 66 |
| HE AD ≥18 years | 0.861 |  | 1.2 | 80 | 73 |

AUC, area under the curve; AD, Alzheimer disease; LE, lower education; HE, higher education.

**Table 5.** Summary of wMetaROI scores adjusted to the MMSE

|  | AUC | $p$ value (two-tailed) | Threshold | Sensitivity, % | Specificity, % |
|---|---|---|---|---|---|
| Whole AD | 0.844 |  | −0.403 | 80 | 70 |
| LE AD <16 years | 0.846 | 0.943239 | −0.405 | 80 | 70 |
| HE AD ≥16 years | 0.843 |  | −0.403 | 80 | 70 |
| LE AD ≤13 years | 0.834 | 0.476093 | −0.29 | 81 | 66 |
| HE AD ≥18 years | 0.870 |  | −0.52 | 80 | 74 |

MMSE, Mini Mental State Examination; AUC, area under the curve; AD, Alzheimer disease; LE, lower education; HE, higher education.

cation-adjusted thresholds with respect to a threshold defined over an "educationally heterogeneous" population.

FDG-PET is largely used as a disease and progression biomarker in AD. The evidence for the clinical validity of different biomarkers for dementia in AD, including FDG-PET, has been recently revised in a validation framework derived from oncology [11, 14–16]. One secondary aim of phase 2 of the validation process, which concerns the use of the assay in the clinical disease phase, is the assessment of factors associated with biomarker status or level in control subjects and in patients. Factors known to affect the biomarker should be tested for their impact on test positivity to define the need of having a factor-adjusted threshold in target subpopulations. While multiple previous studies have shown that education has an impact on the severity of metabolic impairment in AD subjects, this is the first study, to the best of our knowledge, to estimate the impact of the reserve phenomenon on the use of FDG-PET as diagnostic biomarker of AD in terms of diagnostic accuracy and normality threshold.

More educated individuals present at diagnosis with a more severe metabolic impairment the functional impact
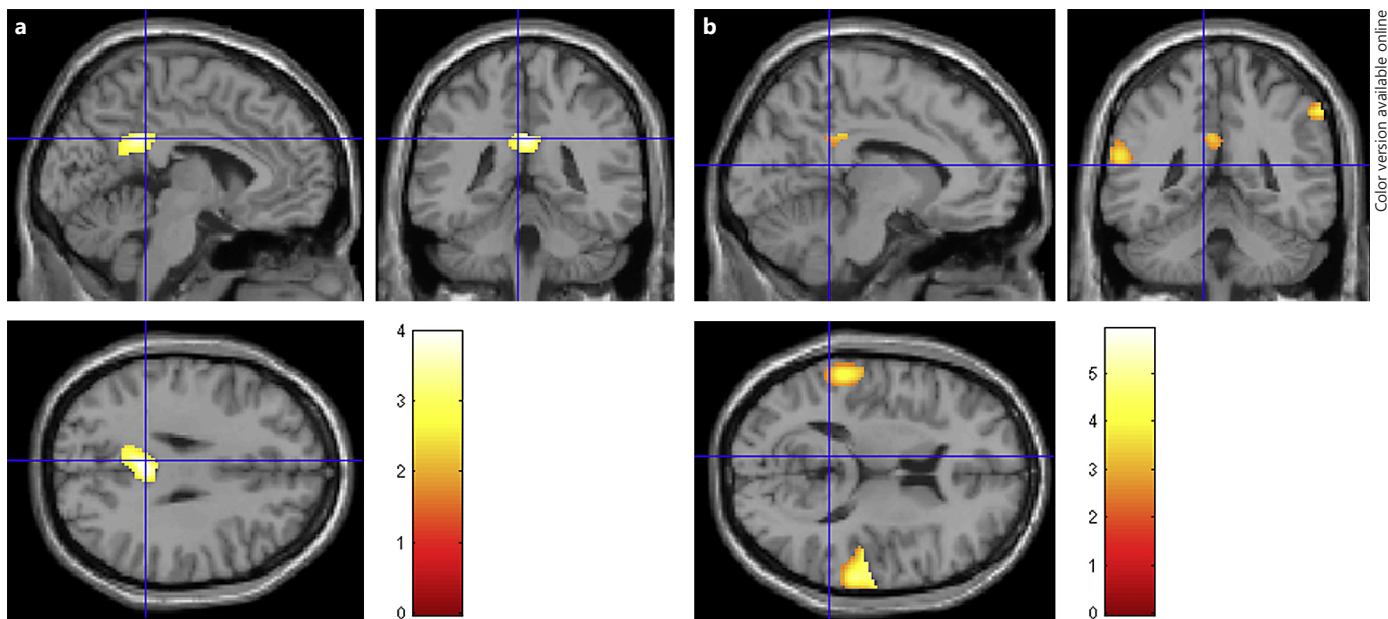
of which is compensated by the reserve capacity both in clinical [17, 18] and in preclinical AD [4, 7]. An individual example of the more extensive metabolism observed in patients with a higher reserve is provided in Figure 3.

Our results are in line with these findings, showing lower, albeit not significant, wMetaROI scores in patients with a higher reserve, despite a higher MMSE score, suggestive of relative compensation in a more advanced stage of the disease. Consistently, there was also an impact of reserve on the threshold to obtain 80% sensitivity, i.e., a less stringent threshold for LE subjects. It should be noted that the educational level of our sample was relatively high, and thus the estimates derived from this population might underestimate the impact of education with respect to a more representative memory clinic population.

However, this difference did not impact the diagnostic accuracy of the test, which was equally high in the whole population and in the LE and HE subgroups, which represent by definition the extremes of the distribution. In addition, the change in specificity was not statistically significant when the threshold derived from the general AD population was applied in the subgroups of LE and HE individuals. Overall, education was not a confounder in the diagnostic accuracy of FDG-PET in AD.

The set of 5 predefined ROI included in the metaROI targets the regions typically affected by AD but not necessarily the regions most touched by reserve phenomena. Differences between LE and HE have been consistently observed, occurring also in regions outside those considered by the metaROI. In a study of 242 patients and 144 controls, the authors found a significant impact of education and occupation on the metabolic activity in the posterior temporoparietal cortex, the precuneus, and the premotor/prefrontal cortex [7]. Other studies have also shown significant associations in the ventral prefrontal metabolism [4] and in the right dorsolateral prefrontal cortex [6].

The association between reserve and the proxy we used has some limitations. Indeed, years of formal education are able to capture only partly the phenomena associated with the reserve capacity, and multiple other proxies, e.g., occupational complexity and leisure activity, contribute independently to the development of reserve [19]. It is possible that in the LE population we misclassified some subjects having a higher reserve capacity provided by other factors not related to education. Ideally, composite measures taking into account, in addition to years of education, various factors such as measures of intelligence, nature and complexity of occupation, and intellectual activities during leisure time would serve best as surrogate

**Fig. 3.** SPM images in an Alzheimer disease patient with a lower reserve (**a**) and in a patient with a higher reserve (**b**). Note that there was a more important (in intensity and extent) decrease in metabolism in the subject categorized as having a higher reserve, while both had a Mini Mental State Examination score of 26/30.

measures of such a multifactorial concept. It should be noted, however, that partial validation of the use of years of education for this purpose comes from the repeated observation of an inverse relationship between reserve quantified by this particular surrogate measure and cerebral metabolism [4, 17, 18]. Our data are also consistent with these observations, measuring higher wMetaROI scores in individuals with LE. In addition, none of the more complex indices could be used in clinical practice for reserve-adjusted diagnostic thresholding. Although we hypothesize that our results are not specific to this proxy, further studies should be carried out on this subject.

There are possible caveats in our study when interpreting the results. We included normal controls without selection on the basis of amyloid status, using the strategy most commonly adopted to define normal reference datasets in imaging. We used a summary metric of metabolism in order to test the impact of reserve on an objective and robust index. Previous studies have shown that visual analyses or visual interpretation of voxelwise outputs might be superior to summary metrics for diagnostic purposes [20]. However, the impact of different levels of education on the visual interpretation of statistical maps would be difficult to estimate given the lack of an objective cut-off.

Finally, we studied subjects with dementia and did not include patients with prodromal AD, where deviations from normality are subtler. This population deserves a targeted investigation.

## Conclusion

In conclusion, the present study directly addressed for the first time the impact of the reserve phenomenon on the diagnostic accuracy of FDG-PET and the normality threshold. Although the effect of cognitive reserve was confirmed in the present group of highly educated patients, the use of an education-adjusted threshold did not significantly impact the diagnostic accuracy of FDG-PET, which was equally high in the whole population and in the LE and HE subgroups.

## Acknowledgement

## Statement of Ethics

This article does not contain any studies with human participants performed by any of the authors.

## Disclosure Statement

## Funding Sources

## References

1 Katzman R, Terry R, DeTeresa R, et al: Clinical, pathological, and neurochemical changes in dementia: a subgroup with preserved mental status and numerous neocortical plaques. Ann Neurol 1988;23:138–144.

2 Stern Y: Cognitive reserve and Alzheimer disease. Alzheimer Dis Assoc Disord 2006; 20:S69–S74.

3 Valenzuela MJ: Brain reserve and the prevention of dementia. Curr Opin Psychiatry 2008; 21:296–302.

4 Ewers M, Insel PS, Stern Y, Weiner MW: Cognitive reserve associated with FDG-PET in preclinical Alzheimer disease. Neurology 2013;80:1194–1201.

5 Fratiglioni L, Wang HX: Brain reserve hypothesis in dementia. J Alzheimers Dis 2007; 12:11–22.

6 Morbelli S, Perneczky R, Drzezga A, et al: Metabolic networks underlying cognitive reserve in prodromal Alzheimer disease: a European Alzheimer disease consortium project. J Nucl Med 2013;54:894–902.

7 Garibotto V, Borroni B, Kalbe E, et al: Education and occupation as proxies for reserve in aMCI converters and AD: FDG-PET evidence. Neurology 2008;71:1342–1349.

8 Perani D, Della Rosa PA, Cerami C, et al: Validation of an optimized SPM procedure for FDG-PET in dementia diagnosis in a clinical setting. Neuroimage Clin 2014;6:445–454.

9 Frisoni GB, Bocchetta M, Chetelat G, et al: Imaging markers for Alzheimer disease: which vs. how. Neurology 2013;81:487–500.

10 Morbelli S, Garibotto V, Van De Giessen E, et al: A Cochrane review on brain [18F]FDG PET in dementia: limitations and future perspectives. Eur J Nucl Med Mol Imaging 2015;42: 1487–1491.

11 Garibotto V, Herholz K, Boccardi M, et al: Clinical validity of brain fluorodeoxyglucose positron emission tomography as a biomarker for Alzheimer's disease in the context of a structured 5-phase development framework. Neurobiol Aging 2017;52:183–195.

12 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.

13 Caroli A, Prestia A, Chen K, et al: Summary metrics to assess Alzheimer disease-related hypometabolic pattern with 18F-FDG PET: head-to-head comparison. J Nucl Med 2012; 53:592–600.

14 Boccardi M, Gallo V, Yasui Y, et al: The biomarker-based diagnosis of Alzheimer's disease. 2. Lessons from oncology. Neurobiol Aging 2017;52:141–152.

15 Frisoni GB, Boccardi M, Barkhof F, et al: Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. Lancet Neurol 2017;16:661–676.

16 Frisoni GB, Perani D, Bastianello S, et al: Biomarkers for the diagnosis of Alzheimer's disease in clinical practice: an Italian intersocietal roadmap. Neurobiol Aging 2017;52:119–131.

17 Morbelli S, Nobili F: Cognitive reserve and clinical expression of Alzheimer's disease: evidence and implications for brain PET imaging. Am J Nucl Med Mol Imaging 2014;4:239–247.

18 Perneczky R, Drzezga A, Diehl-Schmid J, et al: Schooling mediates brain reserve in Alzheimer's disease: findings of fluoro-deoxy-glucose-positron emission tomography. J Neurol Neurosurg Psychiatry 2006;77:1060–1063.

19 Foubert-Samier A, Catheline G, Amieva H, et al: Education, occupation, leisure activities, and brain reserve: a population-based study. Neurobiol Aging 2012;33:423 e415–e425.

20 Morbelli S, Brugnolo A, Bossert I, et al: Visual versus semi-quantitative analysis of 18F-FDG-PET in amnestic MCI: an European Alzheimer's Disease Consortium (EADC) project. J Alzheimers Dis 2015;44:815–826.