Standardisation of Quantitative PET Measures in Alzheimer's Disease Using the Centiloid Scale

Author: Marc Ballestero Ribó*
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Roser Sala-Llonch and Agnès Pérez-Millan

Abstract: Amyloid-beta (Aβ) positron emission tomography (PET) is an established neuroimaging technique for the diagnosis of Alzheimer's disease (AD). Quantitative PET measures can be heavily influenced by particular methodological choices. The Centiloid (CL) scale has been proposed as a metric to standardise these measurements. Nonetheless, the calibration to this scale requires the acquisition of PET scans using the tracer [^{11}C]-Pittsburgh Compound-B (PiB), which is not widely available. The aim of this work is to process and quantify a set of 96 Aβ PET images and test whether a conversion to the CL scale can be done without using PiB acquisitions. Using preprocessed [^{18}F]-Florbetapir (FBP) PET images from the ADNI dataset, we performed quantification using a software tool from Siemens. Results were converted to CL using an externally calibrated formula. These values were compared with standard CL values, also available in the dataset. A high degree of linear correlation was found between the estimated CL values and the standard ones. Nonetheless, the linear regression showed a considerable amount of associated uncertainty. As a conclusion, both methods are comparable in measuring brain pathology. However, further work should explore the origin of uncertainty in the conversion, which should be corrected.

I. INTRODUCTION

Alzheimer's disease (AD) is the most common form of dementia. An estimated 60% to 80% of the dementia cases are caused by AD. It is a neurodegenerative disease characterized by the formation of neuritic plaques as a result of amyloid-beta (Aβ) peptide's accumulation in the brain. Mild cognitive impairment (MCI) represents an early stage of cognitive loss, which might eventually progress to AD.

Neuroimaging techniques are helpful in the diagnosis of AD and MCI. In particular, Aβ positron emission tomography (PET) has a relevant role in AD diagnosis. It is a nuclear imaging technique which involves the injection to the patient of a pharmaceutical compound, called *tracer*, which is formed by a ligand that binds to Aβ plaques labelled with a positron emitter. The positrons released by the tracer interact with the electrons in the patient's cerebral tissue, leading to the emission of a pair of γ rays, which can be detected using a PET scanner. This information is posteriorly reconstructed into a tomographic image. Several tracers are available for Aβ imaging. The most widely used ones are based either on ^{11}C or ^{18}F, such as [^{11}C]-Pittsburgh Compound-B (PiB) and [^{18}F]-Florbetapir (FBP).

Traditionally, PET results were given by visual inspection, which classifies PET scans as either *positive* or *negative*. More recently, quantitative approaches have been developed to allow obtaining regional measurements that fall in a continuous spectrum of values. Quantification is the process that is followed to obtain a numerical measure of the presence of an Aβ load in a patient's brain from an Aβ PET image. Quantitative metrics may be affected by the variability introduced through different methodological choices, such as acquisition and reconstruction technique, the choice of the region of interest (ROI) and reference region (REF) and the Aβ tracer used. The lack of a standard, method-independent, metric for amyloid PET imaging has led to (1) an overly wide range of measures associated with subjects showing a negative Aβ load, (2) the absence of a clear threshold for Aβ positivity, and (3) the inability to compare studies across different Aβ quantification methods [1].

Attempts of developing a standard scale have been made, the most relevant being the development of the *centiloid* (CL) scale, by Klunk *et al.* [1]. Nonetheless, the calibration of a particular method to this scale requires the acquisition of PET scans using PiB as the tracer, which is not widely available in all centres performing Aβ PET imaging. Hence, in such locations, a standard calibration of quantitative measures to the CL scale is not possible.

The aim of this work is to assess and validate a calibration process to the CL scale which does not rely on the acquisition of [^{11}C]-PiB scans, and to evaluate the results in a sample obtained from a publicly available database.

II. BACKGROUND

A. Aβ Burden Quantification

One of the most established and widely used metrics for quantifying brain's Aβ load is the *standardised uptake value ratio* (SUVr), which relates the tracer uptake between the ROI and REF regions in a PET acquisition.

*Electronic address: mballeri7@alumnes.ub.edu

For a given ROI and REF, the SUVr value is given by

$$\text{SUVr} = \alpha_{\text{ROI}} \,/\, \alpha_{\text{REF}}, \tag{1}$$

where α_{X} is the activity concentration of the tracer radionuclide at the region X, that is $\alpha_X = A_{\text{X}} \,/\, V_{\text{X}}$, being A_{X} the activity in the region and V_{X} its volume. Activity concentrations and volumes can be estimated by means of processing the reconstructed PET images.

B. The Centiloid Scale

Klunk *et al.* [1] developed and introduced the CL scale, which standardises quantitative Aβ PET measurements by (1) establishing a standard analysis pipeline for the obtention of cortical SUVr values and (2) scaling the outcome of each particular method to a standard 0 to 100 scale, being 0 an anchor point to a "highly certain" negative Aβ load and 100 a point associated with a "typical" positive Aβ load.

When aiming to implement the CL scale to a local sample, there are three levels of processing. For this project, we will focus on the first two, which are described below.

1. *Level-1: Setting the Anchor Points*

Level-1 CL processing is defined by the authors of the method and was meant to be done once, and hence shall not be repeated by a research group aiming to use the CL scale. However, we describe the formulas used in order to facilitate the part of the procedure that will be implemented locally. This processing level comprises the selection of subjects to define the anchor points of the CL scale, and the establishment of a normalised method for quantification, called the *standard PiB method*.

Klunk *et al.* [1] chose three groups of subjects for the development of the level-1 process: the young healthy control 0-anchor group (YC-0), the AD 100-anchor group (AD-100) and the global cortical target group (CTX).

PET imaging was performed for each participant, using PiB as the PET tracer. The whole cerebellum (WC) was chosen as the REF region for quantification. The global cortical target ROI (CTX ROI) was segmented based on the quantification of images from the CTX group, and comprises brain regions typically showing high Aβ loads in patients with AD.

SUVr values for the CTX ROI were calculated for each participant. The mean SUVr of the YC-0 subjects, ${}^{\text{PiB}}\overline{\text{SUVr}}_{\text{YC-0}} = 1.009$, was set to be 0 CL, and the mean SUVr of AD-100, ${}^{\text{PiB}}\overline{\text{SUVr}}_{\text{AD-100}} = 2.076$, was anchored to 100 CL.

Any SUVr value obtained following the previously detailed standard PiB method can be converted to CL using a linear transformation as

$$\text{CL} = 100 \times \frac{{}^{\text{PiB}}\text{SUVr} - {}^{\text{PiB}}\overline{\text{SUVr}}_{\text{YC-0}}}{{}^{\text{PiB}}\overline{\text{SUVr}}_{\text{AD-100}} - {}^{\text{PiB}}\overline{\text{SUVr}}_{\text{YC-0}}}, \tag{2}$$

which can be numerically expressed as

$$\text{CL} = 100 \times \frac{{}^{\text{PiB}}\text{SUVr} - 1.009}{1.067}. \tag{3}$$

All the anonymised images used for this derivation are publicly available.

2. *Level-2: Calibrating a Non-Standard Method*

The second level of CL processing is the procedure that should be used to calibrate any specific method different from the standard PiB method to the CL scale. This allows the usage of tracers different from PiB, different acquisition times and analysis methods, and other ROI and REF regions.

This process can be divided into two steps.

(1) **Local validation of the level-1 process:** The first step that must be done is replicating the level-1 analysis using the publicly available data to validate the local pipeline. The mean SUVr obtained from this replication for the YC-0 and AD-100 groups should not fall out of a 2% interval around the mean SUVr that was obtained in the level-1 process, published in the original paper. These mean values obtained during the level-2 analysis should then be used to calibrate a new centiloid scale, denoted by CL_{L2}, using a linear transformation analogous to equation (3). Posteriorly, CL_{L2} values shall be estimated for all the subjects. A linear regression should then be done between these CL_{L2} values and the CL values obtained using the standard PiB method and reported in [1]. The slope of the regression line should be between 0.98 and 1.02, the intercept between -2 and 2 CL and the correlation coefficient $R^2 > 0.98$.

(2) **Calibration of the new method:** The centre must locally acquire at least 25 PET scans, both with PiB ($50-70$ min after injection) and the tracer used locally (LTR). Using the standard CTX ROI, ${}^{\text{PiB}}\text{SUVr}$ values must be estimated for the PiB scans. Then, using the local ROI and REF regions, SUVr values should be calculated for the LTR scans, namely ${}^{\text{LTR}}\text{SUVr}$. A linear regression will be performed between the ${}^{\text{LTR}}\text{SUVr}$ and ${}^{\text{PiB}}\text{SUVr}$ values. A correlation coefficient $R^2 > 0.70$ is considered acceptable. The coefficients of the linear regression are then used to convert locally acquired ${}^{\text{LTR}}\text{SUVr}$ values to ${}^{\text{PiB}}\text{SUVr}$ which can be further converted to CL using equation (3).

III. METHODS

A. Data Selection and Preprocessing

The data used in this work to implement the CL standardisation described above was obtained from the

Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset (adni.loni.usc.edu). The ADNI is a multicentric study, designed to develop multimodal biomarkers for the study of AD.

A total of 96 participants of the ADNI dataset were selected for this study, distributed into four groups: healthy controls (CN, $N = 24$, healthy patients), stable MCI (sMCI, $N = 24$, participants diagnosed with MCI during all their participation in the ADNI study), converter MCI (cMCI, $N = 24$, participants with an initial diagnosis of MCI who posteriorly progressed to AD) and AD ($N = 24$, participants with an AD diagnosis). The selection was done ensuring equal distributions of participants in each group, ensuring equal sex distributions and that the age of participants was not out of the interquartile range of the age distribution within their group.

Preprocessed FBP PET images (corregistered dynamic, averaged, standardised image and voxel size, uniform resolution) were downloaded from the ADNI dataset. This preprocessing was done to uniformise the data and eliminate all the differences in the images due to the use of different acquisition systems. Notably, this preprocessing was performed and validated by ADNI researchers. Thus, in terms of CL standardisation, we can consider all the 96 PET images as if they had been acquired in the same centre.

Moreover, demographic data (age, sex, visit dates), the diagnostic of each participant at the time of the selected PET scan, and the CL value derived from the quantification of each scan done following the level-2 CL pipeline, also performed by ADNI researchers [2], were downloaded from the ADNI. We will refer to these level-2 CL values as $\mathrm{CL_A}$ and use them as a standard to test our method.

B. Quantification of PET Scans

Siemens *Syngo.PET*® (hereinafter *Syngo*), available at the Nuclear Medicine Service of the Hospital Clínic de Barcelona (HCB), was used to perform Aβ quantification of the PET images [3]. The *Syngo* software pipeline first involves the affine registration of the PET image to a standard template and spatial normalization. An automatic parcellation of six cortical gray matter ROIs considered relevant for quantification of FBP uptake (Medial Orbital Frontal, Temporal, Anterior Cingulate, Posterior Cingulate and Precuneus) and the WC is performed. Finally, a SUVr analysis using WC as REF is done for each ROI. The SUVr of the composite region of the six ROIs is also reported, and can be interpreted as a measure of the global mean cortical SUVr [4]. We will refer to this quantification method as the *local Syngo method*.

C. Conversion to the Centiloid Scale

The *Syngo* software package does not allow for a proper level-2 calibration of the local Aβ quantification method to the CL scale, as it is uncapable of replicating the standard PiB method. Moreover, no PiB scans were downloaded from ADNI, to replicate the situation where it is not possible to acquire them. As an alternative, an external calibration formula, reported by Navitski *et al.* [5], was used.

Navitski *et al.* [5] implemented the level-2 CL analysis for the usage of FBP as a tracer and the same six ROIs used by *Syngo*. This method has essential similarities with our *Syngo* method: the same tracer is used, and the choice of ROIs and REF is the same. Differences between methods can be mainly attributed to the usage of different preprocessing pipelines and software packages. The analysis of Navitski *et al.* [5] led to a SUVr conversion equation of the form

$$^{\mathrm{FBP}}\mathrm{SUVr} = 0.51 \times {}^{\mathrm{PiB}}\mathrm{SUVr} + 0.45 \tag{4}$$

which, combined with equation (3), can be used to express $^{\mathrm{FBP}}$SUVr values in the CL scale, resulting in

$$\mathrm{CL} = 183 \times {}^{\mathrm{FBP}}\mathrm{SUVr} - 177. \tag{5}$$

For the local *Syngo* method, CL values were estimated with the previous equation using the composite cortical region SUVr. We denote these values as $\mathrm{CL_S}$, in contrast to the $\mathrm{CL_A}$ values reported in the ADNI databases.

To compare both measures, a linear regression analysis between the $\mathrm{CL_A}$ and $\mathrm{CL_S}$ values was performed, through a linear fit equation of the form

$$\mathrm{CL_S} = a \times \mathrm{CL_A} + b. \tag{6}$$

A 95% confidence interval and a 95% prediction interval were estimated for the assessment of the uncertainty in the predictions.

D. The Aβ Level Threshold

The results reported by Navistky *et al.* [5], beyond providing the calibration equation (5), also included a CL threshold for Aβ positivity, which was derived by comparing SUVr obtained from an independent group formed by participants with a clinically verified diagnosis. This threshold was reported to be situated at 24.1 CL.

We used this value with the conversion equation derived from our linear regression (6) to derive an Aβ positivity threshold for our estimated $\mathrm{CL_S}$ values.

E. Statistical Analysis

Differences in SUVr values and $\mathrm{CL_S}$ values between groups were studied using an ANOVA test, corrected post-hoc for multiple comparisons using a Tukey HSD test. All the p-values reported in this work are the corrected values for multiple comparisons. A statistical significance threshold of $p < 0.05$ was adopted.

IV. RESULTS

A. Dataset

Sample demographics are shown in table I. Due to selection criteria, no participants with a visually assessed positive Aβ level are present in the CN group and, conversely, all the participants in the AD group are Aβ positive.

Table I: Group summaries of the study cohort. Age (mean ± standard deviation) and gender data are reported, along CL_A (mean ± standard deviation) values and the number of Aβ positive and Aβ negative participants.

	Group			
Parameter	CN	sMCI	cMCI	AD
N	24	24	24	24
Age (years)	71 ± 2	72 ± 3	73 ± 3	74 ± 3
Sex (M/F)	12/12	12/12	12/12	12/12
CL_A (CL)	−1 ± 11	60 ± 60	80 ± 50	90 ± 40
Aβ level (+/−)	0/24	17/7	20/4	24/0

B. Quantification and SUVr Analysis

The processing and quantification of the selected FBP PET images using *Syngo* resulted in the obtention of the SUVr values for the composite cortical region. Their distribution within each group is displayed in figure 1. Statistically significant differences were found between CN and the rest of the groups, all with $p < 0.001$. No other statistically significant differences were identified.

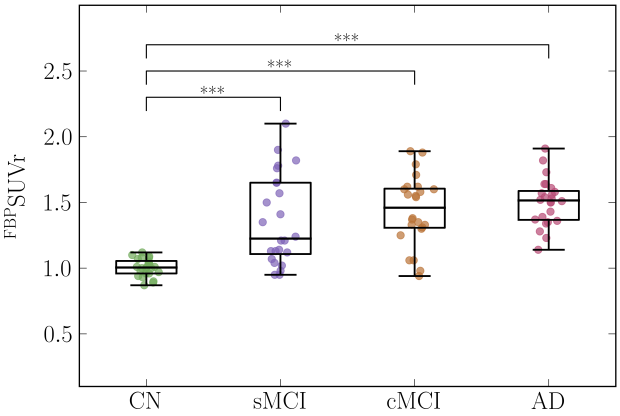

Figure 1: SUVr values of the composite cortical region estimated with *Syngo* for each group. *** are used to indicate statistically significant differences with $p < 0.001$.

C. Centiloid Analysis

The SUVr values for the composite cortical region estimated with *Syngo* were converted to CL using equation (5), as detailed in methods section C, to obtain the so-denoted CL_S values. A linear regression of the form of equation (6) was done between the CL_S and CL_A values. The resulting parameters of the regression are shown in table II.

Table II: Parameters of the linear regression (6).

a (−)	1.00 ± 0.02
b (CL)	11 ± 2
R^2	0.96
$(\delta CL_S)_{reg}$ (CL)	11

The slope of the regression line, a, is equal to 1 with very low uncertainty, whereas the value of the intersection, b, is substantial, as it is of the same order of magnitude as the CL_S values. Note also that the standard error of the predicted values in the regression is of the same order of magnitude as the CL_S values, $(\delta CL_S)_{reg} = 11$ CL. A 95% prediction interval is shown in the plot of figure 2 to illustrate this substantial uncertainty.

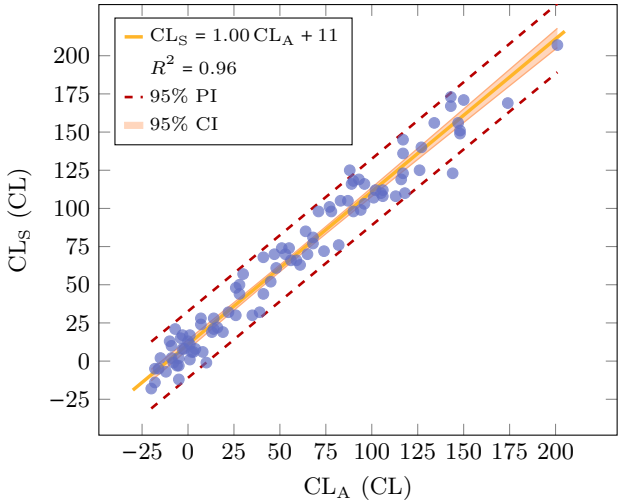

Figure 2: CL_S *vs.* CL_A values, with the linear regression line, the 95% confidence interval (CI) and the 95% prediction interval (PI).

Figure 4 shows the distribution of the CL_S values for each group. We found statistically significant differences ($p < 0.001$) between the CN and all the other groups. No other significant differences were detected.

D. The Aβ Positivity Threshold

The threshold for Aβ positivity in CL_A values was reported to be 24.1 CL by Navistky *et al.* [5].

This threshold can be converted using the linear regression equation (6) to a threshold of Aβ positivity for the CL_S values, as shown in figure 3. This converted threshold was estimated to be (35 ± 22) CL, with the uncertainty given by the 95% prediction interval for the regression. A representation of the group distribution of the CL_S values along this threshold is plotted in figure 4.

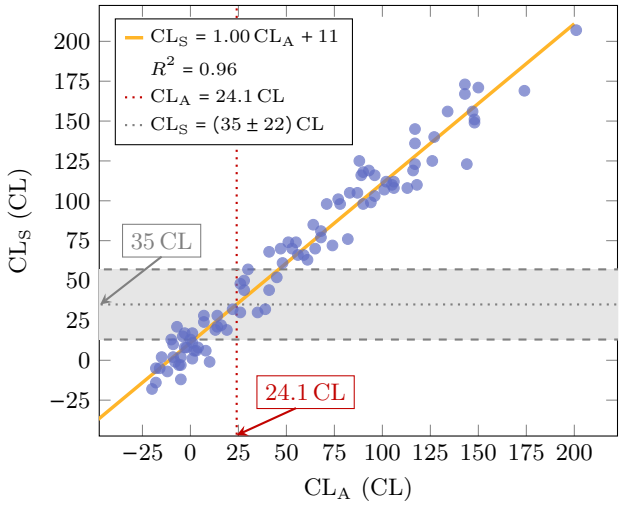

Figure 3: Aβ positivity thresholds for the CL_A and CL_S values, with the corresponding uncertainty range.

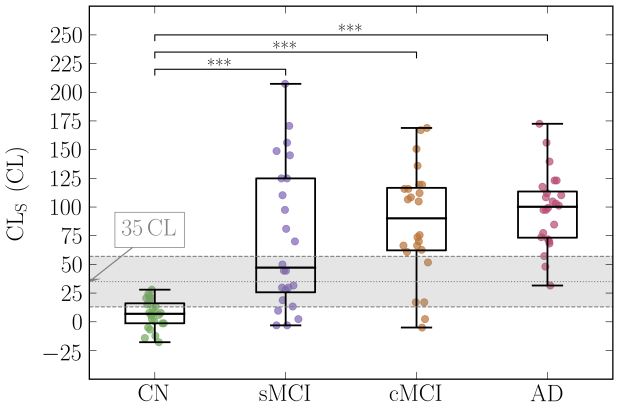

Figure 4: Group distribution of the CL_S values, along the Aβ positivity threshold estimated for this scale. *** are used to indicate statistically significant differences with $p < 0.001$.

V. DISCUSSION

In this work, we implemented a method to estimate CL values by means of converting the local *Syngo* method SUVr values with an external calibration equation. We obtained values that are highly linearly correlated with the CL_A values reported in ADNI and obtained performing a standard level-2 CL process. This linear correlation can be used as a correction equation to convert CL_S to CL_A values. The slope of the linear regression is 1.00 ± 0.02, showing a very low amount of relative uncertainty. Thus, this correction can be implemented by subtracting the bias introduced by the intersection of the regression, $b = (11 \pm 2)\,CL$. This bias can be mainly attributed to the error resulting from the use of a different preprocessing pipeline and software methodology by the local *Syngo* method. Nonetheless, the uncertainty of the values estimated using this conversion equation has the same order of magnitude as the corrected values, since $(\delta CL_S)_{reg} = 11\,CL$, which introduces a considerable error to the correction. This error impacts the derivation of an Aβ positivity threshold for the CL_S values, widening its possible range of values to a point where it may potentially fail in discriminating Aβ positive from Aβ negative participants.

VI. CONCLUSIONS

Here, we explored the possibility of calculating the CL values without relying on a set of PiB scans. The obtained values show a high degree of linear correlation with CL values derived using a standard and validated process. Nonetheless, the considerable amount of uncertainty associated with a conversion equation between the two sets of values does not allow the derivation of a faithful Aβ positivity threshold. Our results indicate that both methods are equivalent in the sense that they can measure the degree of brain pathology in these subjects. However, further experiments might be done to investigate the origin of the uncertainty that appears and to mitigate it.

Acknowledgments

I would like to express my deepest gratitude to my advisors, Dr. Roser Sala-Llonch and Dr. Agnès Pérez-Millan, for their advice, dedication, and help during the development of this work.

I am also grateful to my parents, my partner, and my friends for their encouragement and support all through my studies.

[1] Klunk, W. E. et al. "The Centiloid Project: Standardizing Quantitative Amyloid Plaque Estimation by PET". Alzheimers Dement. **11**(1)-15.e1-4 (2014)

[2] Lee, J. et al. "ADNI Amyloid PET Processing Methods [White Paper]". UC Berkeley and Lawrence Berkeley National Laboratory (2023)

[3] Fleisher, A. S. et al. "Using Positron Emission Tomography and Florbetapir F18 to Image Cortical Amyloid in Patients With Mild Cognitive Impairment or Dementia due to Alzheimer's Disease". Archives of Neurology **68**(11):1404-1411, 11 (2011)

[4] Ghosh, P. et al. "Quantitative Software Evaluation of Beta-Amyloid Brain PET Imaging in Dementia [White Paper]". Siemens Medical Solutions (2012).

[5] Navistky, M. et al. "Standardization of Amyloid Quantification With Florbetapir Standardized Uptake Value Ratios to the Centiloid Scale". Alzheimer's & Dementia **14**(12):1565-1571 (2018)