EMA Scientific Advice – Qualification

Procedure No.: To be advised

Centiloid measure of Amyloid PET to quantify brain amyloid deposition

115952 – AMYPAD
Amyloid imaging to Prevent Alzheimer's Disease

Qualification Dossier

Submitted to the European Medicines Agency
September 2022

Organisation	Synapse SME and AMYPAD consortium
Document	CHMP Qualification Opinion
Procedure No.:	To be advised by EMA
Level of Dissemination	Confidential

Report date	

Table of Contents

LIST OF TABLES

LIST OF FIGURES

List of Abbreviations (to be expanded by AMYPAD)

AD	Alzheimer's Disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
AMYPAD	Amyloid Imaging to Prevent Alzheimer's Disease
ANOVA	Analysis of variance
BPND	non displaceable Binding Potential
BQO	Biomarker Qualification Opinion
BQRT	Biomarker Qualification Review Team
CERAD	Consortium to Establish a Registry for Alzheimer's Disease
CHMP	Committee for Medical Products for Human Use
CI	Confidence Interval
CID	Clinically Important Difference
CL	Centiloid
CSF	Cerebral Spinal Fluid
CT	Computerize Tomography
DLB	Dementia with Lewy Bodies
DPMS	Diagnostic and Patient Management Study
EMA	European Medicines Agency
EPAD	European Prevention of Alzheimer's Dementia
FBB	Florbetaben
FDA	Food and Drug Administration
FMM	Flutemetamol
GAAIN	Global Alzheimer's Association Interactive Network
GCA	Global Cortical Average
Ha	Alternative hypothesis
CoU	Context-of-Use
Ho	Null hypothesis
IDEAS	Imaging Dementia – Evidence for Amyloid Scanning
LCS	Longitudinal Cohort Study

LoI	Letter of Intent
LoS	Letter of Support
MCI	Mild Cognitive Impairment
MMSE	Mini Mental State Examination
MoCA	Montreal Cognitive Assessment
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
PC	Parent Cohort
PET	Positron Emission Tomography
PI	Prescribing Information
PNHS	Prognostic and Natural History Study
p-value	Probability (of obtaining a result equal to or "more extreme" than what was actually observed, when the null hypothesis is true)
QC	Quality Control
RCT	Randomized Controlled Trial
ROI	Region of Interest
SAP	Statistical Analysis Plan
SCD+	Subjective Cognitive Decline Plus
SPECT	Single-Photon Emission Computed Tomography
SPM	Statistical Parametric Mapping
SUVr	Standard Uptake Value Ratio
US	United States
VOI	Volumes of Interest
WM	White Matter

1 EXECUTIVE SUMMARY

Biomarker Qualification Opinion (BQO) for the Centiloid Measure as a universal metric for the assessment of brain amyloid burden:

A robust standardised tracer-independent methodology for measuring amyloid load in subjects with early or established pathology.

Applicant: IMI funded Amyloid Imaging to Prevent Alzheimer's Disease (AMYPAD) consortium

Biomarker Qualification Opinion (BQO) Process:

Developed by the European Medicines Agency (EMA) to facilitate the acceptability of specific use for a novel method or imaging modality to enable progress in the development of novel treatment and management regimes. The opinion process involves the assessment of submitted data and additionally a further public consultation with the scientific community. The process starts with the applicant submitting a detailed proposal and is projected to take approximately 9 months to 1 year.

AMYPAD Program:

AMYPAD is a public-private partnership of 15 European partners who have two active clinical programs in the field of brain amyloid positron emission tomography (PET) imaging, with the ultimate goal to improve knowledge of dementia pathology and clinical progression. One major objective is the development and validation of robust standardised methodology for the measurement of amyloid in the brain. The project is now in its 6th year of funding and was granted a no-cost extension till the 30th of September 2022. The F-18 tracers Vizamyl ([^{18}F]flutemetamol) and Neuraceq ([^{18}F]florbetaben) are approved by EMA and broadly available in Europe, both are being studied in the AMYPAD program. Additional data from the US based IDEAS study which has a large proportion of Amyvid ([^{18}F]florbetapir) scans has also been included in the results section.

What currently exists:

Fluorine-18 labelled brain amyloid PET tracers have been available for routine use in Europe since 2013 and have been validated against Consortium to Establish a Registry for Alzheimer's disease (CERAD) pathology as the standard of truth. Clinical routine use of brain amyloid PET tracers involves categorisation of static scans by visual read as either negative or positive. All three amyloid PET tracers approved in the EU have quantification included in their SmPC as an adjunct to a visual read to assist in the assessment of an amyloid PET scan.

Additionally in the research space, quantitative measures are being employed with many of the standard software packages able to calculate both regional and composite levels of amyloid burden, enabling a continuous measure of amyloid load in addition to the dichotomous read that the visual inspection allows. Methods such as the standardised uptake value ratio (SUVr) yield tracer uptake values, which vary depending upon the chosen reference region and the analytical implementation. In turn, non-displaceable Binding Potential (BP_{ND}) reflects specific tracer uptake, as it takes several technical and physiological factors into account and is therefore considered a more accurate and precise measure. This measure, however, faces a similar dependency on radiotracer and analytical approaches, and requires a longer "dynamic" acquisition protocol, which may limit routine clinical use.

What is the knowledge gap:

Recently, the field of Alzheimer's disease (AD) research has focused on the value of both the topographical distribution and burden of amyloid pathology present, rather than a binary classification of the amyloid status. Studies so far have illustrated the added value of this information for both disease-modifying therapies and clinical use. There is a need to reliably quantify the presence of early amyloid pathology as secondary prevention trials move to treat subjects with low but detectable levels of amyloid. Additionally, there is value to improve the prognostic value of amyloid imaging in clinical routine, by considering the overall pathological load, which could improve subject placement along the AD trajectory. Although controversial regarding the clinical benefit demonstrated so far, the recent Aduhelm approval by the Food and Drug Administration (FDA) also highlights the potential value of a universal metric to assess the amyloid burden by PET as the label was updated in April 2022 to include the following 'confirm the presence of amyloid beta pathology prior to initiating treatment'. This could include both a baseline measure of amyloid to initiate treatment and potentially further scans for the purposes of managing the therapy regime. In addition to Aduhelm, other promising anti-amyloid therapies are in the final stages of closing out Phase III studies and submitting NDA/MAAs in both the USA and Europe. Managing both the inclusion into therapy as well as therapy monitoring across both global territories and with multiple tracers will require a consistent and robust approach.

One method increasingly gaining traction in the dementia neuroimaging space is the Centiloid measure, a tracer independent metric that can be easily grasped beyond Nuclear Medicine as well as providing thresholds to answer different questions. Thus, while visual binary read of global amyloid provides useful information for clinical routine and research purposes, it does not consider the

wealth of information that brain PET scans provide, both from a regional and continuous quantitative measure perspective.

Premise of BQO:

To facilitate the wider utility of standardized, tracer independent, and sensitive methods for 1) measuring cross-sectional levels (and potentially longitudinal changes) of brain amyloid pathology across PET tracers and 2) support amyloid PET biomarker use in both clinical routine and research by providing information on the extent of pathology for differing scenarios. These could include the evaluation of both early and established amyloid pathology as well as the possibility to predict disease trajectory (i.e., prognosis). Currently, the Centiloid measure could be considered the most developed quantitative methodology within the field of amyloid PET and has been reliably implemented in multiple studies, including AMYPAD. Other quantitative methods to optimally measure amyloid burden or accumulation have been proposed, such as Aβ load and Aβ index. However, these approaches are currently less mature, having only been assessed in limited data sets.

Sources of data:

The primary sources of data presented in this BQO is amyloid measures from the two AMYPAD studies (i.e., the Diagnostic and Patient Management study, DPMS; and the Prognostic Natural History Study, PNHS). Additionally, work has been performed by members of the AMYPAD consortium on other cohorts (e.g. ALFA+, ABIDE, IDEAS etc) and will be appropriately referenced. There has also been a large body of data published in the literature and/or presented at recent conferences and this too is considered in this application.

Analysis proposed in this BQO:

A wealth of data and analysis primarily from both the AMYPAD DPMS and PNHS studies are presented in this BQO dossier. The analysis described in Chapter 6 is broadly divided into three sections which cover a) analytical robustness of the quantitation of cortical amyloid, b) cross sectional results of image analysis in the clinical subgroups of DPMS and other studies and c) the longitudinal analysis of amyloid PET in both DPMS and PNHS.

Value to the field of AD:

To provide a framework for the validation of quantitative assessment of amyloid burden, which is suitable for use/implementation by the general dementia field. The Centiloid method is the example for this BQO. The approach by AMYPAD

has been endorsed by the European Association of Nuclear Medicine (EANM) (see letter of endorsement in Appendix A).

The application could also provide a template for further methodologies to be introduced as well, as future uses of amyloid PET are expected, e.g., more widespread applicability of longitudinal scanning to monitor therapeutic efficacy of cases with developing pathology.

The use of the Centiloid method allows the dementia field to use a central, universal metric, which is valid across all three approved brain amyloid PET tracers. This method aligns the use of target and reference regions and harmonizes the outcome measures.

Additionally, the BQO will demonstrate best practice PET acquisition parameters for the acquisition and reconstruction of amyloid PET images gained via collection and analysis of over 2000 images acquired in the AMYPAD program (either newly acquired for prospective AMYPAD studies or in collaboration with other consortia).

2 REGULATORY HISTORY

Summary of Previous Regulatory Interactions

Table 1 Summary of AMYPAD Regulatory Interactions with EMA

Date	Event	Summary
10th Nov 2016	EMA Scientific Advice Procedure No.: EMEA/H/SAH/070/1/2016/SME/III	Advice on design and endpoints of Diagnostic and Patient Management Study (DPMS)
28th Feb 2019	EMA Scientific Advice Procedure No.: EMEA/H/SA/4003/1/FU/1/2019/SME/II	Discussion and advice on concepts (including the Centiloid measure) for the quantitation of brain amyloid burden using amyloid PET tracers Q4 from this EMA SA provided an initial opportunity to discuss the CL concept and provided baseline guidance for this Qualification Opinion

2.1 Previous Biomarker Qualification Opinions in the Neuroimaging Space

The EMA issued in 2018 a full Qualification Opinion for the use of dopamine transporter imaging as an enrichment biomarker in Parkinson disease trials targeting subjects with early motor symptoms. Exclusion of subjects without evidence of dopaminergic deficit (SWEDDs) in future clinical trials targeting early motor PD subjects aimed to enrich clinical trial populations with idiopathic PD patients, improve statistical power, and exclude subjects who are unlikely to progress clinically from being exposed to novel test therapeutics.

The structure and content of this current dossier for our AMYPAD submission of 'Centiloid measure of Amyloid PET to quantify early and established amyloid deposition' is based upon the template provided by the DAT imaging document (Stephenson et al 2019)

3 PROPOSED CONTEXT OF USE STATEMENT

Context of Use: Centiloid measure as a universal metric for the assessment of brain amyloid burden.

- **General Area**

Robust standardised tracer-independent methodology for measuring amyloid load in subjects with early or established pathology.

- **Target Population for Use**

Subjects on the AD pathology continuum where knowledge of the amyloid load specifically leads to a) diagnostic decisions such as therapeutic intervention/monitoring or b) inclusion or exclusion from clinical research.

- **Stage of Implementation**

18F-labelled Amyloid PET Tracers are available for routine use where the etiological assessment of amyloid burden is required, but CL methodology as an adjunct to visual inspection is not yet embedded in standard of care/image analysis. Currently the technology is widely used in clinical trials across the full AD continuum. Preparation of the clinical field for newly approved therapeutics where a measure of amyloid burden is required to ensure consistent and reliable patient outcomes.

- **Intended Application**

There are a number of possible applications of the Centiloid measure which can be envisaged:

- Selection of individuals with established pathology who are targets for novel (in research) and future approved amyloid targeted therapies.
- Optimal therapy monitoring (including treatment decisions/prognosis from baseline etc).
- Identification of subjects with preclinical AD or at risk of cognitive decline for early treatment or secondary prevention.

- **Critical Parameters for Context of Use**

The Centiloid method allows the dementia field to use a central, universal metric which is valid across all three approved brain amyloid PET tracers so all researchers/clinicians/regulators can use a common set of clearly defined units. The methodology has been widely published in recent years in many studies. However new data from AMYPAD not only answers some critical questions but also alludes to the robustness of the procedure.

The Context of Use for this BQO application is summarised below in Figure 1.

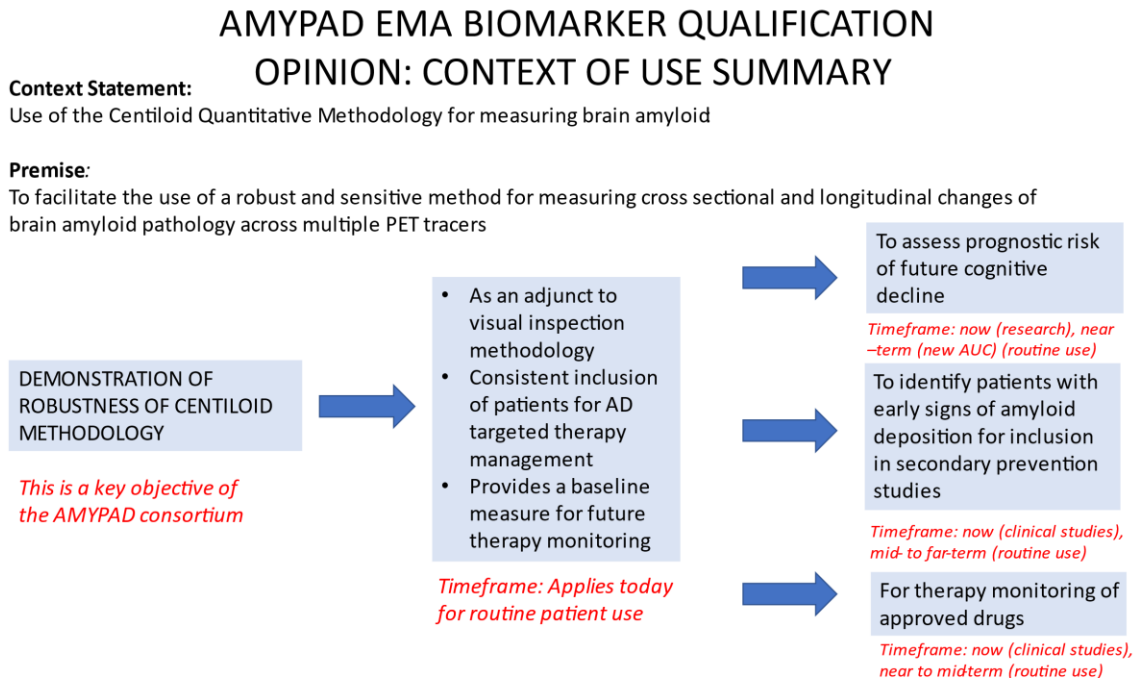


Figure 1 AMYPAD Biomarker Qualification Opinion: Context of Use Summary

4 CONCEPT AND LITERATURE OVERVIEW

4.1 Amyloid PET

The use of amyloid PET allows for the *in-vivo* visualization and quantification of Aβ protein fibrillary deposits in the brain, directly providing information on the total load and spatial distribution of Aβ pathology. Three fluorine-18 amyloid PET tracers are currently available for routine clinical use and have been validated against CERAD pathology as the standard of truth (SOT). These radiotracers are [^{18}F]florbetapir (Amyvid™; Avid Radiopharmaceuticals; approved in the EU since 2013) (Clark et al., 2012), [^{18}F]flutemetamol (Vizamyl™; GE Healthcare; approved in the EU since 2014) (Salloway et al., 2017) and [^{18}F]florbetaben (Neuraceq®; Life Radiopharma Berlin GmbH; approved in the EU since 2014) (Sabri et al., 2015). Each of these radiotracers has been approved by the FDA and EMA for routine clinical use and have local regulatory approval in other countries such as Japan and Korea. The tracers are also widely used by the research community. In addition, other known compounds such as the Carbon-11 labelled Pittsburgh compound B ([^{11}C]PiB) (Klunk et al., 2004) and [^{18}F]NAV4694 (Rowe et al., 2013; Villemagne et al., 2021) are available for investigational use only.

4.2 Centiloid Scaling

As the use of different amyloid PET tracers grew in both clinical and research settings, there was a need for inter-tracer standardization of the SUVr metric in multi-centre collaborations. To this end, the CL scale was developed (Klunk et al., 2015), which is an unbound scale with anchor points at 0 CL (mean grey matter signal of young healthy controls, which are assumed to be free of Aß plaques) to 100 CL (mean of a typical AD patient signal, when Aß burden peaks) that conveys a single patient's amyloid burden based on using the [^{11}C]PiB SUVr from the Global Alzheimer's Association Interactive Network (GAAIN) reference dataset (http://www.gaain.org/CL-project). The main aims of the CL scale have been to: (i) simplify and expedite direct comparison of Aβ PET results across sites and studies; (ii) outline the earliest thresholds for amyloid positivity and define the range of positivity in AD; (iii) robustly quantify longitudinal change; and (iv) facilitate inter-tracer comparisons (Klunk et al., 2015). Since then, several studies have tested the scale's validity and used it to improve the harmonisation and standardisation of Aβ PET quantification across tracers, scanners, and analytical implementations (Battle et al., 2018; Bourgeat et al., 2021, Bourgeat et al., 2018; Cho, Choe, Kim, et al., 2020; Cho, Choe, Park, et al., 2020; Leuzy et al., 2016; Navitsky et al., 2018; Rowe et al., 2017, Rowe et al., 2016; Schwarz et al., 2018; Su et al., 2018, Su et al., 2019; Tudorascu et al., 2018; Yun et al., 2017).

The CL approach allows any site using amyloid PET to follow a multi-step process to generate a CL scaling from their own local Aβ PET data. The basic principle is to scale the ^{18}F-labelled tracers' SUVr to equivalent [^{11}C]PiB SUVr, this is further transformed to the scale mentioned above. This process consists of a validation of the local pipeline using the GAAIN data and then the application to a new tracer (Klunk et al., 2015; Rowe et al., 2016). PET processing for CL quantification is often implemented through statistical parametric mapping (SPM) but other methods are available, including those without the use of an accompanying MRI (Bourgeat et al., 2021; Buckley et al., 2019). Routinely, PET images are first co-registered to their corresponding T1-weighted MRIs and subsequently transformed to MNI space. Next, PET images are intensity normalized often using the whole cerebellum as the primary reference region, other reference regions include pons, cerebellar grey matter and whole cerebellum plus brainstem. Finally, CL values are generated using the mean values of the standard CL target region based on a previously calibrated transformation (Klunk et al., 2015). The team behind the CL project and producers of the approved fluorine-18 labelled radiotracers have made progress in deriving and verifying conversion formulae that enable translation of non-[^{11}C]PiB Aβ PET semi-quantitative values to standardized [^{11}C]PiB measures (Battle et al., 2018; Bourgeat et al., 2021; Navitsky et al., 2018; Rowe et al., 2017), see Table 2 for conversion equations using the standard CL processing pipeline.

Table 2 Conversion Equations Applicable to the Standard CL Processing Pipeline for Generating CL Scores with the Most Commonly Used Tracers

Tracer	Variance (CL SD) Young Controls	Variance Ratio (Tracer SD/PiB SD)	Slope (Tracer SUVR to PiB SUVR)	Intercept	R^2	CL equation CL =
[^{18}F]Florbetapir (Navitsky et al., 2018)	12	4.6	0.54	0.5	0.89	$175.4*SUVR_{fbp}-182.3$
[^{18}F]Flutemetamol (Battle et al., 2018)	5.4	1.54	0.78	0.2	0.95	$121.4*SUVR_{flute}-121.2$
[^{18}F]Florbetaben (Rowe et al., 2017)	6.8	1.96	0.61	0.4	0.96	$153.4*SUVR_{fbb}-154.9$
[^{11}C]PiB (Rowe et al., 2016)	3.5	n/a	n/a	n/a	n/a	$93.7*SUVR_{pib}-94.6$

Ref: Adapted from (Krishnadas et al., 2021)

4.2.1 Implementation

Since its development in 2015, the CL scale has been widely implemented in research studies, including both AMYPAD studies and various clinical trials, see Figure 2 (Amadoru et al., 2020; Battle et al., 2018; Bullich et al., 2021; Collij et al., 2021; Farrell et al., 2018, Farrell et al.,2021; Jack, Wiste, Weigand, Therneau, Knopman, et al., 2017; Jack, Wiste, Weigand, Therneau, Lowe, et al., 2017; Klein et al., 2021, Klein et al., 2019; La Joie et al., 2019; Lopes Alves et al., 2020; Matsuda et al., 2021; M. Milà-Alomà et al., 2021; Marta Milà-Alomà et al., 2021; Mintun et al., 2021; Rowe et al., 2017, Rowe et al 2016; Royse et al., 2021; Salvadó et al., 2021, Salvadó et al., 2019; van der Kall et al., 2021).

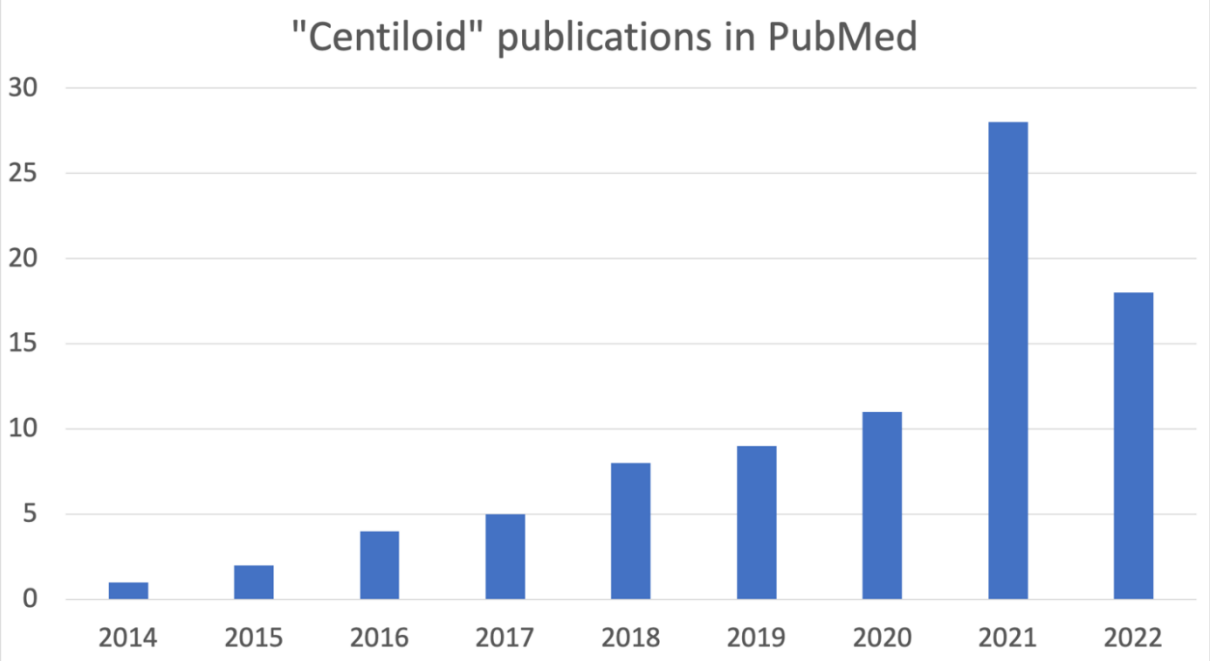


Figure 2 Bar Graph Showing the Increasing Use of CLs in Academic Publications

Ref: Numbers were obtained through a PubMed search for "Centiloid" in all fields on 27 June 2022 (hence not a complete year for 2022).

One of the key advantages of an 'absolute' metric of amyloid burden is generalisation of quantitative thresholds across tracers and pipeline implementations. Universal cut-off or threshold values to denote amyloid status can be applied alongside visual reads and in longitudinal multi-centre studies to facilitate inter-centre and inter-tracer comparisons. The CL approach has been validated against neuropathology (Amadoru et al., 2020; La Joie et al., 2019) where CL<10 correlates with absence of neuritic plaques, CL>20 specified at least moderate plaque density, and >50 CL best confirmed both neuropathological and clinicopathological evidence of AD. Clinical studies have also validated thresholds for amyloid PET positive status (Collij et al., 2021; Jack, Wiste, Weigand, Therneau, Lowe, et al., 2017; Royse et al., 2021, Dore et al., 2019), defined "grey zone" patient cut-offs (Bullich et al., 2021) and derived CL cut-offs to detect early amyloid abnormalities in cognitively unimpaired individuals (Jack et al., 2017; M. Milà-Alomà et al., 2021; Marta Milà-Alomà et al., 2021; Salvadó et al., 2021). Predictive models using the CL scale have been developed for calculating rate of cognitive decline in cognitively normal subjects (Farrell et al., 2018, Farrell et al., 2021; van der Kall et al., 2021). In addition, Hanseeuw et al. 2021 found that a CL threshold of 26 in memory clinic patients optimally predicts progression to dementia 6 years after PET (Hanseeuw et al., 2021).

The adjunctive use of quantitative information for image interpretation including both the Centiloid- and SUVr-based approach as metrics for the quantitation of amyloid load was recently validated and approved by EMA for Neuraceq PET scan assessment with CE-marked software packages. At the time quantification was added to the Amyvid and Vizamyl SmPCs the tracer dependent SUVr measures were used as the Centiloid-based quantification was incorporated in CE-marked software only very recently.

The three CE-marked software packages employed for Neuraceq quantification used the whole cerebellum as reference region and amyloid load was estimated with SUVr (Hermes Brass v.5.1.1, Neurocloud v.1.4) or Centiloids (MIMneuro v.7.1.2). All scans were quality controlled to ensure correct positioning of regions of interest; cases that did not pass quality control were excluded from the analysis (on average 2.6% of the cases analysed with CE-marked software out of 673 PET scans analysed). The mean sensitivity and specificity in three CE-marked amyloid quantitation software packages was 95.8±1.8% and 98.1±1.4%, respectively. The thresholds for amyloid quantitation were derived from samples with post-mortem confirmation of brain amyloid status as the standard of truth (from pivotal clinical autopsy cohort) using receiver operating characteristics (ROC) curve analysis. In a second dataset, the derived thresholds were used to categorise a test cohort and to compare the binary quantitative assessment and visual read. In a quality checked dataset, the average concordance between visual read and the CE-marked software packages was 91.2±1.7% and 96.2±1.8% in a subset where a group of readers had consensus in the visual assessment, i.e., all readers assessed the scans in the same way. The updated Neuraceq SmPC will available shortly at the EMA website. (Ref. EMA/CHMP/710821/2022 and EMEA/H/C/002553/II/0038)

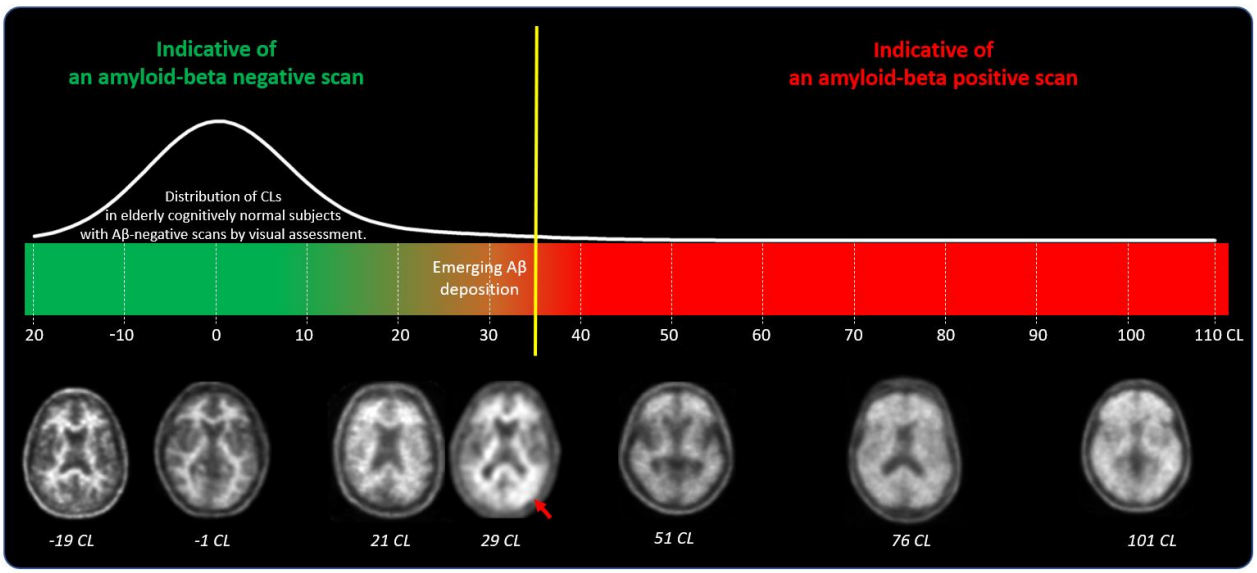


Figure 3 Illustrative display of the use of quantitative information to supplement visual assessment of Neuraceq PET scans in the EU.

In Figure 3 it is illustrated for Neuraceq PET scans that Centiloid values above 35 CL indicate established amyloid-beta pathology corresponding to a density of moderate and frequent neuritic plaques by neuropathology. Centiloid values below 20 are compatible with the Centiloid values found in elderly cognitively normal subjects with negative amyloid-beta scans by visual assessment. Centiloid values in the range between 20 to 35 CL are more likely to be ambiguous, can be either negative or positive by visual assessment and correspond to subjects with emerging amyloid-beta deposition. The readers should review such scans carefully to identify subtle amyloid accumulation that can be focal and unilateral (see red arrow).

4.3 AMYLOID PET FOR CLINICAL TRIAL USE AND THERAPY MONITORING

4.3.1 Previous Use of AMYLOID PET in Clinical Trial Data

Meta-analysis investigating the rate of amyloid positivity have shown in multiple papers that approximately 30% of subjects being clinically diagnosed with AD have a negative amyloid scan when scanned with amyloid PET (Barthel 2017, Fantoni et al., 2018, Shea et al., 2018, Kim et al., 2018) pointing to the possible failure of earlier amyloid targeted therapies in Phase III trials due to deficiencies in the early inclusion strategies. Indeed, early sub-studies with the research tracer C-11 PIB (e.g., by Sevigny et al., 2016) discounted nearly 40% of potential trial participants based upon visual inspection methodology, whilst early studies with bapineuzumab highlighted the futility of including amyloid negative subjects in a study of an amyloid targeted therapy. A subgroup screened with C-11 PIB identified 6.5% of APOE4 carriers and 36.1% of noncarriers with a clinical diagnosis of probable Alzheimer's dementia (NINCDS-ADRDA criteria) as having no evidence of brain amyloid burden (Liu et al., 2015) indicating the limitations of using clinical diagnosis alone. In more recent studies (see below) the measure of an amyloid burden either by amyloid PET or CSF has become routine for participant inclusion.

4.3.2 Current Routine Use (in USA only)

Aduhelm. The Biogen drug aducanumab (Aduhelm, human IgG1 anti-AB mab) was approved by FDA in 2021 under an accelerated approval mechanism which is based on the reduction of amyloid load as observed in patients treated with Aduhelm and which was measured by amyloid PET (Florbetapir) demonstrating active target engagement of

the therapy. Inclusion of amyloid status for initial diagnosis was included in an update of the US PI in April 2022 (i.e., to confirm the presence of amyloid beta pathology prior to initiating treatment) and hence could provide an opportunity for amyloid PET to provide baseline (t=0 information) for subsequent monitoring if required. Data showing the reduction in brain amyloid load from baseline is presented in both SUVr units (which are florbetapir specific) and Centiloid values which are tracer independent. Over 78 weeks of dosing with Aduhelm the amyloid load declined by approximately 60 CL units in the high dose group. On top of being tracer independent, CL values also provide an absolute metric of change. For example, if there is a reduction of 0.25 SUVr units this is tracer and reference region specific, but if a reduction of 60 CL is seen (as observed with Aduhelm) this gives a more general indication of how much plaque in a typical AD patient has been removed (Biogen Inc., 2022) (Figure 4).

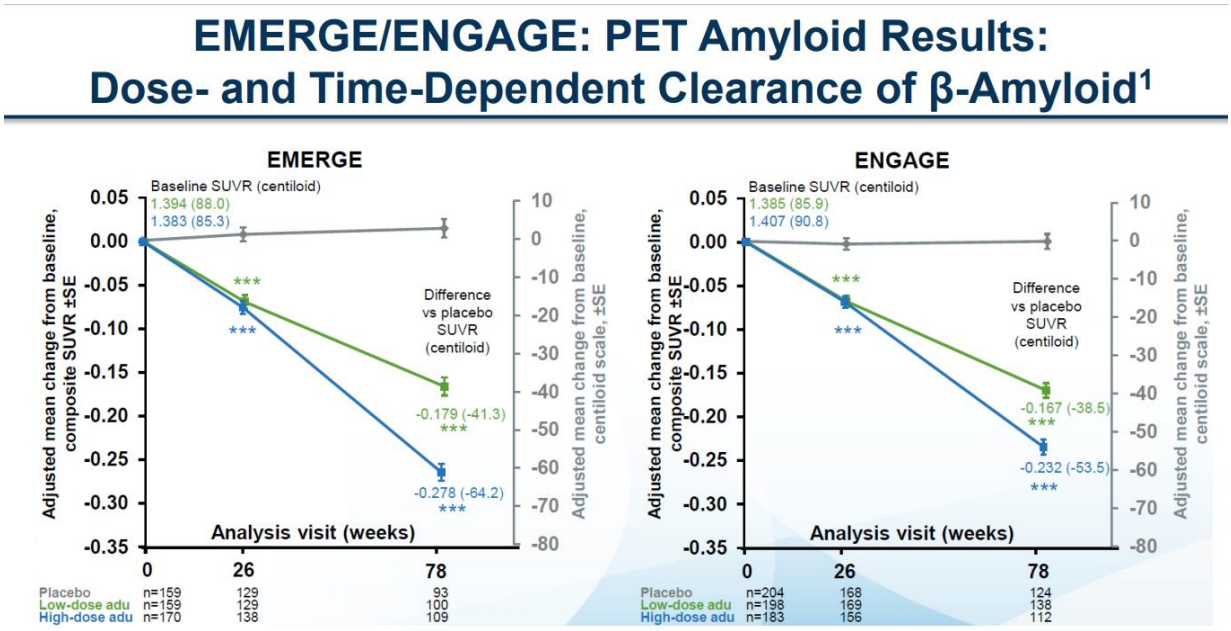


Figure 4 Example Use of Amyloid PET Showing Measurement of Baseline and Post-therapy Amyloid Burden by Both SUVr and Centiloid Measures

Ref: Graphs taken from ADPD 2022 Biogen Presentations (made available to conference participants). Data also presented in Aduhelm PI (Biogen Inc., 2022).

Additional data was presented by Biogen at the CTAD (Clinical Trials in Alzheimers Disease) Conference 2021 and demonstrates when there is a treatment gap (in this case of approximately 1.6 years) there was minimal redeposition of amyloid load as measured with a further amyloid PET scan (see Figure 5 below). From a utility perspective this demonstrates the opportunity that amyloid PET might have to contribute

to active patient management should anti-amyloid therapies become more routine in clinical care.

Reduction of Amyloid Plaque Levels Was Maintained During Treatment Gap From End of Feeder Studies to EMBARK Baseline[1]

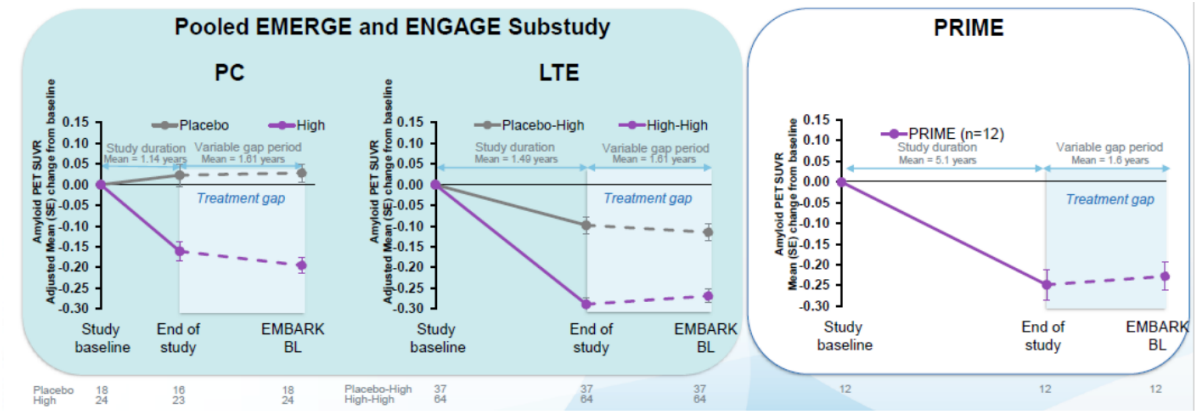


Figure 5 Amyloid PET Also Used to Show Amyloid Levels Did Not Rise During Interrupted Therapy Regime

Ref: Biogen data presented at CTAD 2021 and AAIC 2022 (Presentations made available to conference participants)

4.3.3 Possible Routine Future Use

4.3.3.1 Ongoing Therapy Trials with Other Humanised IgG Anti-AB-mabs

Eisai was recently (July 2022) granted priority review status by FDA based upon data from their Phase II program showing that lecanemab similarly reduced amyloid load (figure 6) with 81% of participants showing a negative visual read after 18 months of treatment in parallel with a reduced Aβ42/40 plasma ratio as well as a slowing down of ADCOMS* measure. The larger Phase III Clarity trial with over 1700 early AD patients is due to report out end 2022. Positive topline results were announced on Sept 28, 2022. The primary endpoint was met as a statistically significant reduction of clinical decline was demonstrated. Also, all key secondary endpoints were met including the change from baseline at 18 months compared with placebo of treatment in amyloid levels in the brain measured by amyloid PET.

*ADCOMS consists of 4 Alzheimer's Disease Assessment Scale–cognitive subscale items, 2 Mini-Mental State Examination items, and all 6 Clinical Dementia Rating—Sum of Boxes items. ADCOMS demonstrated improved sensitivity to clinical decline over individual scales in pAD, aMCI and in mild AD dementia.

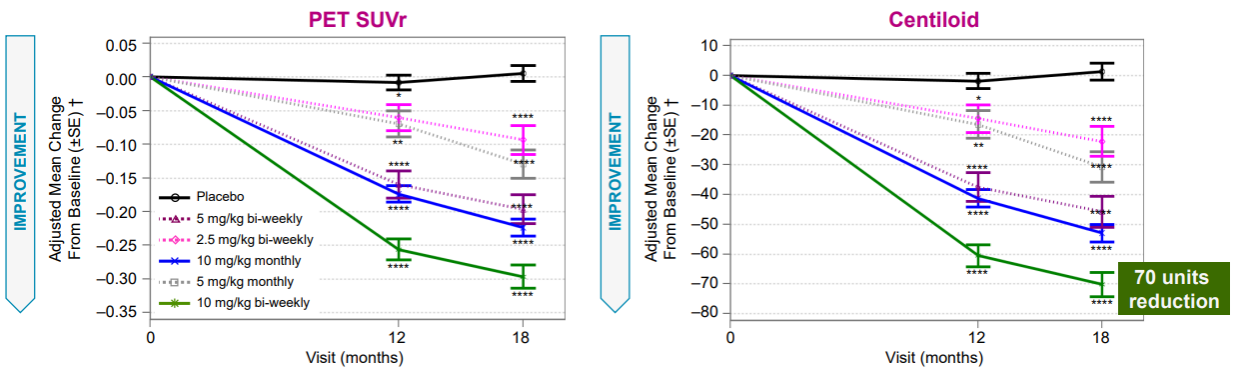


Figure 6 Dose-dependent reduction in amyloid PET observed in the lecanemab (BAN2401) Phase 2 study (CTAD 2018)

An earlier Eisai study (the elenbecestat Mission AD Phase 3 program in early AD patients) used the three amyloid PET tracers Amyvid, Neuraceq or Vizamyl to verify inclusion criteria (positive brain amyloid status) and to assess the amyloid load at baseline in Centiloid. Although the study was terminated a large cohort of nearly 10,000 subjects were screened for eligibility, of which amyloid PET was performed in almost 3,500 subjects. Overall, 386, 2548 and 558 subjects received Amyvid, Neuraceq or Vizamyl as amyloid PET tracer, respectively. Centiloid conversion and combining data of the 3 tracers yielded a mean of 1.5 CL for amyloid negative subjects and 82.7 CL for amyloid positive subjects screened in this study (Figure 7, left). The assessment of the baseline CL is consistent between the three tracers further demonstrating the robustness of the Centiloid metric as a tracer-independent measure of amyloid burden in the brain. (Figure 7, right) (Ref Roberts et al. 2020).

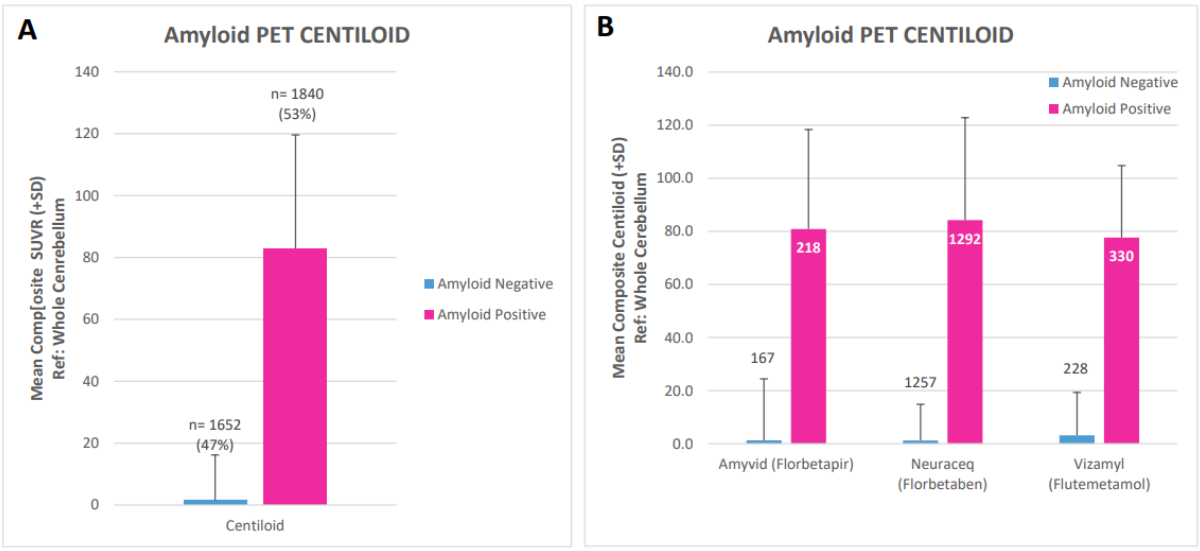


Figure 7 Illustration on the use of different amyloid PET tracers as study inclusion criteria in the elenbecestat Phase 3 Study (Roberts et al., 2020)

The investigational Lilly drug donanemab has also shown extensive loss of amyloid in their TRAILBLAZER-ALZ Phase II trial with nearly an 80 CL decline (figure 8) being recorded as well as significant slowing of changes in the following cognition scores (iADRS, CDR-SB, ADAS-Cog) (Mintun et al 2021). The trial also included a novel measure of tau to streamline recruitment of the amyloid positive subjects with only intermediate levels of tau included and both no tau (40-60% of subjects) or high tau (10-15%) patients excluded.

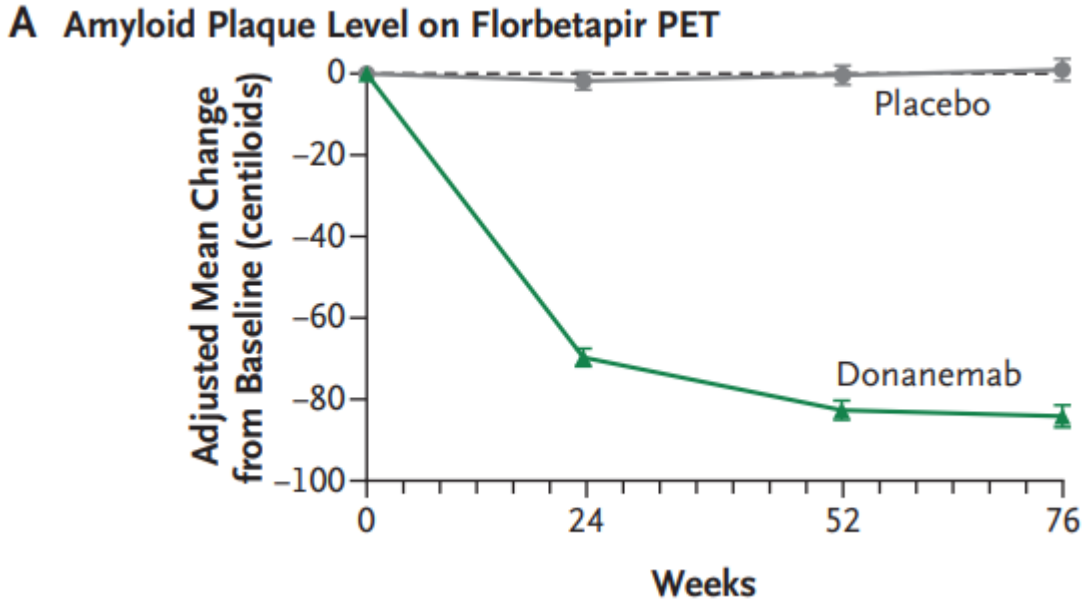


Figure 8 18F-florbetapir PET and Centiloid analysis as part of the secondary biomarker outcomes from the TRAILBLAZER-ALZ Phase II study. The change of the amyloid PET signal for placebo and donanemab treated patients from baseline to 76 weeks was assessed (Mintun et al., 2021)

Recent post-hoc analysis of this study indicates that the likelihood to fully clear amyloid in a certain time period is dependent on the baseline amyloid levels (in Centiloids), providing another relevant rationale for amyloid quantification and standardization. The Figure 9 shows three logistic regression curves to estimate the probability of achieving complete amyloid clearance with donanemab after 24, 52, and 72 weeks of donanemab treatment. The black bars represent the histogram of participants with corresponding baseline amyloid levels. Complete amyloid clearance is defined as having less than 24.1 centiloids (CL).

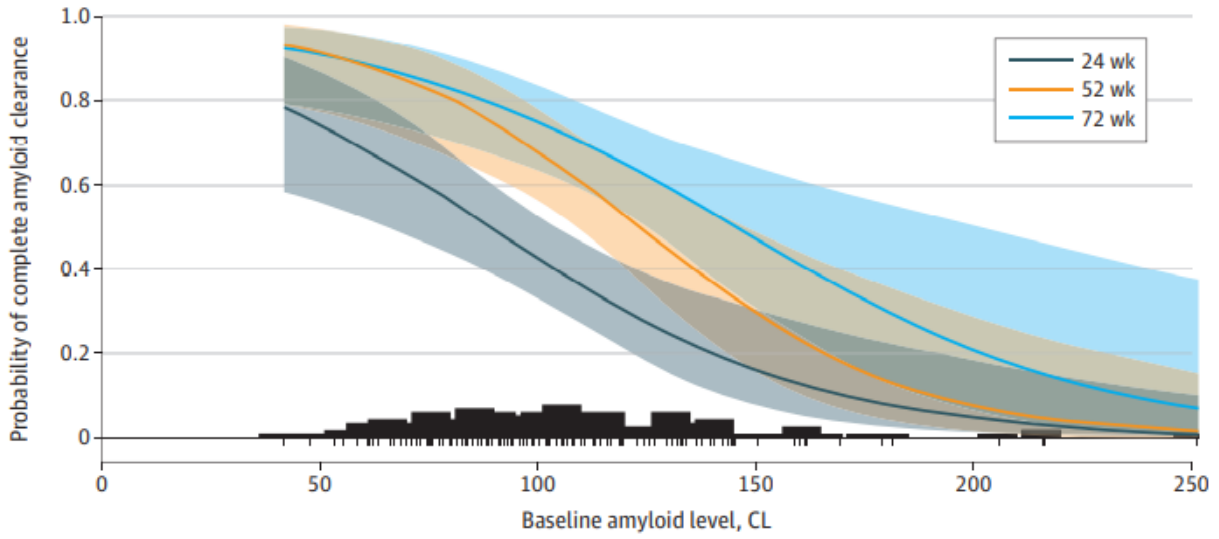


Figure 9 Probability of achieving complete amyloid clearance with donanemab treatment dependent on the baseline amyloid load measured by Centiloid (Shcherbinin et al., 2022)

The donanemab Phase III studies (TRAILBLAZER 2) have been designed based upon these inclusion measures and aim to report out end of 2022/beginning of 2023. Lilly has also initiated the TRAILBLAZER 4 study aiming to compare the rates of amyloid loss between donanemab and aducanumab over 18months of treatment. Results from this study are expected also by the end of 2022.

The Roche drug gantenerumab is an anti-amyloid therapeutic that is administered by subcutaneous route rather than by intravenous infusion. The two GRADUATE studies of nearly 1000 patients each will report out at the end of 2022 with the current primary endpoints relating to change from baseline in CDR-SB.

Different therapy monitoring strategies have emerged from the above clinical trials. In addition to monitoring the loss of amyloid over the dosing period (observed in all studies), following possible changes/re-emergence of amyloid load have been performed after cessation of therapy in both (EMBARK, PRIME, with antibody Aduhelm) as well as in lecanemab studies highlighting the opportunity that follow up amyloid PET could bring to managing future treatment regimes.

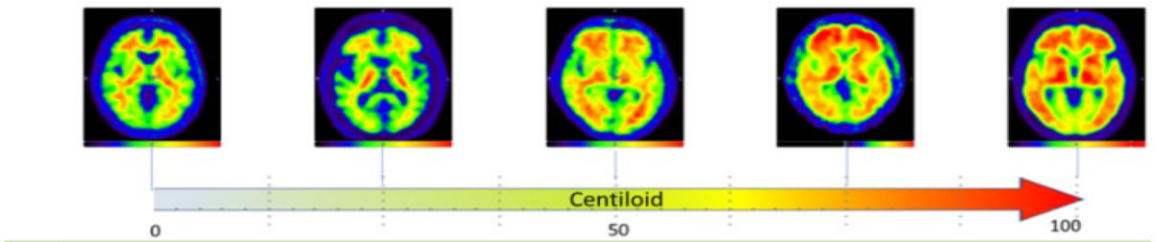
4.3.3.2 Ongoing Therapy Trials in Preclinical AD

Lecanemab, donanemab and gantenerumab all have ongoing studies involving the recruitment of preclinical AD. The AHEAD 3-45 study recruits preclinical subjects with

amyloid load between 20-40 CL units (with the aim to slow Aβ accumulation) whilst those over 40 CL are followed with the PACC-5 measure with the aim to show a slowing of cognitive decline. All subjects will receive lecanemab over a 4-year period. TRAILBLAZER 3 aims to see if short term donanemab treatment at the start of the study can slow progression to AD. Although enrolment is expected to be complete by the end 2022 the subjects will be monitored over the following 4 years to assess changes in cognition using the CDR-SB metric. Finally, Roche recently announced SKYLINE, a new 1200 subject preclinical gantenerumab study, aiming to use the PACC-5 score as an outcome metric over 4 years of study observation.

4.4 Summary

In clinical trial settings and future routine use, the quantification of the amyloid load by means of PET could provide important information to identify the optimal window for therapeutic intervention (Bischof & Jacobs, 2019). This is illustrated by the AHEAD 3-45 Study, which requires participants to have specific levels of amyloid pathology, either 'intermediate' (between 20-40 CL) or 'elevated' (>40 CL), signifying the added value beyond binary classifications (Aisen et al., 2020). The CL scale has been used in clinical trial settings to track therapy response measure (Bateman et al., 2020; Klein et al., 2021, Klein et al.,2019; Mintun et al., 2021; Roberts et al., 2021; Salloway et al., 2021), determine strategies for reducing AD prevention trial sample sizes (Lopes Alves et al., 2021) and improve patient selection for trials (Knopman et al., 2021; Weiner et al., 2017) and could assist in treatment endpoint decisions (Lopes Alves et al., 2020). Various cut-offs established in the literature are summarised in Figure 10.



Neuropathology [a, b, c]	**<10** = neuritic plaques absent, rule out AD **>20** = at least moderate plaque density **~50** = strongest correlation with AD diagnosis		
"Grey-zone" Visual Read	**12**----------**30** = equivocal, from earliest detectable (12)[b,e] to established Aβ (30)[d,e] **17** = optimal visual cut-off for highly experienced readers [f] **26** = high correlation with positive visual read[a]		
Disease Progression	**19** = "reliable worsening" of CL rate of change[g] **26** = optimal prediction of progression to dementia 6 years after PET[h]		
Clinical trial inclusion criteria (AHEAD 3-45[i])	**20**----------**40** "intermediate Aβ" Early preclinical AD	**>40** = "elevated Aβ"	

[a] Amadoru *et al.*, 2020. [b] La Joie *et al.*, 2018. [c] Doré *et al.*, 2019. [d] Salvadó *et al.*, 2019. [e] Bullich *et al.*, 2021. [f] Collij *et al.*, 2021. [g] Jack CR *et al.*, 2017. [h] Hanseeuw *et al.*, 2021. [i] https://clinicaltrials.gov/ct2/show/NCT04468659

Figure 10 Summary of Various CL Thresholds Established in Literature and In Use for Clinical Current Clinical Trial Inclusion (Pemberton et al., 2022)

5 DATA SOURCES

Summary of Main Data Sources for this Biomarker Qualification Opinion Document

The AMYPAD program has been comprised of 2 clinical studies providing amyloid imaging data with the 2 tracers Vizamyl and NeuraCeq to support this Centiloid Biomarker Qualification Opinion document.

The full protocol to the first AMYPAD Clinical Study DPMS (Diagnostic and Patient Management Study) can be found in Appendix B. Further summary detail on DPMS is to be found in Section 5.1.1.

In total 862 scans were performed (including 116 repeat scans) on 244 SCD+, 343 MCI and 257 dementia patients. Subjects were imaged with either Vizamyl (47%) or NeuraCeq (53%).

The full protocol to the second AMYPAD Clinical Study PNHS (Prognostic Natural History Study) can be found in Appendix C Further summary detail on PNHS is to be found in Section 5.1.2.

Final Status of PNHS (August 2022): 1419 scans (of which 226 were follow up images) have been prospectively collected from September 2018 to August 2022. An additional 1315 scans were contributed by other cohorts joining AMYPAD for a total of 2734 scans for analysis. The distribution of tracers used in PNHS was approximately Vizamyl (70%) or NeuraCeq (30%).

Additional analysis was performed by one of the centres (Barcelona) with a special interest in testing the robustness of the Centiloid methodology. Further detail can be found in Chapter 6 (see analysis components A1 and A4).

5.1 AMYPAD DPMS Summary

5.1.1 DPMS Study Overview

In summary, the AMYPAD-DPMS has investigated the clinical utility and cost-effectiveness of amyloid-PET in Europe.

Participants with subjective cognitive decline plus (SCD+), mild cognitive impairment (MCI), or dementia were recruited in 8 European memory clinics from April 16th 2018, to

October 30th 2020, and randomized into 3 arms: ARM1, early amyloid-PET; ARM2, late amyloid-PET; ARM3, free-choice.

A total of 844 participants were randomized (244 SCD+, 343 MCI, 257 dementia). We observed no relevant differences in sociodemographic/clinical features among recruiting memory clinics or with previous studies. The randomization procedure assigned 35% of participants to ARM1, 32% to ARM2, and 33% to ARM3; cognitive stages were equally distributed in the 3 arms.

The features of the AMYPAD-DPMS participants have been assessed and demonstrated to be consistent with those expected for a memory clinic population (Altomare et al., 2022). The analysis supports the generalizability of future study results.

Purpose of the Study:

This phase 4, multicentre, open-label, randomised study explored the impact of amyloid PET imaging on diagnostic thinking in the workup of patients with SCD+ (subjective cognitive decline associated with features that increase the likelihood of preclinical AD), MCI, or dementia where AD is in the differential diagnosis.

Primary Objectives:

To test the hypothesis that the proportion of participants for whom the managing physician reaches an etiologic diagnosis with very high confidence (≥90%) at 12 weeks after baseline is higher for participants who have not yet undergone amyloid PET imaging (i.e., participants scheduled to undergo amyloid PET imaging at 8 months [±8 weeks] after baseline). (According to Johnson et al., 2013, the managing physician is a dementia expert trained and board-certified in neurology, psychiatry, or geriatric medicine who devotes a substantial proportion [≥25%] of patient contact time to the evaluation and care of adult-acquired cognitive impairment or dementia.)

Table 3 Overall Summary of Amyloid PET Scans Collected in DPMS Study

Sites	Arm 1	Arm 2	Freechoice	Arm 1 Follow-Up
UNIGE	58	48	56	9
VUMC	44	43	41	32
CHUT	36	34	44	12
BBRC	33	29	29	15

Sites	Arm 1	Arm 2	Freechoice	Arm 1 Follow-Up
UKK	33	30	23	18
UCL	16	16	17	1
CHUV	23	14	10	14
KI	27	16	18	11
Total	270	232	244	116

Overall total scans collected = 862

Tracer balance; Vizamyl 47%/NeuraCeq 53%

Ref: data from AMYPAD (c/o data from Geneva team)

Summary of Image Analysis Workflow in DPMS

DPMS

The analysis workflow and tools used for the DPMS study are different from those used for the PNHS study as Magnetic Resonance scans are not acquired in this protocol. Figure 11 shows the workflow for analysis in the DPMS study. Briefly, PET only data was supplied to the GEHC core lab for analysis. The analysis was performed with Amypype - the PET only analysis system developed by GE Healthcare based upon the proprietary Cortex ID software tool. Cortical VOIs implemented in Amypype for DPMS quantification, fused with an MRI template shown in Figure 11. The system outputs a series of visual QC slices for each dataset. These are the PET images with the VOI overlays for the purpose of inspecting VOI placement. Special emphasis was given to the delineation of the reference region.

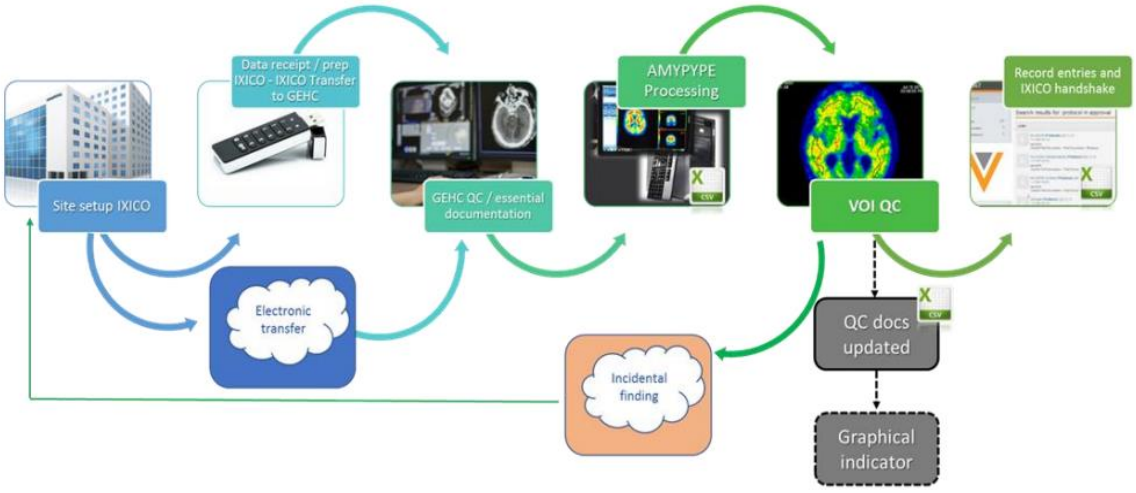


Figure 11 Processing Workflow for Data Generated in the DPMS Study

Ref: AMYPAD manuscript in preparation

5.1.2 AMYPAD PNHS

The AMYPAD Prognostic and Natural History Study (PNHS) was planned as an open label, prospective, multicentre, cohort study linked to several European Cohorts such as the main collaborator and sister project European Prevention of Alzheimer's Dementia (EPAD) Longitudinal Cohort Study (LCS) (ClinicalTrials.gov Identifier: NCT02804789). For the purpose of phenotyping and disease modelling, the EPAD LCS employs the concept of an Alzheimer's Disease (AD) risk probability spectrum that comprises four main dimensions – research participant clinical outcomes (e.g., cognition), disease biomarkers, and traditional risk factors (genetic and environmental) and change in these. With the incorporation of longitudinal change scores as relevant, these dimensions are used to estimate an individual's overall predicted probability of AD-related decline in terms of a variety of outcomes.

As one application of this strategy, AMYPAD PNHS is a natural history study that evaluated how amyloid imaging might help improve the understanding of the natural course of AD. The study assessed amyloid PET imaging as an additional and potentially relevant AD biomarker to complement the phenotyping and disease modelling efforts of parent cohorts (PC) such as EPAD LCS. As measured from [^{18}F]flutemetamol and [^{18}F]florbetaben PET images, brain amyloid load was quantified at baseline and mean temporal change in amyloid load was estimated. The ability to accurately estimate *in vivo* amyloid load could lead to a better understanding of disease evolution, earlier detection of the disease, and enable researchers to objectively monitor change in amyloid load to measure the impact of novel therapies.

Objectives:

The aim of this study was to evaluate the additional value of quantitative amyloid imaging analysis for modelling and assessing Alzheimer's Disease (AD) dementia risk in individuals without dementia, compared to a range of existing cognitive, imaging, laboratory and genetic biomarkers. Risk modelling is being performed to determine the optimal combination of quantitative amyloid imaging and other biomarker measures to determine placement of individuals on an AD risk probability spectrum.

Primary Objective:

To evaluate the value of quantitative PET amyloid imaging measures for predicting progression within an AD risk probability spectrum (derived from four different dimensions: cognition, other biomarkers, traditional genetic and environmental risk factors and temporal changes in these dimensions) based on quantitative PET amyloid imaging measures, with or without other biomarkers.

Table 4 Overall Summary of Amyloid PET Scans Collected in PNHS Study

Sites	Parent Cohorts	Time Point		
		Baseline	FU1	FU2
All Sites (wave-1 + wave-2)	EPAD	515	227	0
VUMC (Wave-1)	EMIF-AD (60++) - BL + FU	198	141	0
	EMIF-AD (60++) - BL	4	0	0
	EMIF-AD (90+)	64	21	0
	DPMS - BL	23	0	0
	DPMS - BL + FU1	16	16	0
	DPMS - BL + FU1	8	8	0
	DPMS - BL + FU1 + FU2	7	7	7
BBRC (Wave-1)	ALFA+ - BL + FU	211	211	0
	ALFA+ - BL	2	0	0
UNIGE (Wave-1)	Microbiota - only BL	43	0	0
	Microbiota - BL + FU1	5	5	0
Leuven (Wave-2)	FPACK	48	48	0
Fundaciò ACE (Wave-2)	FACEHBI - BL + FU1 + FU2	117	117	117
	FACEHBI - BL + FU1	52	52	0
	FACEHBI - BL + FU1	28	28	0
	FACEHBI - BL	34	0	0
UCL Louvain (wave-2)	UCL-2010-412	158	28	0
UGOT (wave-2)	H70	6	0	0
DZNE (collaborator)	Delcode	120	42	0
Total AMYPAD Prospective Scans (blue highlighted)	1419	570	725	124
Total AMYPAD Scans	2734	1659	951	124

Ref: Data from AMYPAD consortium (c/o Amsterdam team)

Summary of Image Analysis Workflow in DPMS

Quantification of the PNHS PET scans is performed with the aid of a structural Magnetic Resonance (MR) scan and using the analytic pipeline (LEAP) designed by IXICO for this trial. A schematic representation of the different steps of the pipeline is shown in Figure 12. Briefly, PET scans are co-registered and aligned to the MR scan and all quantification is performed in the participant's native space (i.e., not in the 'normalized' or space). Parcellation of the cortical and subcortical gray matter brain areas is performed using the LEAP pipeline and mean tracer uptake is measured in these regions. In addition, tracer uptake is also measured in regions typically showing high amyloid load in Alzheimer's patients. This region is based on the reference cortical Volume of Interest (VOI) from the Centiloid website (http://www.gaain.org/centiloid-proiect). This region in the MNI space is brought to the participant's native space by means (Global Cortical Average; GCA). On top of this region, a secondary target VOI is created by pooling LEAP regions overlapping with the GCA. In addition, white matter (WM) is also segmented from LEAP, further refined with morphological operators and used as a reference region, on top of the other standard ones in the Centiloid method which are available from the LEAP parcellation (Whole cerebellum, Cerebellar Gray and Pons). Finally, SUVr values are calculated as the ratio of tracer uptake between the GCA and the different reference regions and converted to Centiloids with calibrated linear transformations specific for each region and tracer.

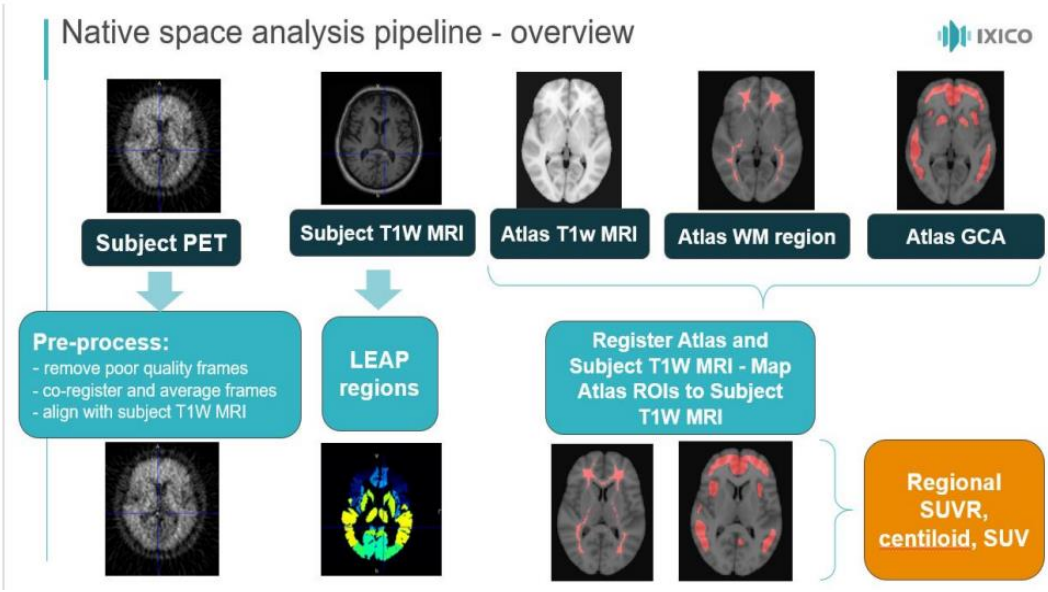


Figure 12 Processing Workflow for Data Generated in PNHS Study

Ref: AMYPAD manuscript in preparation

Other Data Used for Biomarker Qualification Opinion

The AMYPAD partners also used other sources of data available to them to support the BQO data generation exercise. In particular clinical development data were used from both NeuraCeq (see current SmPC https://www.ema.europa.eu/en/documents/product-information/neuraceq-epar-product-information_en.pdf)

and Vizamyl (see current SmPC https://www.ema.europa.eu/en/documents/product-information/vizamyl-epar-product-information_en.pdf)

Additional data for this BQO document was also derived from the IDEAS (Rabinovici et al., 2018) and ABIDE (de Wilde et al., 2019) studies.

6 ANALYSES, METHODS, AND SUMMARY OF RESULTS

6.1 Introduction to BQO Analysis Components

This chapter is subdivided into 3 sections; A) robustness of the Centiloid (CL) metric, B) cross-sectional evaluation, and C) measuring longitudinal change to support the context of use presented in Figure 1 . An overview of the presented analyses and key findings for the BQO are presented in Table 5. Each of the sections of work shown below in each row in Table 5 is presented in this chapter as a standalone piece which is composed in short abstract format comprising objectives, methods, results and summary. Final conclusions are also presented at the end of each section (A, B, and C).

Table 5 Overview of Presented Analyses and Key Findings for Biomarker Qualification Opinion

No	Title	Key Analysis	How results support BQO
A	**Robustness of the Centiloid Metric**		
A1	Evaluating the sensitivity of Centiloid quantification to pipeline design	Head-to-head comparison of 32 calibrated CL pipelines	Technical assessment of the robustness of the CL methodology
A2	Impact of error propagation in the development of the Centiloid conversion equation	Simulations of calibration data sets to compare error propagation to published CL equations	Low propagation (±3 CL) error at values around the negative/positive threshold indicate relative accuracy of measures
A3	CL measurement of AMYPAD data in a range of pipelines	Measurement of CL values from 80+ cases (flutemetamol)	Demonstration of robustness of measure using either a PET only or PET-MRI pipeline
A4	CL stability using CSF as anchors: Identify factors affecting CL and their impact (in CL units)	Using an independent biomarker of amyloid burden as a proxy for head-to-head analysis	Accuracy of CL transformation across different amyloid PET tracers
A5	CL measurement in a range of pipelines for florbetaben PET scans	Determining diagnostic performance of CL quantification in 589 cases	Demonstration of robustness of CL measurement using 5 different analytical pipelines
B	**Cross Sectional Evaluation**		
B1	[^{18}F]florbetaben (FBB) and [^{18}F]flutemetamol (FMM) PET analysis renders comparable CL estimates of amyloid burden from both DPMS and PNHS	Gaussian Mixture Modelling of cross-sectional CL values from routine patients	Consistent anchoring of negative subjects at CL=0 for both FBB and FMM. Provides robust CL threshold for routine use

No	Title	Key Analysis	How results support BQO
B2	Centiloid Measures from US IDEAS Study	Use of a central metric to compare quantitation to local visual read in an independent USA based clinical cohort	Harmonised use of Centiloid to assess the consistency of local visual reading against a pathology driven threshold
B3	Centiloid Quantification in a Clinical Cohort (ABIDE)	ROC between Centiloid quantification and visual read by an expert reader in a clinical population across the *continuum*	High agreement between expert reader and CL metric, optimal cut-off = 21 CL for visual positivity.
B4	Visual assessment of [^{18}F]flutemetamol PET images can detect early amyloid pathology and grade its extent	ROC between Centiloid quantification and visual read of 3 expert readers in an cognitively unimpaired population	Ability to detect early amyloid by expert readers (cut-off CL =17) and consistency of positive read at 25 CL
B5	Visual and Quantitative amyloid-PET measures in the AMYPAD DPMS Study	Influence of read experience on agreement between visual and CL measures	Value of CL metric in less experienced readers as an adjunct to image interpretation
C	**Measuring Longitudinal Change in Amyloid Load**		
C1	Estimation of longitudinal within-subject variability	Analysis of CL values in subjects projected to be stable	Estimation of inherent intra subject noise in the amyloid PET scan
C2	A Centiloid window to help predict true amyloid accumulation	Estimating annualized rates of change and estimate optimal CL values for subjects with significant changes in amyloid load	Value of the CL metric to demonstrate potential for amyloid accumulation within a defined time period

6.2 A) Robustness of Centiloid Measure

This section presents five (A1 to A5) pieces of work that illustrate the robustness of the Centiloid measure between tracers, sub-populations across the full AD continuum, various analytical pipeline scenarios and additionally using CSF measures as an anchor/reference standard against an independent biomarker of amyloid-β burden (due to the Covid pandemic it was not possible to add in a head to head comparison between tracers).

A1. Evaluating the sensitivity of Centiloid quantification to pipeline design

Relevance to BQO

Though Centiloid scaling pipeline and pre-processing steps are defined in standard (i.e., MNI) space, using a predefined reference region and cortical target mask (i.e., standard GAAIN, Klunk et al 2015), the Centiloid pipeline could deviate from standard design if different variables are implemented. Due to the importance of reliable quantification for both research and routine use, evaluating the impact of potential differences underlying in pre-processing pipeline design on CL quantification is crucial.

The Centiloid method includes one 'Standard' quantification pipeline based on routines in the Statistical Parametric Mapping (https://www.fil.ion.ucl.ac.uk/spm/) neuroimaging suite. Still, it opens the possibility for other pipelines (i.e. CE marked) calibrated to render Centiloid units. These alternative pipelines may perform better than the 'Standard' one in specific contexts of use (e.g. lower inter-site bias in multicentre studies, insensitivity to brain atrophy in patients, etc…) but, by design, CL units should be comparable across calibrated and validated pipelines.

Several design options are available when designing such pipelines. Some of them, like the selection of the **reference region** are partly included in the Centiloid method which provides 4 regions: Whole cerebellum (the primary one), Cerebellar Gray, Pons, and Cerebellum plus Brainstem.

Other typical pipeline design options are the following:

- **Space**: Analysing images in Standard MNI space (as the Standard Centiloid pipeline) or in 'Native' Space, that is in the original anatomy of the patient, thus retaining the information of brain volume, size etc.
- **Definition of Target and Reference VOIs**: The pipeline can use the standard GAAIN VOIs (as in the Standard Pipeline) or opt to refine them with subject-based segmentations of gray and white matter to minimize signal spill in from undesired tissues.

Aim

In this experiment, we aimed at estimating the potential bias associated to CL pipeline design options. To this end, we designed, calibrated and validated 32 different pipelines, combining the following design options:

- 4 reference regions (Whole Cerebellum, Cerebellar Gray, Whole Cerebellum + Brainstem, Pons)

- 2 target VOI (standard GAAIN vs subject-based)

- 2 reference region types (standard GAAIN vs subject-based)
- 2 analysis spaces (native vs MNI)

All the pipelines were implemented in the SPM environment, as the Standard CL pipeline that was used as the reference.

Included population and analyses performed

This analysis was performed primarily in 330 participants of the DPMS (207 acquired with flutemetamol and 123 with florbetaben). Amyloid PET scan, T1-weighted MRI, MMSE score, and APOE carrier ship status were available for each subject (**Figure 13)**.

Amyloid PET scans were quantified using the 32 calibrated CL pipelines. The impact of the pipeline design factors was assessed with the following statistical model and that was estimated using Generalized Estimating Equation (GEE) analyses.

Centiloid~ Intercept + Space + Tracer + Target Type +Reference region+ Reference region type+ visual read+MMSE

The model included all the pipeline design options as factors, plus two variables (visual read and MMSE) to account for global amyloid load in the participants. The impact of pipeline design option was measured as the difference in the marginal means of the factor with respect to the Standard pipeline. Finally, Tracer was also used as a Factor to assess whether the pipeline design options impacted CL values differently as a function of the radiotracer used.

From this statistical model, two main outcomes were used. The first one is the chi-squared value that accounts for the variability in the model and which was used to order the relevance of these factors with respect to their impact on CL values. Then, for every factor in the model, residual mean squared differences in CL units between the different pipeline design options (factor levels) and the reference pipeline were measured. This difference, in CL units, will be compared to the test-retest variability of CL values (approximately 2.5 CL, Battle et al 2018) and the statistical significance differences will be considered when the 95% confidence intervals do not overlap ($p<0.05$).

The demographic information of participants is shown in Table 6. Participants included from DPMS in this analysis had a mean age of 70.5 years, with a mean MMSE score of 25.7, and 32.3% prevalence of *APOE*-Ɛ4 carriership.

Table 6. Demographic Information of Participants Included for Evaluating the Sensitivity of Centiloid Quantification to Pipeline Design

Demographic	DPMS
Age (Mean±SD)	70.52±7.23
Sex (Female%)	138 (41.8%)
MMSE (Mean±SD)	25.67±4.14
Clinical status	SCD+(110, 33%) MCI (134, 40.6%), & Dementia (86, 26.1%)
Centiloid (Mean±SD)	46.33±48.86

Table 7 shows the results of the GEE model and the impact of each dependent factor on the Centiloid quantification. Higher chi-squared values indicate that the factor explains a higher percentage of the observed variability across pipelines (i.e. that the factor has a higher impact on determining CL values).

Table 7 GEE Model Results

Tests of Model Effects			
Source	**Type III**		
	Wald Chi-Square	**df**	**P-value**
(Intercept)	71.684	1	<0.001
Visual read	516.25	1	<0.001
MMSE	19.076	1	<0.010
Reference region	164.191	3	<0.001
Reference region type (GAAIN vs tissue-based)	84.601	1	<0.001
Target (GAAIN vs AAL-composite)	36.668	1	<0.001
Space (MNI vs Native)	9.564	1	0.002
Tracer	0.321	1	0.571
Dependent Variable: Centiloid			

Model: (Intercept), Visual read,MMSE,Target Type, Reference region,Reference region type, Space, Tracer

The factors 'Visual Read' and 'MMSE' are very significant, as expected, but irrelevant to the purposes of this study, as they were only entered to model the observed distribution of CL values in the studied sample. The 'intercept' only means that the average CL value (46.33) is significantly different from zero in the studied sample.

Among the remaining relevant factors, it can be observed that 'Reference Region' is the most significant one.

Table 8. Estimates from Reference Region

Estimates				
Reference Region	**Mean**	**Std. Error**	**95% Wald Confidence Interval**	
			Lower	**Upper**
Whole Cerebellum	42.115	1.504	39.167	45.0633
Cerebellum grey matter	45.480	1.590	42.362	48.598
Whole cerebellum + Brainstem	39.066	1.467	36.190	41.942
Pons	29.688	1.612	26.527	32.848

Table 8 shows that 'Cerebellum Grey' and 'Whole cerebellum + Brainstem' provide similar values (± 3 CL, approximately) to the primary reference region ('Whole Cerebellum'). Since their 95% Confidence intervals overlap, these differences are not statistically significant at the $p<0.05$ level. In addition, taking into account that the test-retest variability of CL values is in the range of 2.5 CL, these differences can be considered to be inconsequential. On the other hand, using the 'Pons' as the reference region resulted in significantly lower CL values (13 CL approximately) than the primary one, systematically across pipelines. This result might stem from the poorer performance of spatial normalization algorithms in this region, as reported in the original Centiloid paper (Klunk et al., 2015), in which the pons also showed the highest variability.

The following factor in order of importance was how the reference region was delineated (whether the ROI is refined with an individual tissue segmentation). In this case, the estimated bias of this design option was of approximately 3.5 CL which is close to the rest-retest variability (Table 9).

Table 9. Estimates based on Reference Region Type

Estimates				
Reference Region Type	**Mean**	**Std. Error**	**95% Wald Confidence Interval**	
			Lower	**Upper**
Tissue-based	37.299	1.479	34.398	40.200
Cerebellum grey matter	40.875	1.481	37.972	43.778

The following factors had decreasing impact on CL values. Using tissue segmentations in the target region provided average differences of approximately 2.5 CL and the quantification space (MNI) of 1.2 CL.

Importantly, Tracer was the least important factor, with mean differences in the order of 1 CL, irrespective of pipeline design options. Actually, this was the only factor that did not reach statistical significance (p=0.571).

Conclusion

Our results show that the Centiloid metric is robust against pipeline design options; the only exception being the use of the Pons as a reference region. The rest of factors resulted in mean marginal means in the order of the rest-retest variability of the CL metric. Of note, very similar CL values were obtained with the two amyloid PET tracers in the study, irrespective of pipeline design options, thus highlighting the robustness of the CL method to render comparable values across tracers.

A2. Impact of error propagation in the development of the Centiloid conversion equation

***Relevance*:**

The CL equation for each tracer is derived from a linear regression fitted to SUVR values of that tracer and of [^{11}C]PiB, obtained in a head-to head study [Battle et al 2018, Rowe et al 2017]. However, the estimation of this linear regression is affected by random noise, limited sample size, and distribution of SUVR values in the sample, leading to uncertainty in CL conversion equations, hampering the comparability of CL values across tracers. The objective of this analysis was to estimate the impact of error propagation in the initial generation of the CL conversion equations for [^{18}F]florbetaben (FBB) and [^{18}F]flutemetamol (FMM), using a simulation-based approach.

Methods

Linear regressions fitted to the head-to-head GAAIN datasets (http://www.gaain.org/centiloid-project) of [^{18}F]FBB and [^{11}C]PiB SUVRs, and of [^{18}F]FMM and [^{11}C]PiB SUVRs, were used to generate simulations (n=10000 reproductions of the calibration dataset) by adding heteroscedastic gaussian noise (bootstrap simulations). The CL calibration equation was derived for each noise realization, and the error propagation to CL values with respect to the theoretical ones was evaluated. Then, the maximum expected difference due to error propagation when pooling data from both tracers was estimated. A Jackknife approach was used to confirm the obtained results.

Results

The errors (95% CI) in CL estimates with respect to the theoretical values were associated to the sample size of the dataset used to develop the CL calibration equation: ± 3 CL (CL=0), ± 7 CL (CL=100), for N=35; ± 1 CL (CL=0), ± 5 CL (CL=100), for N=74.

The maximum difference (95% CI) between [^{18}F]FBB and [^{18}F]FMM CL values due to error propagation in the CL equations was small for low CL values: ± 3.5 CL (CL=0 and 25 CL), ± 4.5 CL (CL=40), and it grew at larger CL values: ± 10.5 CL (CL=100).

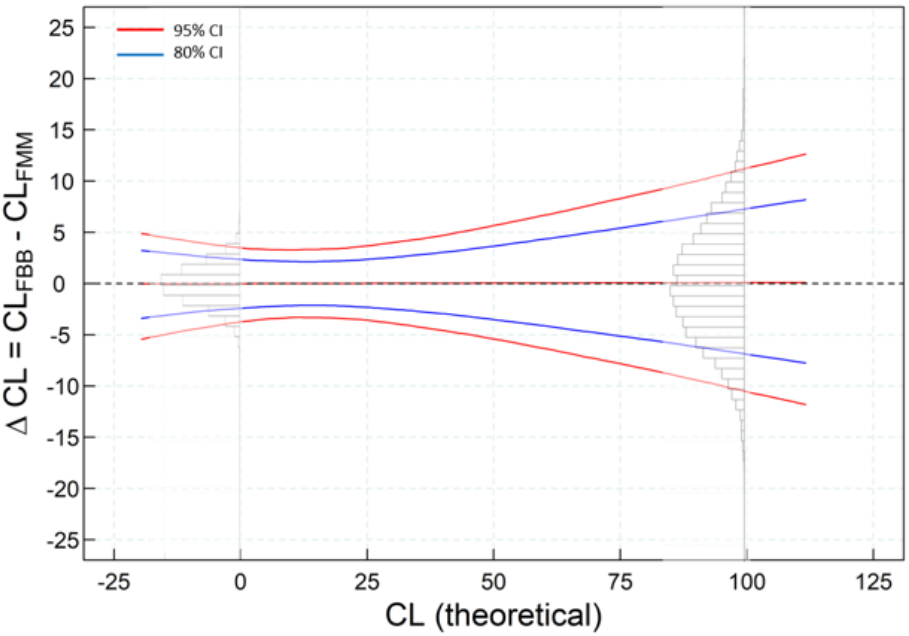


Figure 13. **Potential differences in CL values as estimated with the two different tracers, following the bootstrap-based estimation**

Conclusions

The propagation of errors in the generation of [^{18}F]FBB and [^{18}F]FMM CL calibration equations is expected to have a **small influence in the lower CL range** (<3.5 CL) but it increases to approximately 10% at the top end of the Centiloid Scale. The use of large sample sizes in the development of the CL calibration equations would reduce the likelihood of errors due to propagation of uncertainties.

A3: Cross comparison of Centiloid values from analysis pipelines used in AMYPAD

Relevance to BQO: To quantify amyloid PET images in Centiloids through different process pipelines used in the AMYPAD studies and compare the outcomes. By doing so the outcomes can be used to demonstrate the robustness of the CL metric to different pipelines as this is a likely future situation if quantification becomes more routine.

Methods: The Centiloid atlas was applied to a cohort of 82 [^{18}F]flutemetamol subjects, selected from the AMYPD DPMS and PNHS studies, using two independent Centiloid Pipelines; BBRC (c/o Barcelona) and AmyPype (c/o GE Healthcare). Both pipelines had been validated following the methods described by Klunk *et.al.* (Klunk et al., 2015). The Centiloid outputs from the pipelines was compared for variability and robustness.

Results: Mean CL for subjects processed through the BBRC pipeline was 13.2±24.9 CL (range -11.8 to 114.4). The AmyPype Centiloid output gave a mean of 11.9±26.7 CL (range -17.7 to 115.2). Analysis of the data gave a mean absolute deviation of 5.4±4.0. Correlation of the data was good, with an R^2 of 0.9369 (Figure 14). The relationship between the data is also illustrated the Bland-Altman plot (Figure 15).

Conclusions: The BBRC pipeline employs the GAAIN Whole cerebellum reference region, with the GAAIN Cortical mask VOI. PET and MRI images are co-registered and processed in MNI space. AmyPype employs the same GAAIN Whole cerebellum reference region and GAAIN Cortical mask VOI however the pipeline is PET-Only, with the images co-registered to the MNI152 T1-weighted average structural MRI template.

From the data we see a strong correlation (R^2 0.94) between the CL values generated from the BBRC and AmyPype Centiloid process pipelines. The low mean absolute deviation (5.4±4.0) is consistent with the noise levels found for Centiloid longitudinal/test-retest conclusions (see C1 section in this chapter). Therefore, from this analysis we can say that data generated from either pipeline is comparable and both pipelines are robust in their utility and output of Centiloid values for amyloid PET images.

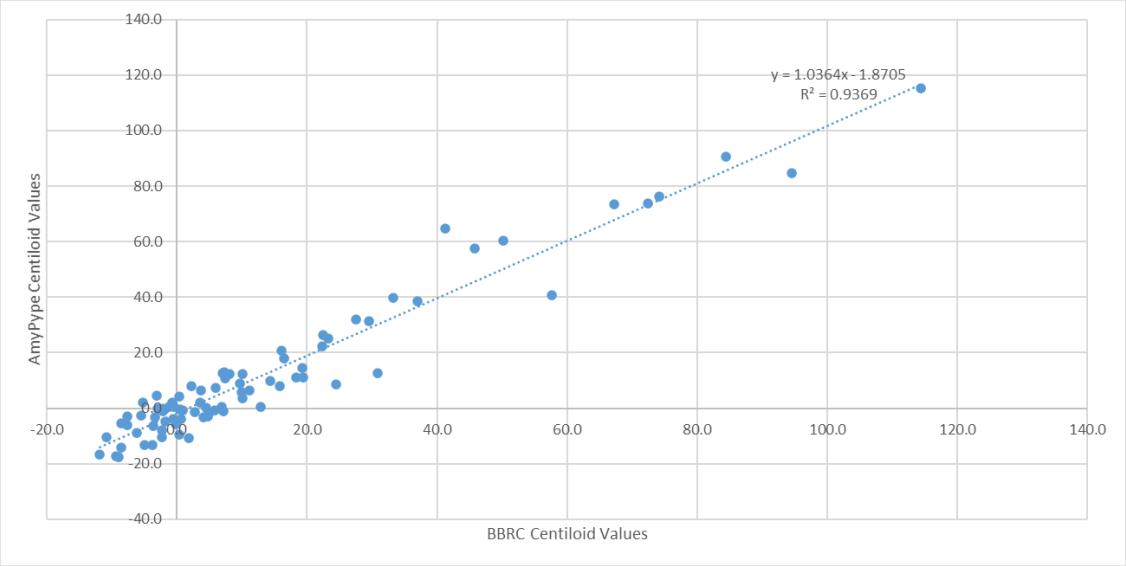


Figure 14. Correlation between BBRC Centiloid outputs and AmyPype Centiloid outputs in a cohort of AMYPAD subjects. N=83

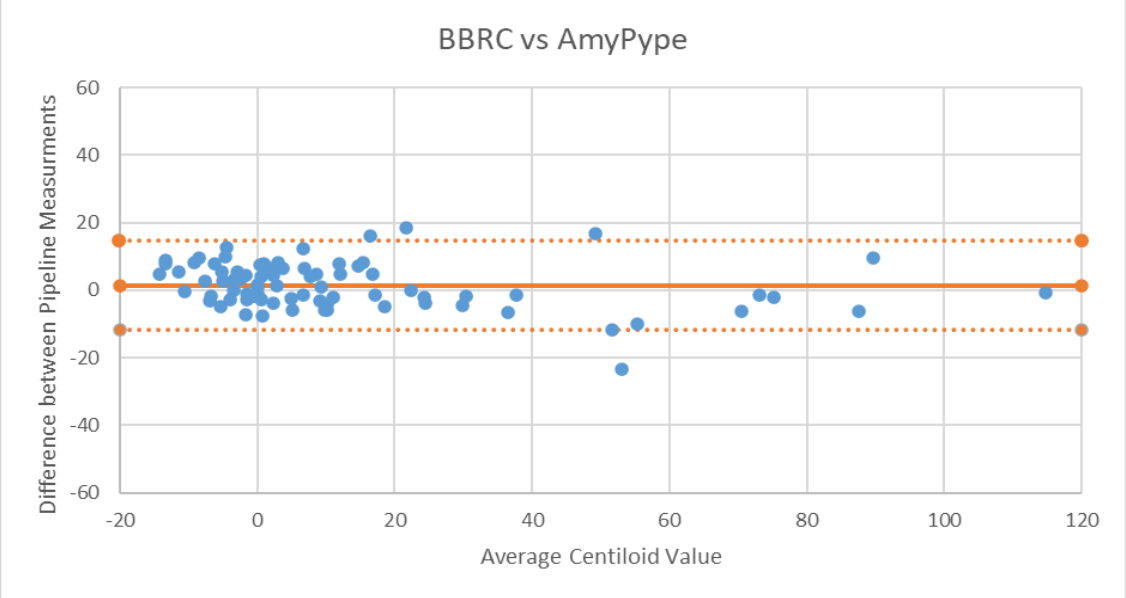


Figure 15. Bland-Altman plot showing the variation between CL outputs from BBRC and AmyPype process pipelines. N=83.

A4. CL stability using CSF measures as anchors: Identify factors affecting CL and their impact (in CL units)

Background / aim:

The Centiloid scale has been proposed to provide a universal metric to measure amyloid load, irrespective of the PET tracer. However, studies assessing the accuracy of the CL method for this purpose are lacking. In this work, we examined the accuracy of the CL method to render comparable CL values when using two different amyloid PET tracers. CSF measures were used as an independent measure of brain amyloid burden. This analysis was performed as an alternative to direct head-to-head comparison between tracers.

Methods

This analysis included 153 participants of the AMYPAD PNHS study with available CSF Aβ42 and pTau determinations (Roche Elecsys) and amyloid PET scans performed within 1 years' time (average difference 62±94 days). Amyloid PET scans were acquired with flutemetamol (n=125) and florbetaben (n=28) and quantified with a validated CL pipeline (IXICO LEAP method). Cerebral amyloid load was estimated using the CSF pTau/Aβ42 ratio as a proxy, as this ratio is a better predictor of Amyloid PET positivity and has a more linear association with CL values than CSF Aβ42 (Fig 16)

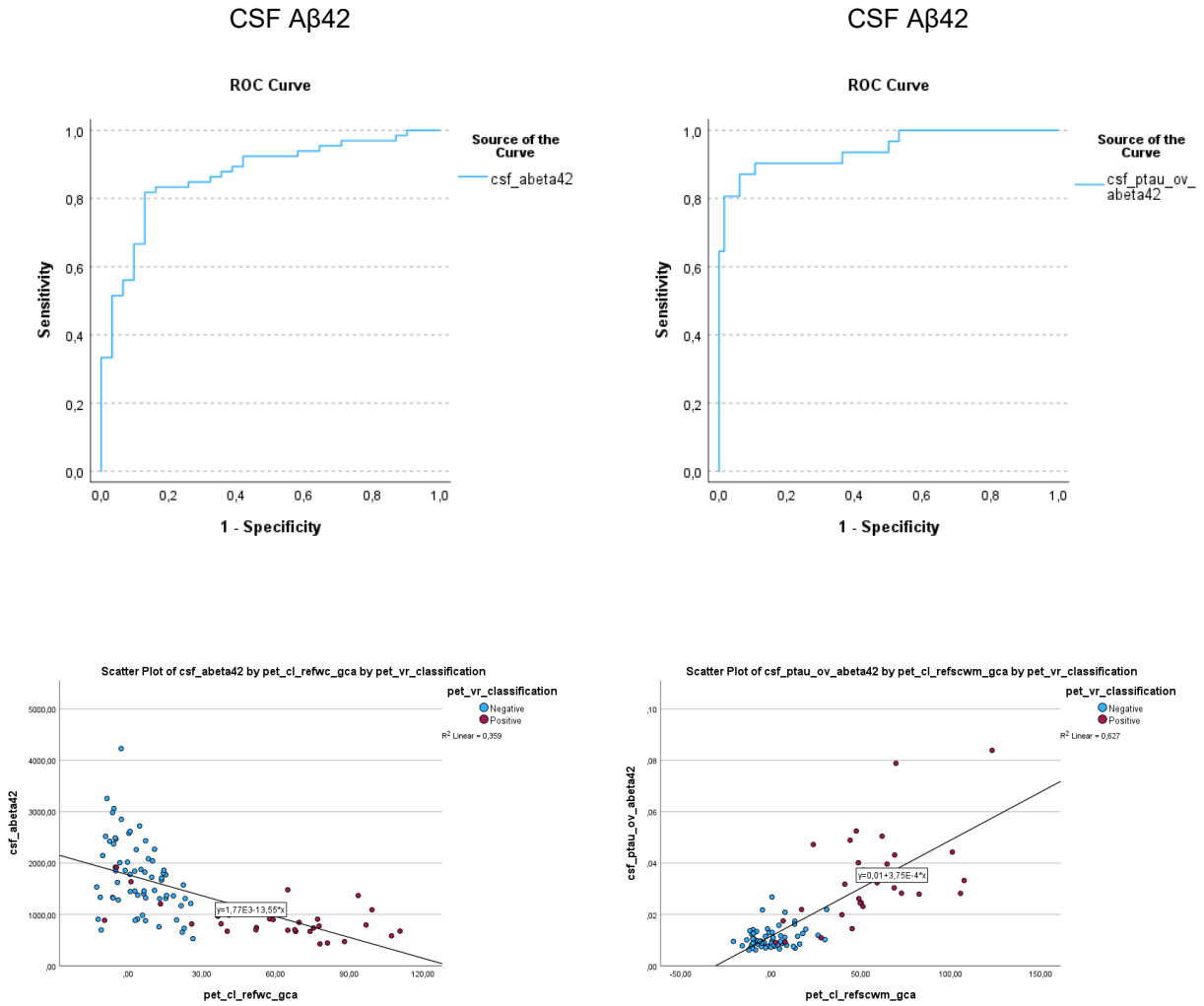


Fig 16: Top: Capacity of CSF Aβ42 (left) and CSF pTau/Aβ42 (right) to predict a positive visual read. It can be observed how CSF pTau/Aβ42 achives a significantly better predictive capacity (AUC = 0.945) than CSF Aβ42 (AUC = 0.870). Bottom: Associations between CSF Aβ42 and CSF pTau/Aβ42 with CL values. It can be observed that CSF pTau/Aβ42 has a more linear association with CL values.

Figure 16. Comparison of CSF pTau/Aβ42 and CSF Aβ42 for Measurement Cerebral Amyloid Load

In Figure 16, it can also be observed that the variance of the CSF pTau/Aβ42 ratio increases with CL values. In order to stabilize the variability and avoiding

heteroscedasticity issues with linear models, both the CSF pTau/Aβ42 and CL values were log10 transformed for statistical analysis.

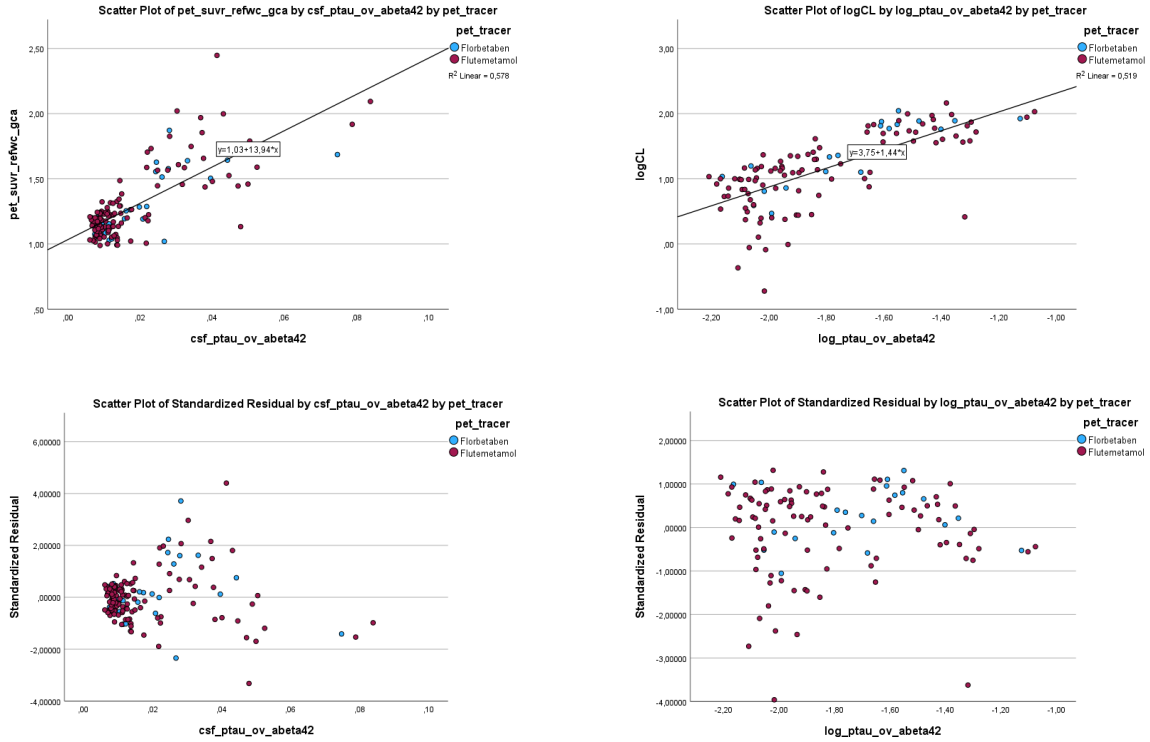


Fig 17: In the scatterplot in the top-left, it can be observed that the variability of the association of CSF pTau/Aβ42 is higher with higher CL values (bottom-left). In order to avoid heteroscedasticity in the linear models, the statistical analysis has been conducted with log10-transformed variables. The top-right figure shows the scatterplot of the log10-transformed variables and its residuals at the bottom-right. It can now be observed that the variability is kept constant through the full range of CL values.

Figure 17 Association between CSF pTau/Aβ42 and CL Values

A linear model was set up with log10(CL) values as the outcome and tracer and the log10 of CSF pTau/Abeta42 ratio and Tracer as predictors:

$$\log10(CL) \sim 1 + \log10(\text{CSF p-tau/Ab}) + \text{Tracer}$$

The difference in CL of the estimated marginal means associated with the two tracers was used a measure of the accuracy of the CL transformation between tracers. To this

end, the estimated marginal mean rendered by the statistical model was untransformed to render CL units.

Results

The difference in the estimated marginal means with flutemetamol and florbetaben was of 1.09 CL (95% Confidence Interval: [0.84, 1.42]), after accounting for the amyloid load with the CSF p-Tau/Abeta42 ratio. All tests for heteroscedasticity were negative ($p>0.05$).

Conclusions

Differences in CL values associated with the tracer are within the tolerance limits of the Centiloid transform (-2, 2) and within the range of the test-retest variability (2.5 CL) of amyloid PET imaging. This result confirms that the CL transformation accurately render comparable quantitative estimates of amyloid load irrespective of the PET tracer used.

A5. Validation of centiloids as an adjunct to visual assessment of [18F]florbetaben PET - a multi-software analysis (Source: LMI – data on file, abstract submitted)

Background

Amyloid positron emission tomography (PET) with [18F]florbetaben is an established tool for detecting Aβ deposition in the brain *in vivo* and has been approved for routine clinical use based on visual assessment (VA) of PET scans since 2014. Quantitative measures are however commonly used in the research context, with several available PET software packages capable of calculating amyloid burden both on a global or regional level, allowing continuous measurement of amyloid burden in addition to the approved dichotomous VA.

Objectives

This study aimed to provide scientific evidence of the robustness and additional value of florbetaben PET quantification using centiloids. The diagnostic performance (i.e., sensitivity and specificity) of quantification against the histopathological confirmation of Aβ load was estimated and compared to the effectiveness of the approved VA method. Additionally, the concordance between visual and quantitative evaluation of florbetaben PET scans was assessed. The reliability and comparability of the different analytical pipelines was further tested.

Methods

This is a retrospective analysis of florbetaben PET images that had been acquired in previous clinical trials. The study population consisted of 589 subjects with at least one available florbetaben PET scan from previously completed clinical studies. Florbetaben PET scans were quantified with five analytical methods reporting centiloids (MiMneuro (Piper et al., 2014), standard centiloid pipeline (Klunk et al., 2015, Rowe et al., 2017), neurology toolkit, SPM8 (PET-only), CapAIBL (Bourgeat et al., 2018), non-negative matrix factorization (Bourgeat et al. 2021). All the scans were quantified in batch mode to minimize operator intervention. The operators were different for each software package and blinded to the diagnosis of subjects, demographics, visual PET assessment, histopathology results and all other clinical data. All the results were quality controlled.

Results

The mean sensitivity, specificity and accuracy for all the quantitative methods was 96.1±1.6%, 97.4±1.2% and 96.7±1.2%, respectively. The mean percentage of

agreement between binary quantitative assessment and visual majority assessment was 93.2±0.4%. Substantial agreement was observed across software packages.

Conclusion

Results from this retrospective analysis demonstrate that centiloid values can complement visual assessment of florbetaben PET images. Such robust, validated methods could enable readers to augment their visual analysis with optional quantitative tools. Adjunct use of quantification software tools could be beneficial when images are assessed with relatively low confidence based on visual assessment alone, or when amyloid levels of patients are close to "pathology" thresholds.

Overall conclusion section A

The five pieces of work in this first section of results demonstrate in a variety of ways the robustness of the Centiloid metric as a tracer-independent measure of amyloid burden in the brain. AMYPAD has shown that pipeline design options do not markedly influence the CL measure nor did the propagation of error in development of the initial CL calibration equations (particularly in the low CL range). Additionally, when using CSF p-Tau/Abeta42 as a central anchor there was minimum differences in the CL levels between tracers too.

6.3 B) Cross-Sectional Applications

This section includes 5 pieces of work to illustrate the utility of the Centiloid measure in clinical populations across the AD disease *continuum*, and importantly the robust identification of a CL cut-off in high agreement with visual read positivity.

B1. The Centiloid scale applied to [^{18}F]florbetaben and [^{18}F]flutemetamol PET renders comparable estimates of amyloid burden in both memory clinic patients and those in the natural history study.

Relevance: Amyloid burden in Centiloid units for DPMS and PNHS

The goals of AMYPAD rely on the assumption that amyloid burden can be accurately quantified irrespective of the radiotracer that was used for the acquisition of the PET scans. In this regard, the Centiloid scale (CL) has been proposed to provide an absolute scale to quantify amyloid burden which can be compared across tracers and quantification pipelines. This scale assigns a CL value of zero to the lack of amyloid burden in a young control group and a CL value of 100 to the typical amyloid load of mild-moderate AD patients.

Methodology

In order to verify the assumption that CL values are comparable across the 2 tracers used in AMYPAD, we have conducted a Gaussian Mixture Modeling (GMM) exercise on the distribution of CL values in the DPMS and the PNHS. GMM is a data-driven statistical technique that is capable of estimating the parameters of a finite number of Gaussian distributions that underlie the observed distribution of values. GMM has been widely used to model global estimates of amyloid burden as measured by PET (Farrell et. al., 2021). It is well established in the literature that the distributions of amyloid load values, when recruiting memory clinic patients, show a bimodal distribution with one Gaussian modelling the distribution of 'negative' scans and another one that of the 'positive' ones (Klunk 2011, Nordberg et. al., 2012). Such a bimodal distribution fits well with the clinical use of the amyloid tracers that are typically rated visually as positive or negative. On the other hand, when scanning cognitively unimpaired individuals at high risk of AD, the distribution of values is dominated by a negative Gaussian that is skewed towards higher values (Farrell et. al., 2021). Such a distribution violates the assumption of the GMM that the data points are follow a finite number of Gaussian distributions. In addition, GMM presents another limitation such as it being sensitive to the initialization parameters. Another limitation of GMM is that it only provides point estimates of the Gaussian parameters (relative proportion, mean and standard deviation) but not spread estimates (i.e., the 95% confidence interval [95%CI] of such parameters).

In order overcome such limitations and robustly model the distribution of CL values even in challenging distributions such as that of the PNHS, we have introduced several methodological innovations to the modelling. First, in order to overcome the dependency of initial estimates and in order to provide spread estimates of the Gaussian parameters, we have implemented a bootstrapped version of the GMM. This method performs a GMM with random initial parameters in 100,000 'bootstrap' samples from the original distribution. Bootstrapping is a technique that randomly resamples a given distribution with replacement a high number of times. By doing so, it mimics the sampling of the recruited population in the study and is therefore capable of providing measures of accuracy to the sample estimates. Using this method, we are able to overcome the lack of spread estimates to the Gaussian parameters and by randomly initialization each of the bootstrap samples, we compensate the dependency of the GMM to the initial parameters.

In order to overcome the limitation of some the distributions not resulting from a finite number of 'pure' Gaussian distributions, a non-Gaussian distribution has been added to the GMM to model the intermediate CL values in the so-called 'gray-zone'. Such distributions were originally proposed to model the intensities of partial volume voxels in magnetic resonance imaging (Santago and Gage 1993; Santago and Gage 1995; Laidlaw 1998; Ruan 2000; Grabowski 2000). The distribution of the gray-zone values is linked to the positive and negative Gaussians, as it shares with them their mean and standard deviation values. Only the relative proportion of the 'gray-zone' values is estimated by the GMM.

Results and Conclusions of GMM for DPMS and PNHS by overall and tracer distribution.

Using this procedure, we modelled the distribution of the DPMS study as well as those after stratifying by tracer (Figure 18).

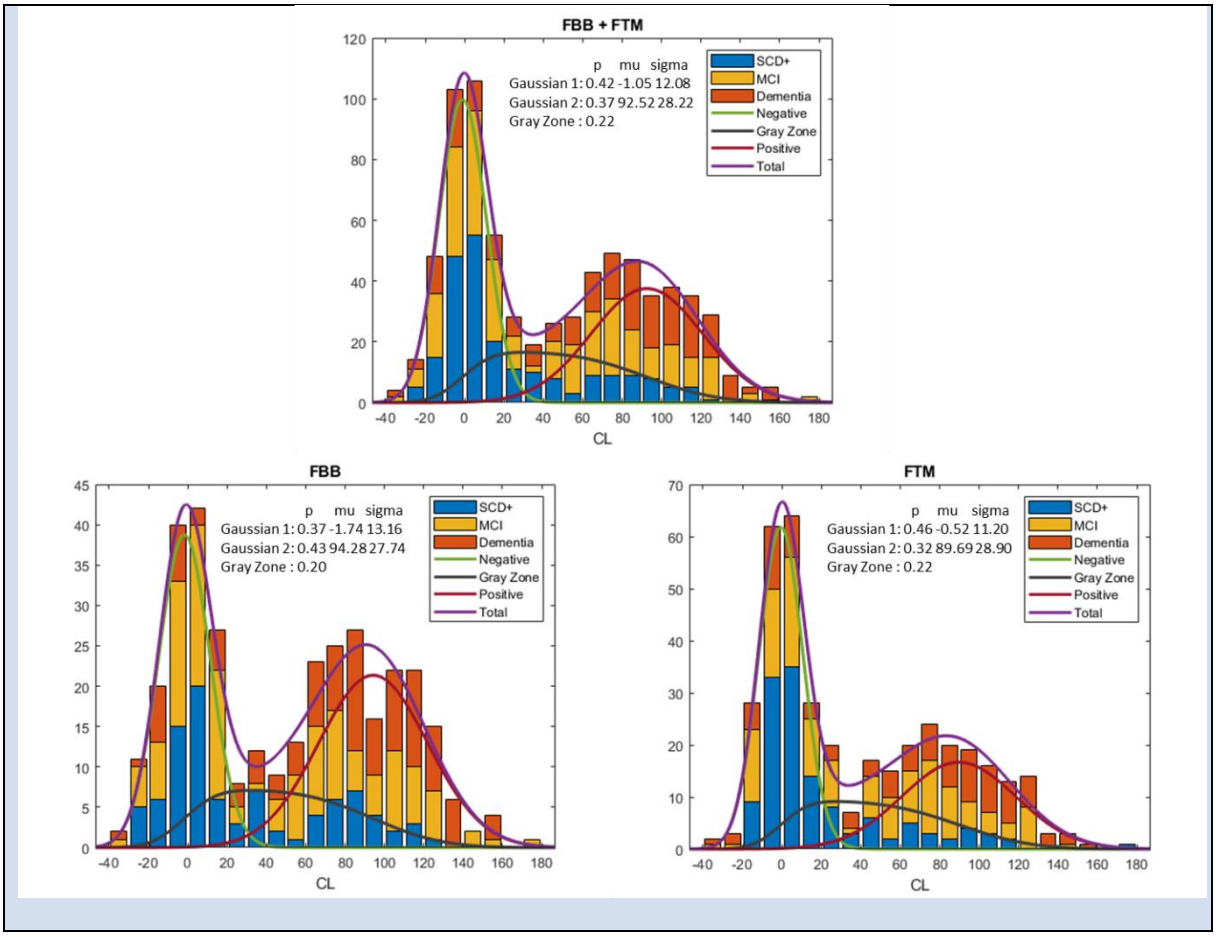


Figure 18 GMM: Distribution of CL Values in DPMS

Ref: Data from Amypad c/o BBRC presented at EANM October 2022.

Of note, our version of the GMM could estimate a negative Gaussian with 95%CI of the mean including the zero, as expected. Also as expected, the mean of the positive Gaussians was close to 100, albeit a bit lower as the amyloid positive participants of the DPMS might harbour slightly lower amyloid load than those used to calibrate the CL scale. Of note, when stratifying by tracer, the 95%CI of all parameters overlapped between the 2 tracers, thus confirming that the CL scale provides comparable estimates of amyloid burden across the 2 tracers in the DPMS.

Regarding the PNHS, the developed GMM method could also adapt to the expected distribution that was dominated by a negative Gaussian that was skewed towards higher values, as expected from a recruited population of primarily preclinical cases. In this case, it can be observed that the relative proportion of the 'gray-zone' distribution is

significantly higher than that of the positive one. In this more challenging scenario, the 95%CI of mean value of the negative Gaussian also included the zero, as expected. However, when the distribution was split by tracer, minor differences appeared in the estimated negative means (-2.57 CL for ^{18}F-Florbetaben and 2.25 CL for ^{18}F-Flutemetamol) (Figure 19). Still, these differences are in the range of the test-retest variability of CL values (Battle et. al., 2018) and have a negligible impact on CL quantification.

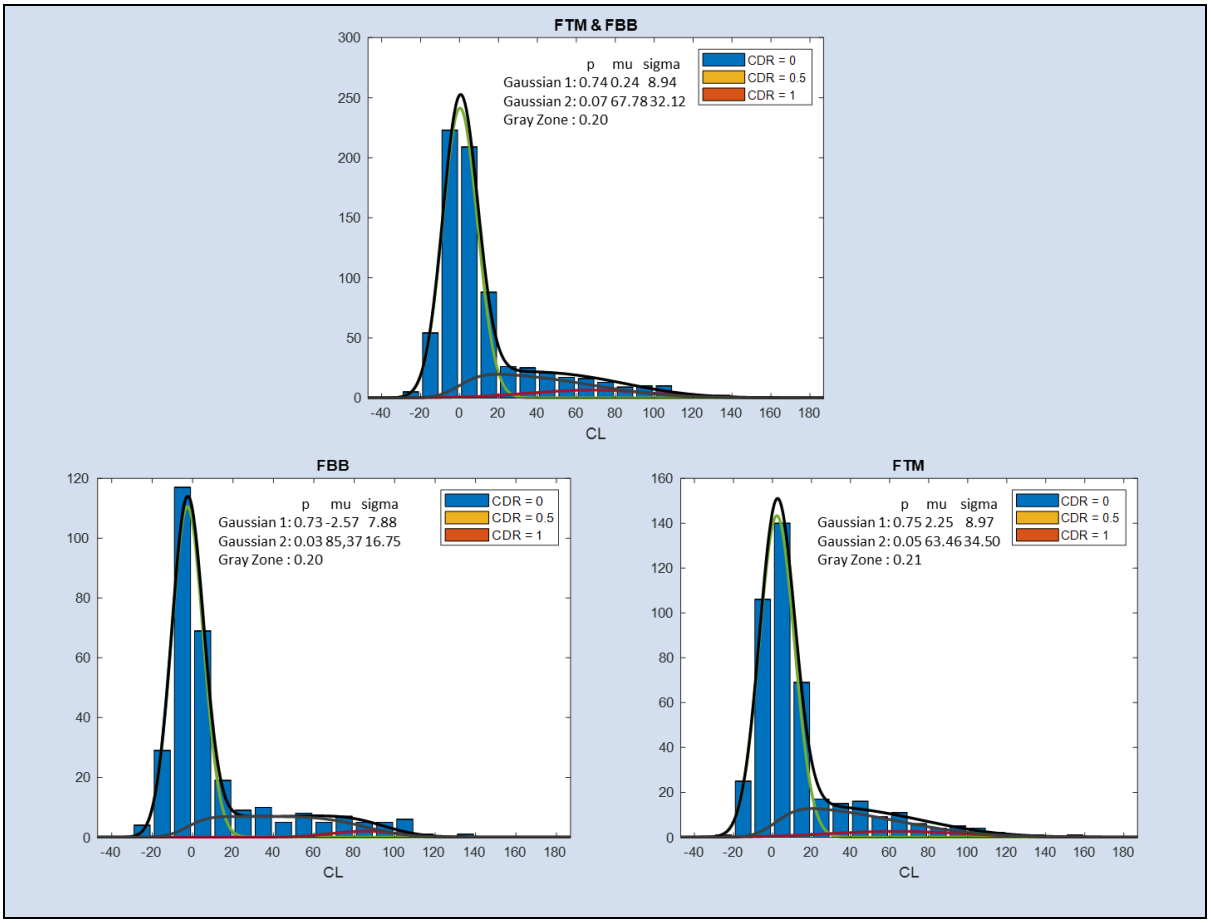


Figure 19. GMM: Distribution of CL Values in PNHS

Ref: Data from AMYPAD consortium c/o BBRC

B2. Quantitative Analysis of 6150 Real world amyloid PET scans from IDEAS (Data presented at AAIC 2022 by UCSF group).

Relevance: Centiloid Quantification was performed in an independent Clinical Cohort (IDEAS). Use of the tracer independent Centiloid metric to harmonise the quantitative measure of brain amyloid PET burden to assess the validity of amyloid PET in clinical setting with real world heterogeneity of patients and interpreting clinicians.

Methods.

Amyloid PET scans (n=6150) from all 3 approved tracers (florbetapir, florbetaben and flutemetamol) were collected from the IDEAs real world study and processed centrally at USCF. Centiloids were measured using the MRI-free pipeline rPOP (Laccarino et al., 2022) and compared to local visual reads using a pathology-based CL threshold of 24.4CL units (La Joie et. al., 2019) to define positivity independent of the visual read results

Results:

Centiloids presented in the violin plot below showed a bimodal distribution with a minority surrounding the 24.4 CL threshold. There was an 86.5% agreement between the local readers and quantification-based positivity. 53.3% of the visual reads were positive by both quantification and visual inspection whilst 33.2% were both negative. From a discordance perspective there were approximately equal amounts of visual negative/quantification positive (6%) and 7.5% being visual positive/quantification negative. The average CL value of the visually negative scans was approaching the zero CL value as expected (3 CL$\pm$27) whilst the CL value of the positive scans was 72 $\pm$41).

Conclusion: The Centiloid Metric was used in a large real-world study to provide a harmonized metric across 3 tracers to compare to local visual image interpretation and demonstrate the consistency of the image interpretation to a high degree.

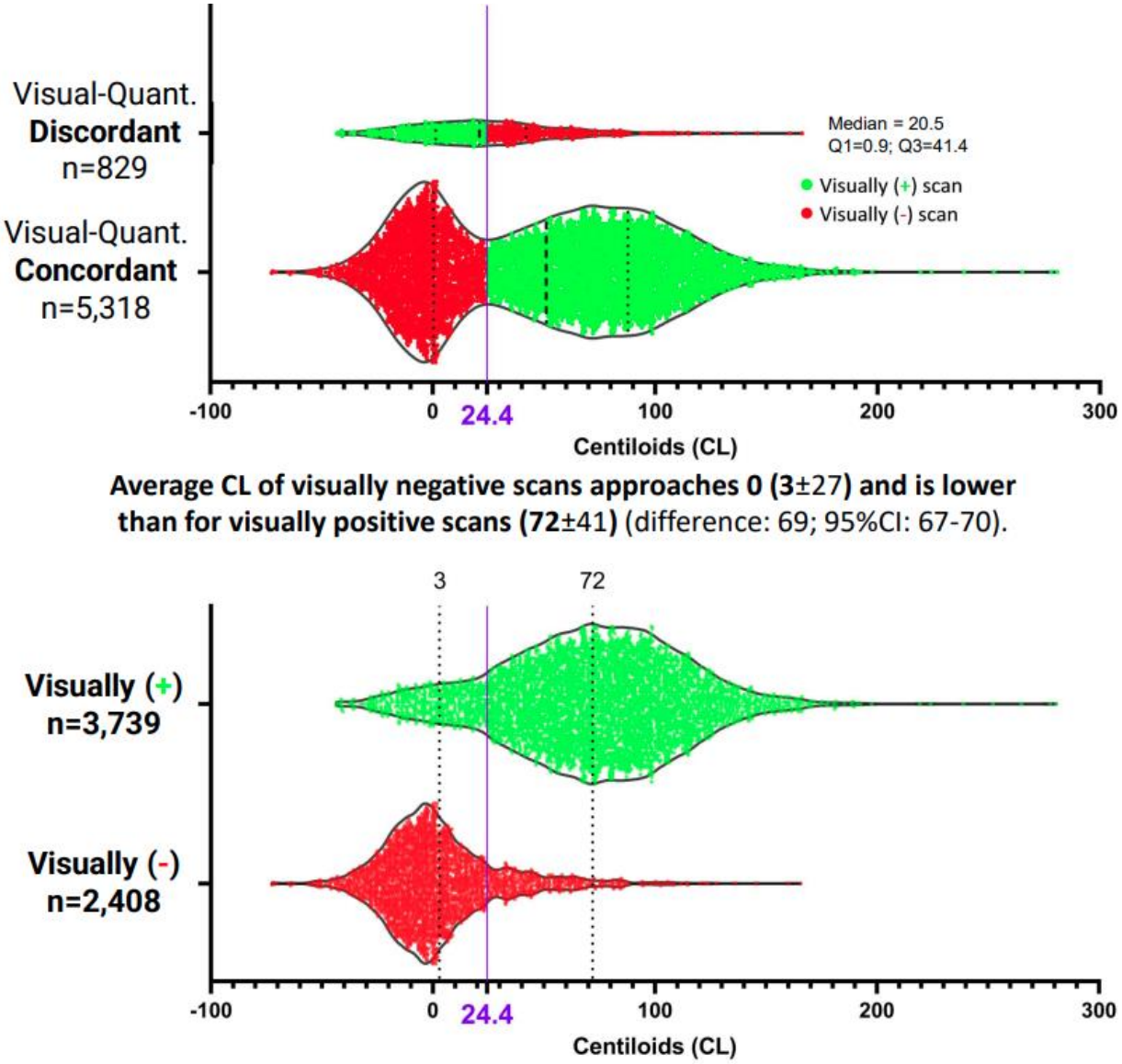


Figure 20. Centiloid distribution for IDEAs (all three tracers) showed a bimodal distribution. The CL threshold for visual positivity was 24.4 CL

Ref: Data from AAIC 2022 presentation, poster presented by Zeltzer et al

B3. Centiloid Quantification from a second independent Clinical Cohort (ABIDE)

Relevance

Centiloid (CL) quantification provides a more fine-grained picture of pathological burden and could hold complementary information to visual assessments in the clinical routine for diagnosis. However, little is known on CL measures in a heterogeneous memory clinical population, as most previous studies were performed in highly controlled research populations.

Methods

To assess the value of CL quantification in clinical routine, Collij et al., performed visual read (VR) and CL quantification of [^{18}F]florbetaben amyloid-PET images from an unselected memory cohort (i.e., ABIDE study), reflecting a heterogeneous clinical population including patients with AD and non-AD dementia, mild cognitive impairment (MCI) and subjective cognitive decline (SCD). First, they investigated the agreement between the 2 measures. Next, they determined the association of amyloid positivity and level of pathology with clinical stage and primary etiological diagnosis.

In total, 348 patients had a PET acquisition of sufficient quality for quantification purposes and were included in the current study. The final sample included 130 (37.4%) patients with SCD, 63 (18.1%) with MCI, and 155 (44.5%) with dementia. PET images were intensity normalized using the whole cerebellum as the reference region using the mask provided by the CL method (Klunk et al., 2015)(http://www.gaain.org/centiloid-project). Global cortical CL values were calculated using the standard GAAIN target region.

Results

Amyloid positivity based on VR was associated with a higher CL burden compared to visually negative patients images (VR-: 3.0±14.2; VR+: 61.7±27.9, F=624.3, p<0.001, Figure 16). Within VR+ patients, CL burden increased with disease severity (SCD: 51.5±27.2; MCI: 54.8±29.7 dementia: 66.6±26.5, F=4.42, p=.014). The optimal CL cut-off as indicated by the Youden index (VR reference) was CL=21 (J=0.86, AUC=.958, 95%CI:.935-.981; sensitivity=92.4%, specificity=93.7%). Utilizing this cut-off, 24 (6.9%) patients were considered discordant between VR and CL, with 12 (50%) patients assessed as VR+ but with a CL below 21 (i.e., CL-) and 12 (50%) patients assessed as VR- and a CL value above 21 (i.e., CL+).

The amount of amyloid pathology as expressed in CL units was associated with the primary etiological diagnosis after correcting for age, sex, *APOE*-ε4 carriership, and clinical stage, with AD patients showing the highest amyloid burden (62.9±27.5), followed by DLB (25.3±35.5) and CVD (16.7±24.5), and finally FTLD (5.0±17.22, t=-

12.66, p<0.001, Figure 17). Importantly, CL remained predictive of etiological diagnosis (t=-2.41, p=0.017) within the VR+ population (N=157), with mean CL burden being 64.1 (±26.5) for AD patients, 44.3 (±30.3) for CVD, and 49.9 (±35.9) for DLB.

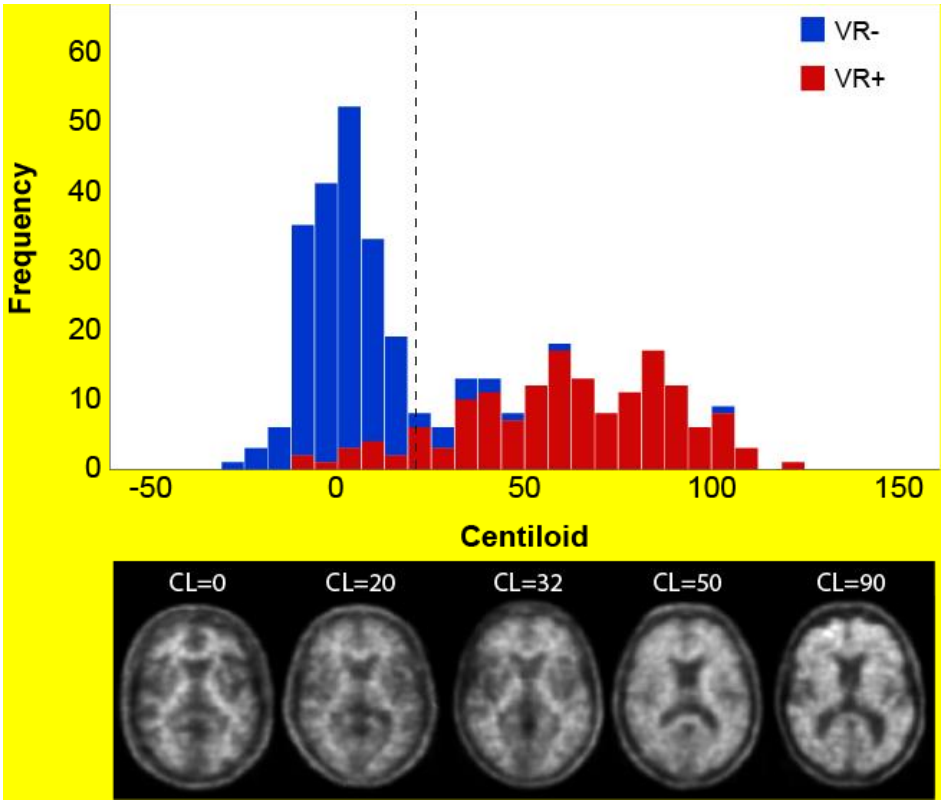


Figure 21 Centiloid Distribution Against Visual Assessment

Histogram displaying the frequency of Centiloid values colour-coded for visual read (VR) status. A bimodal distribution can be appreciated. Dashed line illustrates the optimal CL cut-off compared to VR as determined by the ROC analyses (i.e., CL=21). Below some illustrative [^{18}F]florbetaben scans are shown across the continuum.

Ref: Data under review in Alzheimer's & Dementia

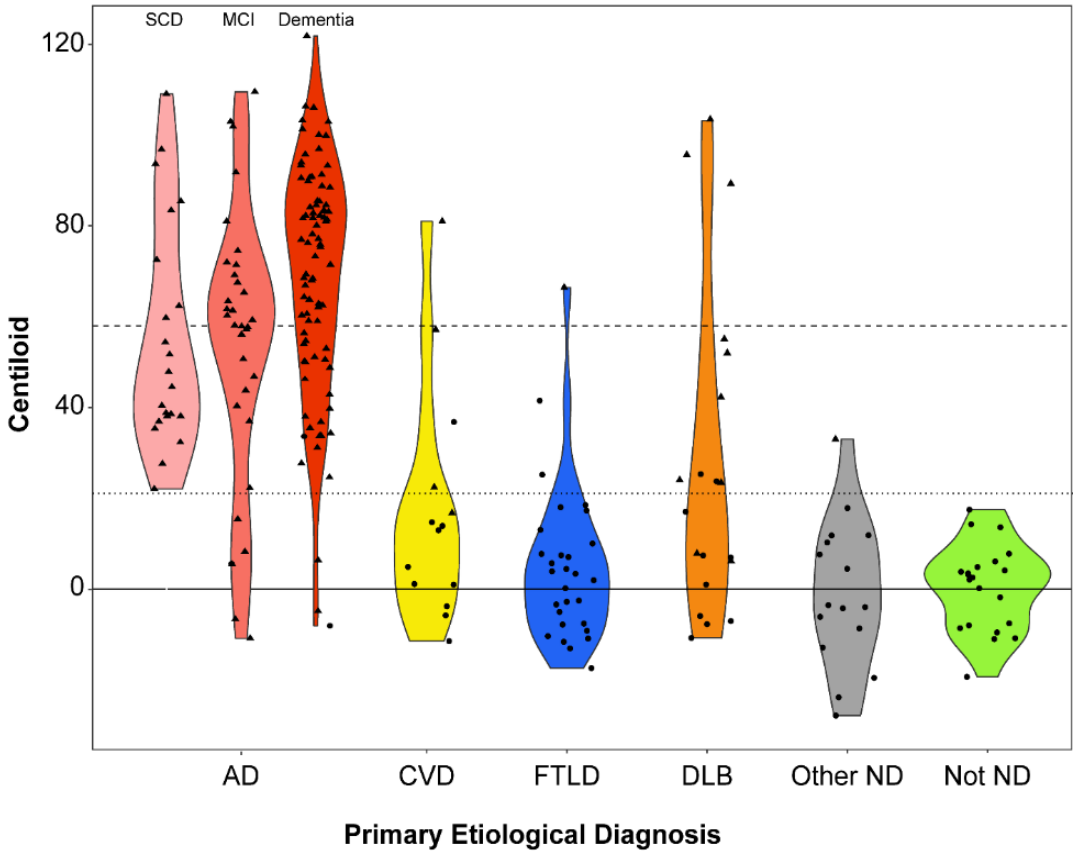


Figure 22 Centiloid Burden Across Etiological Diagnoses

Violin plot displaying the distribution of amyloid burden as measured in Centiloid units per primary etiological diagnosis. The AD population is further stratified based on clinical stage, showing a step-wise increase in amyloid burden. For the other diagnoses, data of only MCI and dementia patients are shown. The dotted line illustrate the optimal Youden index based cut-off against visual read (CL=21). Majority of patients with a non-AD primary etiological diagnosis fall below this latter cut-off. Visually positive patients are shown as triangles, while visually negative patients are illustrated as circles.
AD=Alzheimer's disease, CVD=Cerebrovascular disease, FTLD=Frontotemporal Lobar Degeneration, DLB=Dementia with Lewy Bodies, ND=Neurodegeneration, SCD=Subjective Cognitive Decline, MCI=Mild Cognitive Impairment.

Conclusions

CL quantification was in high agreement with VR status, with an optimal cut-off for positivity defined at CL=21. In addition, in visually amyloid positive subjects, the amount of pathology as expressed in CL units was associated with clinical stage and with primary etiological diagnosis, with AD patients showing the highest amyloid burden, followed by patients with DLB. In patients with dementia, CL could support assignment of a primary diagnosis of AD or non-AD. In this unselected heterogeneous memory

clinical population, the utility of CL quantification to complement VR of amyloid-PET is illustrated. Continuous CL measures could further increase diagnostic confidence, support differential diagnosis in clinically advanced cases, and provide improved prognostic information in preclinical subjects.

B4. Visual assessment of [^{18}F]flutemetamol PET images can detect early amyloid pathology and grade its extent

Relevance

This work illustrates the agreement between Centiloid quantification and consensus read of 3 expert readers in a population with emerging amyloid burden.

Methods

[^{18}F]Flutemetamol PET images of 497 subjects (ALFA+ N = 352; ADC N = 145) were included. Scans were visually assessed according to product guidelines by 3 expert readers. Scans were quantified using the standard and regional CL method. The agreement between VR-based classification and published CL-based cut-offs for early (CL = 12) and established (CL = 30) pathology was determined. An optimal CL cut-off maximizing Youden's index was derived. Results were confirmed using 28 post-mortem cases from the [^{18}F]flutemetamol phase III study.

Results

VR showed excellent agreement against CL = 12 (κ = .89, 95.2%) and CL = 30 (κ = .88, 95.4%) cut-offs. ROC analysis resulted in an optimal CL = 17 cut-off against VR (sensitivity = 97.9%, specificity = 97.8%) (Figure 23), with consistent positivity across readers at CL=25.

Conclusion

Centiloid quantification could support the identification of very early amyloid pathology. Consistent positivity was observed at CL=25, which could be considered an optimal cut-off for clinical use.

Frontal	-	+	+	-	+	+	+	+	+	+
PC/PCC	-	-	-	+	+	+	+	+	+	+
Lat.Temporal	-	-	-	-	-	-	+	+	+	+
Parietal	-	-	-	-	-	-	-	-	+	+
Striatum	-	-	-	-	-	-	-	+	+	+
Centiloid	2	17	19	27	32	35	38	47	57	81

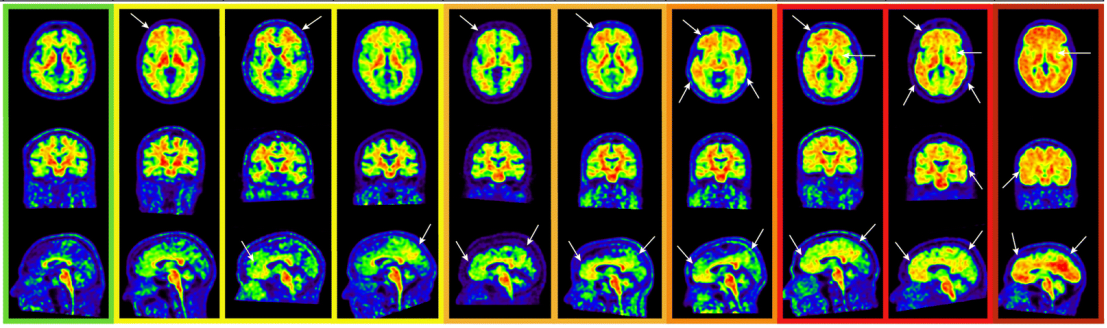


Figure 23 Example [^{18}F]flutemetamol Images

A series of 10 [^{18}F]flutemetamol scans form the ALFA+ cohort ordered based on Centiloid values are shown. Upper panel illustrates which regions were visually assessed as positive. From top to bottom, axial, coronal, and sagittal images are provided. White arrows highlight specific regional amyloid uptake. Note, that the main differences between VR- (left panel) and early amyloid accumulation (second to fourth panel) can be observed basal frontally on the axial image and in the orbitofrontal and precuneal regions on the sagittal images.
PC/PCC: precuneus/posterior cingulate cortex; VR: visual read
Ref: Collij et. al., 2021

B5. Visual and Quantitative amyloid-PET measures in the AMYPAD DPMS Study

Relevance:

In clinical practice, visual assessment of amyloid-PET images results in a binary classification (negative or positive) of amyloid status. While this approach has shown high clinical utility, the expected approval of disease-modifying treatments warrants the need for objective amyloid burden quantification, to further support a high-confidence diagnosis of AD and treatment decisions. Two widely available quantification methods, i.e., CL and z-scores to optimize clinical use of amyloid-PET were investigated in AMYPAD DPMS.

Methods:

Patients with subjective cognitive decline plus (SCD+; N=220), mild cognitive impairment (MCI; N=293) or dementia (N=216) enrolled in the AMYPAD-DPMS that have undergone amyloid-PET imaging and passed quality control were included in the study. Acquired static (i.e., 90 to 110 minutes post-injection) [^{18}F]flutemetamol (N=380) or [^{18}F]florbetaben (N=349) PET images were processed using GE Healthcare's AMYPYPE PET-only pipeline, providing global CL and z-scores, and z-scores for 7 cortical ROIs. Local readers provided visual read (VR) status and regional assessments.

Results:

Demographics for our DPMS study are shown in Table 9. In total, 364 (49.9%) patients were assessed as amyloid VR-positive and showed significantly higher CL (U=3954, p<0.001) and global z-score (U=4197, p<0.001) values across clinical stages compared to VR-negative patients (Figure 24A). The proportion of VR-positive patients increased with clinical stage (SCD+: 30.0%, MCI: 48.8%, dementia: 71.8%) and global amyloid burden (CL: H=24.97, p<0.001; z-score: H=24.43, p<0.001) was also associated with disease severity (Figure 24B). VR-positive patients with a primary etiological diagnosis of AD showed higher overall higher amyloid burden (CL: H=6.93, p=0.008; z-score: H=5.54, p=0.019) and specifically in the pre-frontal cortex (H=8.18, p=0.004) compared to patients with a non-AD primary etiological diagnosis, though percentage of frontal VR-positivity was highly similar between groups (AD: 97.2% vs. non-AD: 96.6%). Global CL and z-scores were highly correlated (ρ=.987, p<0.001), by a factor of 10 (Figure 24C).

Conclusion:

Our results demonstrate high agreement between CL and z-scoring quantitative measures in a clinical population. While initial results do not indicate differences in patterns of regional positivity between AD and non-AD subjects, differences in global amyloid burden seem to be driven by known AD-related early accumulating regions.

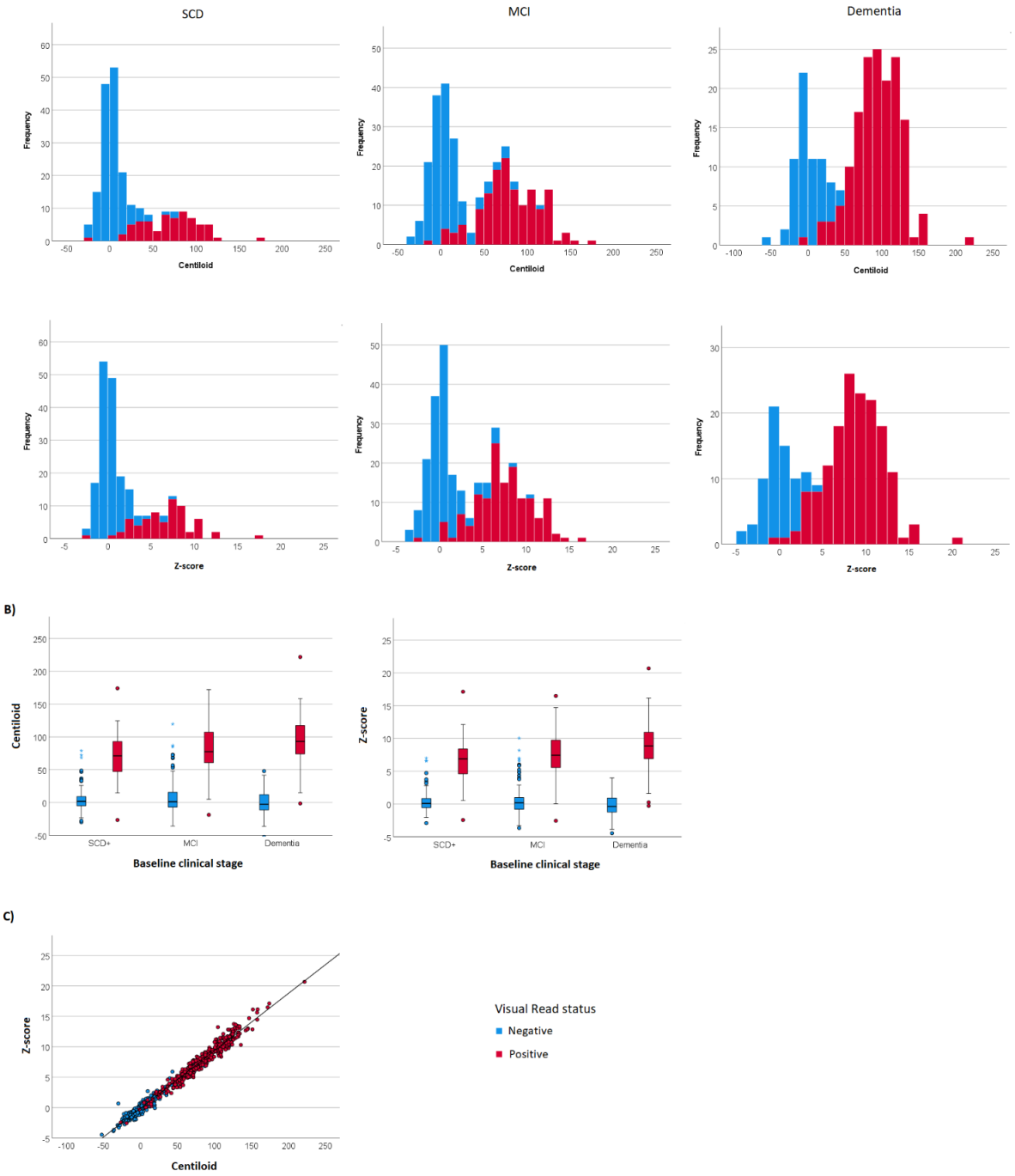


Figure 24 Quantification Against Visual Read Status Within the AMYPAD Diagnostic and Patient Management Study

A) Bar graphs showing the distribution of Centiloid (top) and global z-scores (bottom) for patients with SCD+, MCI, and dementia, color coded for visual read status (blue: negative; red: positive). As expected, the proportion of VR-positive cases increases with clinical stage. In addition, a bimodal distribution of both

Centiloid and z-score quantification can be appreciated across the cohort, though more clearly for MCI and dementia patients. **B)** Boxplot illustrating that the amount of amyloid burden in visually-positive cases increases depending on clinical stage. **C)** Scatterplot between Centiloid and z-score quantification, showing excellent correlation and a factor 10 relationship.

Ref: Data taken from AAIC 2022 Featured research session (manuscript in preparation).

Table 9 Overview of Participant Demographics and Clinical Features

	Study population	SCD+ *n*=220 (30.2%)		MCI *n*=252 (40.2%)		Dementia *n*=257 (29.6%)		*p*-value
		VR–	VR+	VR–	VR+	VR–	VR +	
	N=729 (100%)	*n* = 154 (70.0%)	*n* = 66 (30.0%)	*n* = 150 (51.2%)	*n* = 143 (48.8%)	*n* = 61 (28.2%)	*n* = 155 (71.8%)	
Age, years	72 [65-77]	67 [63-72.75]	71 [67-75]	70 [62-76.75]	74 [69-78]	73 [70-76]	75 [67.5-79]	p <0.001[a,c,d], p <0.05[b]
Sex, female (%)	325 (44.6)	67 (43.5)	27 (40.9)	59 (39.3)	66 (46.2)	20 (32.8)	86 (55.5)	p <0.05[a]
Education, years	13 [10-16]	15 [12-17]	13 [11-15.75]	12 [10-15]	13 [10-16]	12 [9-14]	12 [9-15]	p <0.001[a,c], p <0.05[b]
MMSE	27 [23-29]	29 [28-30]	29 [27-30]	27 [25-29]	26 [23-28]	24 [21.75-26.25]	21 [18-24.75]	p <0.001[a,b,c,d]
Etiological diagnosis*	346 (47.5)	0 (0.0)	43 (65.2)	15 (10.0)	125 (87.4)	13 (21.3)	150 (96.8)	p <0.001[a,c,d]
AD (%)	245 (33.6)	88 (57.1)	14 (21.2)	92 (61.3)	13 (9.1)	36 (59.0)	2 (1.3)	p <0.001[b,c,d]
Non-AD (%)	138 (18.9)	66 (42.9)	9 (13.6)	43 (28.7)	5 (3.5)	12 (19.7)	3 (1.9)	p <0.05[a], p < 0.001[b,c,d]
NA (%)								
Centiloid value	33.8 [0.727-84.9]	1.79 [–4.93-9.05]	71.1 [47.0-92.7]	1.22 [–7.04-15.3]	77.4 [60.7-107.2]	–2.71 [–11.4-12.0]	93.2 [73.9-117.5]	p <0.001[a,b,c,d]
Global Z-score	2.70 [0.038-7.86]	0.078 [–0.537-.843]	6.88 [4.60-8.39]	0.174 [–0.800-.972]	7.44 [5.55-9.73]	–0.373 [–1.23-.918]	8.85 [6.92-10.9]	p <0.001[a,b,c,d]

SCD+: subjective cognitive decline plus; MCI: mild cognitive impairment; VR–: negative visual read status; VR+: positive visual read status; AD: Alzheimer's disease; non-AD: non-Alzheimer's disease; NA: etiological diagnosis not yet achieved. Data is shown as the median value [interquartile range]. P-values are derived from either a binomial or Mann-Whitney U test. Only significant *p*-values are shown (i.e., p<0.05).

[a]comparing SCD+ and MCI groups

[b]comparing MCI and dementia groups

[c]comparing SCD+ and dementia groups

[d]comparing VR status (i.e., negative and positive) across clinical stages

*Final etiological diagnosis.

Ref: AMYPAD data from VUMC. Publication in Preparation

Summary Conclusion for Section B

Section B contains 5 pieces of strong evidence to demonstrate the utility of the Centiloid measure in both clinical and research (ie earlier) populations across the AD continuum. We present data from 4 independent cohorts (AMYPAD DPMS and PNHS as well as IDEAS and ABIDE) which show the negative scans are anchored at the CL=0 point, and the positive scans range from 40 to 80+ CL depending upon the clinical-subtype examined. Data from IDEAs used the CL measure to harmonise quantitative values across the three approved tracers to measure the consistency of local visual inspection results relative to quantitation.

The threshold of positivity for routine scans is approximately 21-24 for routine readers, with evidence that more experienced readers can be more sensitive for early depositing amyloid and read down to 17 Centiloids. Work from the AMYPAD DPMS study also shows the utility of the Z-score metric (which some users may have access to in lieu of the CL measure) as being approximately comparable to the CL measure (although requiring a normal data base to anchor the calculation).

6.4 Section C: Measuring longitudinal change in amyloid using the Centiloid Metric

This section describes two linked analyses regarding the longitudinal behaviour of Centiloid quantification as measured from both DPMS and PNHS and its possible use in clinical settings.

C1. Estimation of longitudinal within-subject variability

Relevance

Longitudinal PET-based imaging endpoints are commonly integrated in clinical trials of disease-modifying drugs for Alzheimer's disease (AD) and could become a key component of therapy response monitoring. We assessed the within-subject variability of Centiloid (CL) in the context of longitudinal studies acquired with either ^{18}F-flutemetamol (FMM) or ^{18}F-florbetaben (FBB).

Methods

Longitudinal amyloid-PET imaging from two cohorts was used. 107 participants from the Diagnostic and Patient Management Study (DPMS), which are representative of a memory clinic population were included (participants with subjective cognitive decline (SCD+, *N*=35), mild cognitive impairment (MCI, *N*=44) and dementia (*N*=28)). PET scans were performed at two time points using either FMM (*N*=49) or FBB (*N*=58) with an average scan time interval of 1.3 years. The second cohort consisted of a selection of cognitively unimpaired individuals with at least two PET scans from the Prognostic and Natural History Study (PNHS, *N*=680). This sample would be representative of an at-risk population which may benefit from prevention trials. PET scans were acquired with FMM (*N*=424) or FBB (*N*=256) and average scan time interval in this cohort was 2.1±0.3 year (follow-up 1) and 4.8±1.0 years (follow-up 2). Global CL quantification was performed using a PET-only pipeline for the DPMS while the PNHS used an MR-based pipeline. To capture the inherent measurement variability over time, a subset of individuals expected to be stable across the follow-up time was selected as a means to estimate test retest (DPMS inclusion criteria: SCD+ at both time points, baseline CL < 10; PNHS inclusion criteria: CL <10 and VR negative at both time points, CSF aβ42/40 or aβ42 and CSF ptau negative, ApoE ε4 non carrier). Estimates of the within- and between-subject variability were obtained through an Intraclass Correlation Coefficient (ICC) analysis. In addition, the expected change in the predicted stable subsets was estimated using the 95th percentile of the annualised rate of change (ARC).

Results

The subset expected to be stable comprised of 22 subjects from the DPMS and 46 subjects from the PNHS (Table 10). The within-subject variability was approximately ~3 CL for both cohorts (Table11). Moreover, the reliability analysis revealed that both the DPMS and PNHS have an ICC close to 1, indicating that the variability observed between the two scans is mostly attributable to an underlying difference in protein deposition across individuals rather than measurement error. The 95th percentile of ARC was below 5 CL/year across studies (4.8 CL/year in the DPMS; 3.9 CL/year in the PNHS). No significant difference was found across tracers.

Conclusion

The CL was found stable and robust across tracers with a longitudinal (i.e., test retest) variability estimated around ~3 CL/year, and very low likelihood of it surpassing 5 CL. Of relevance, no differences were detected in annualised rates of change of CL between tracers.

Table 10 Demographic characteristics of the DPMS and PNHS longitudinal cohorts

		DPMS		PNHS	
		All	'Stable' subset	All	'Stable' subset
N		107	22	680	46
Gender (% female)		44.9%	36.4%	60%	61%
Age (y (median (Q1-Q3)))		70 (64 - 76)	65 (61 - 70)	65 (60 - 70)	67 (63 - 69)
Education (y)		13.0 ± 3.9	15.0 ± 2.9	14.8 ± 4.1	15.0 ± 4.2
MMSE	at BL	25.8 ± 3.8	28.6 ± 1.3	29.2 ± 1.0	29.3 ± 0.9
	at FU	25.1 ± 4.8	28.3 ± 2.2		
Centiloid at BL		41.7 ± 47.7	-0.8 ± 4.7	12.9 ± 23.4	-1.2 ± 6.4
ARC (CL/year)		1.4 ± 6.4	-0.2 ± 3.3	1.2 ± 4.0	-0.3 ± 2.3
FU time (y)		1.3 ± 0.2	1.3 ± 0.3	2.1 to 4.8 ± 1.0	2.1 ± 0.1

Values are provided as mean ± SD, unless otherwise specified. Annualised rate of change (ARC) was computed as ($CL_{follow\text{-}up} - CL_{baseline}$)/dt with dt the time interval in between timepoints in years.

Abb - y = years, MMSE = Mini Mental State Examination, BL = baseline, FU = follow-up

Table 11 Repeatability and reliability of the Centiloid scale (CL), estimated with subjects expected to be stable over the follow-up time interval

		DMPS	PNHS
N		22	46
Sample Mean (CL)		-0.7	1.5
Sample SD (CL)		4.9	6.3
	RMS between subjects (CL)	7.0	8.9
	RMS within subjects (CL)	3.0	3.3
RMSD		4.2	4.7
ICC		0.82 [0.57, 0.92]	0.86 [0.75, 0.92]

Abb - ICC = intraclass correlation coefficient, RMS= root mean squares, RMSD= root-mean-square deviation

C2. A Centiloid window to help predict true amyloid accumulation

Relevance

Serial amyloid PET acquisitions enable the close monitoring of AD pathophysiology, in particular in the pre-symptomatic phase of the disease targeted in recent clinical trials. We assessed the ability of global Centiloid (CL) quantification to predict amyloid accumulation based on PET visual read (VR) and rates of amyloid accumulation in a cognitively unimpaired (CU) population.

Methods

Longitudinal amyloid-PET was performed on 686 healthy participants of the PNHS, using either ^{18}F-flutemetamol (FMM, *N*=427) or ^{18}F-florbetaben (FBB, *N*=259). Scans were acquired at 2 or 3 timepoints, with average follow-up times of 2.1 ± 0.3 and 4.8 ± 1.0 years respectively (Table 12). Quantification of scans was performed using an MR-based CL pipeline. Longitudinal change in CL was modelled using generalized estimating equations (corrected for age, sex, and ApoE ε4 carriership) based on categorizing subjects with different approaches. The first was based on visual read (VR), with subjects classified as Stable VR-, Converters or Stable VR+. The second approach was based on categorising subjects based on their baseline CL load, either Negative (CL<12), in a Grey-zone (12≤CL≤50) or Positive (CL>50). The third classification was based on whether subjects showed evidence of amyloid accumulation, based on their annualised rates of change (ARC) being above the within-subject variability of the projected stable negative subjects. Differences in trajectories between tracers was also investigated.

Results

First, we found three distinct longitudinal CL trajectories according to the VR groups. At baseline, 86% of cases were considered VR-, out of which 10% converted to VR+. Individuals in the Stable VR+ displayed the highest baseline CL load, followed by Converters and Stable VR- (Figure 25 A). Converters and Stable VR+ group showed a greater ARC than Stable VR- ($ARC_{Stable\ VR-}$= 0.3±3.1CL/year, $ARC_{Converters}$= 5.4±5.1CL/year, $ARC_{Stable\ VR+}$= 3.4±7.6CL/year, p<.001). When restricted to individuals with a baseline CL in the Grey-zone (12≤CL≤50), the patterns were similar ($\beta_{Stable\ VR-}$=0.3; $\beta_{Converters}$=4.9, p<.005) (Figure 25 B). Subjects considered Converters and subjects with CL in Grey-zone category showed a higher proportion ApoE ε4 carriers compared to the whole dataset. Furthermore, individuals considered ‘accumulators’

displayed a higher baseline CL value (36.3±28.8 CL versus 11.8±19.4 for 'Non accumulators', p<.001, Figure 25 C). Considering individuals with baseline VR- and CL in the Grey-zone, 30% will convert to amyloid positivity, while this number increases to 65% for subjects also found to be 'accumulators'. Finally, no significant differences in baseline CL value or in ARC was found between FBB and FMM (Figure 25 D).

Conclusion

The CL methodology is robust across tracers in a longitudinal setting. Baseline CL can help identify subjects more likely to accumulate pathology and could therefore assist subject selection and therapy response monitoring in clinical trials.

Table 12 Demographics characteristics of the PNHS

	PNHS		
	All	^{18}F-florbetaben	^{18}F-flutemetamol
N	686	259	427
Age (y (median (Q1-Q3)))	65 (60 - 70)	67 (61 - 71)	64 (60 - 69)
Gender (% female)	60	61	60
ApoE ε4 (% carriers)	41	33	46
Education (y)	14.8 ± 4.1	15.5 ± 4.2	14.4 ± 3.8
MMSE at BL	29.2 ± 1.0	29.2 ± 1.1	29.2 ± 1.0
***N* Timepoints**	2 (*N*=583) 3 (*N*=103)	2 (*N*=156) 3 (*N*=103)	2
Centiloid at BL	12.9 ± 23.4	14.1 (24.2)	15.5 (21.0)
ARC	1.2 ± 4.0	1.3 (4.6)	1.2 (3.7)
FU time (y)	FU1: 2.1 ± 0.3 FU2: 4.8 ± 1.0	FU1: 2.1 ± 0.2 FU2: 4.8 ± 1.0	2.0 ± 0.4 years

Values are provided as mean ± SD, unless otherwise specified. Annualised rate of change (ARC) was computed as ($CL_{follow\text{-}up} - CL_{baseline}$)/dt with dt the time interval in between timepoints in years.
Abb - y = years, MMSE = Mini Mental State Examination, BL = baseline, FU = follow-up

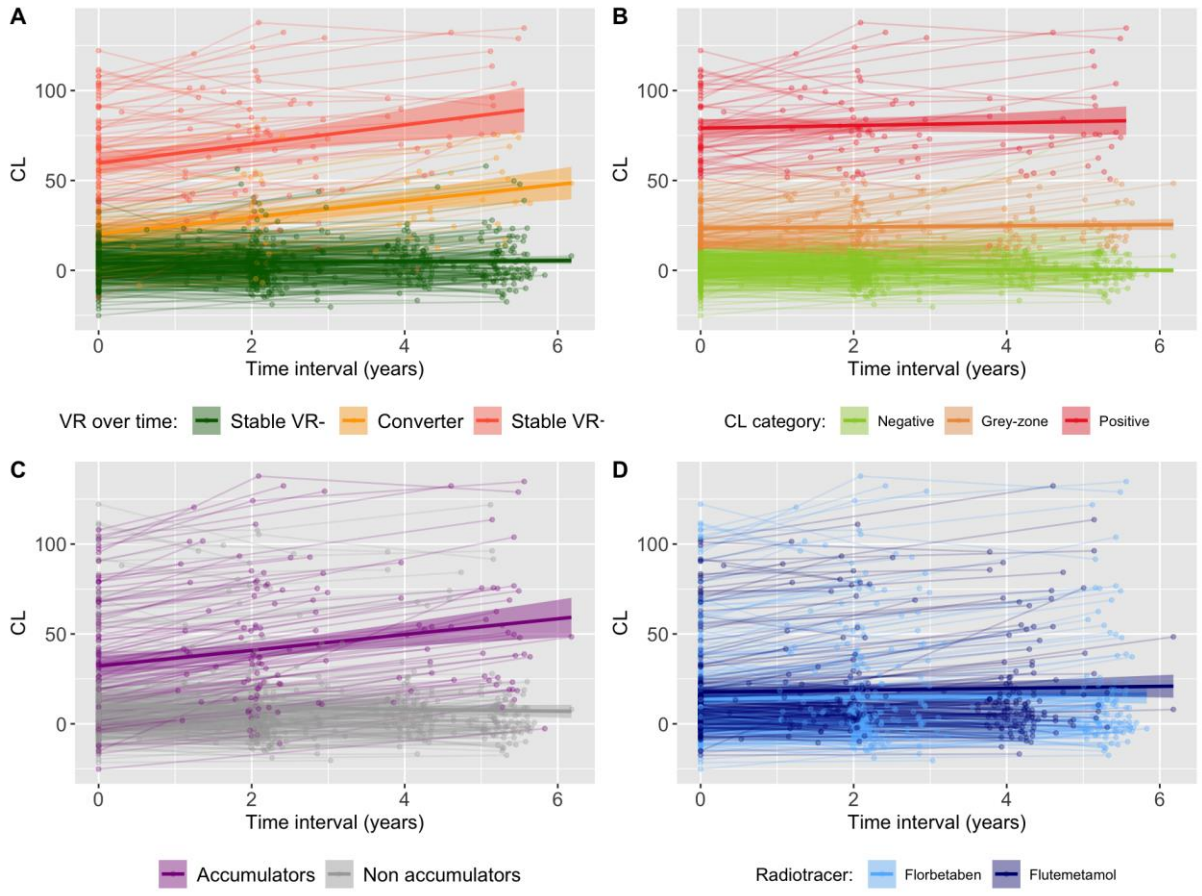


Figure 25 Longitudinal Trajectories of Amyloid Accumulation

A) Trajectories based on visual assessment, subjects were classified as Stable VR- (*N*=332, VR- VR-), Converter (*N*=42, VR- to VR+) or Stable VR+ (*N*=61, VR+ to VR+); **B)** Trajectories based on baseline CL category, established with the following cut-offs: CL negative (CL<12), grey-zone (12≤CL≤50) and positive (CL>50); **C)** Trajectories based on the annualised rate of change (ARC) above expected measurement variability, 'accumulators' are defined as individuals with an ARC ≥ 3.3 CL/year; **D)** Trajectories per tracer

Overall Conclusions Section C

Longitudinal analysis of DPMS and PNHS data has indicated several important factors which are relevant for understanding the behaviour of PET tracers over time. The CL was found stable and robust across tracers with a within-subject (i.e., test retest)

variability estimated around ~3 CL/year, and very low likelihood of it surpassing 5 CL. Analysis has also shown there is a Centiloid window of approximately 12-50 CL units which is predictive of those subjects where a significant rise in amyloid load might be expected (compared to those in the 10 CL region where the amyloid load is unchanged). Importantly, results were consistent irrespective of tracer and pipeline used, thus confirming that the estimation of the rates of change of amyloid load as measured with the CL method are independent of the tracer used.

7 IMAGING METHODOLOGY FOR ASSESSMENT OF CENTILOID MEASURES

Section 7 in this BQO is a reference to the general guidelines for the use (camera set up, image acquisition, image reconstruction etc) of the amyloid PET tracers.

7.1 Amyloid PET tracers

Two amyloid PET tracers Vizamyl [^{18}F]flutemetamol (c/o GE Healthcare, Amersham UK) and Neuraceq [^{18}F]florbetaben (c/o Life Radiopharma Berlin GmbH, Berlin, Germany) have been employed in the AMYPAD studies. Both tracers have been routinely available in Europe since their marketing authorization in 2014.

Further details on the Vizamyl product can be found in the current EU SmPC.

https://www.ema.europa.eu/en/documents/product-information/vizamyl-epar-product-information_en.pdf

Further details on the NeuraCeq product can be found in the current EU SmPC.

https://www.ema.europa.eu/en/documents/product-information/neuraceq-epar-product-information_en.pdf

Notably, both tracers, as well as Amyvid [^{18}F]florbetapir, have language in the SmPC that allows quantitation as an adjunct to visual reads. This language was accepted after data was submitted to EMA validating the robustness of quantitation by various methodologies and supports the use of quantitation as a adjunct to visual reading.

7.2 Imaging Guidelines

A summary of the Imaging Guidelines for Vizamyl and NeuraCeq can be found in Appendices D and E.

7.3 RSNA QIBA Profile for Amyloid PET as an Imaging Biomarker for Cerebral Amyloid Quantification

This document (Smith et al 2022) is a very recent and comprehensive overview which aims to facilitate a standardised approach to the use of F-18 labelled amyloid PET tracers across PET centres. The paper focuses on the SUVr unit as the central metric of choice although acknowledges that the approach is applicable to the Centiloid Unit too.

The context of use is similar to that covered in this BQO dossier (ie the objective measure of longitudinal change, staging and prediction of cognitive changes). The paper highlights the importance of the choice of reference region. The key message in this QIBA profile relates to a justification of a within-subject coefficient of variation (wCV) of ≤1.94% when measuring brain amyloid load when paying attention to tracer, scanner, acquisition and analysis protocols.

7.4 Tips from AMYPAD for optimal image acquisition/image processing/reconstruction

The standard CL method is highly robust against difference in image resolution when using the Whole Cerebellum and the Global Cortical Target VOI provided in the GAAIN website. As long as this reference region and target regions are being used in the CL method, a wide range of clinical reconstruction parameters are not expected to render significantly different CL values. Nevertheless, best practice for optimal accuracy is to use a quantitative methodology based on Hoffmann phantom acquisitions (e.g. https://pubmed.ncbi.nlm.nih.gov/26249138/).

Similar considerations apply to multicentre studies. The CL method provides very robust values irrespective of reconstruction parameters in multi-centre studies in which the image resolution has not been prospectively harmonized, as long as the standard target and Whole Cerebellum VOIs are used. Best practice for optimal quantitative results involves the prospective harmonization of reconstruction parameters across sites to obtain comparable image quality, using existing methods for this purpose (https://pubmed.ncbi.nlm.nih.gov/33517517/).

7.5 Training for Image Interpretation

Training for Image Interpretation can be found at the following sites

https://www.readvizamyl.com/

http://www.neuraceqreadertraining.com/learn

7.6 Software Tools where the Centiloid Unit is/will be available

- MIM Neuro MIM Software's *MIMneuro* (https://www.mimsoftware.com/nuclear_medicine/mim_neuro)

- Syntermed NeuroQ (https://syntermed.com/neuroq). Release including Centiloids planned for SNMMI 2023
- cPET (part of cNeuro suite) c/o Combinostics (https://www.combinostics.com/) Centiloid/SUVr/Z-scores. Software CE approvals aimed for end 4Q22 and 510(k) clearance in 1Q23.
- rPOP (https://github.com/leoiacca/rPOP) c/o the UCSF Neurology group
- NiftyPET/AmyPET (https://github.com/AMYPAD/AmyPET) – available but still being refined for research use

7.7 Technical information on the Centiloid Unit

The use of [^{11}C]PiB is accepted as the "Gold Standard" for amyloid imaging. PiB utility in routine clinical application however is limited by the short half-life of the Carbon-11 radioisotope and the requirement for an on-site cyclotron.

F-18 Tracers have longer radioactive half-life (110 minutes) compared to PiB and offer greater clinical utility in amyloid imaging. Three tracers; [^{18}F]florbetapir (Amyvid), [^{18}F]florbetaben (NeuraCeq) and [^{18}F]flutemetamol (Vizamyl) have all been validated through post-mortem studies and are approved by regulatory authorities.

These tracers all show increased cortical retention in AD subjects. However, each tracer has its own unique set of cortical and reference regions, methods for visual evaluation and positive PET cut-off points associated with quantitative use. Furthermore, differences in their dynamic ranges, kinetics and non-amyloid white matter binding all add up to make comparing data sets across studies, and across groups scanned with more than one of these tracers, challenging.

As such, there has been an unmet need to standardise quantitative amyloid imaging. Standardized units would also allow better interpretation of longitudinal changes, improving how sites monitor disease progression, and any potential therapeutic effects.

The Centiloid Scale, conceived by Prof. WE Klunk and colleagues, was proposed as a means of comparing different β-amyloid imaging tracers with different kinetics in AD imaging[1]. They developed the Centiloid Atlas for PET image quantification, based on PiB uptake in young healthy controls (YC) and typical AD patients. Essentially, VOIs templates are applied to co-registered PET and MR images using the software package SPM8. SUVRs can then be derived for the cortical structures and other key regions, using the whole cerebellum as the refence region. These SUVR values are then converted to Centiloid values (CL). On the scale YC Subjects have a mean value 0 CL, and AD Subjects have a mean value of 100 CL.

The Centiloid methodology is a multi-step process, essentially split in to 2 parts: Validation of Pipeline (Figure 26, Left) and Application to tracers (Figure 26, Right). There are a number of "Stop / GO" points along the process ensure that Centiloid scaling is being correctly applied.

Firstly, in order to validate their pipelines, each Site must obtain the PiB PET and MR images and Centiloid VOIs from the Global Alzheimer's Association Information Network (GAAIN, www.gaain.org/centiloid-project) and follow the steps described by Klunk *et al*. to apply the VOI atlas. Site must ensure their PiB SUVRs correlate well (Slope = 0.98-1.02, R^2 >0.98) with the published values, and the mean SUVR are within tolerance (±2%) for the YC-0 and AD-100 groups.

PiB SUVRs are then converted to Centiloids using the following equation:

$$CL=100X\ (^{PiB}SUVR_{IND} - {}^{PiB}SUVR_{YC\text{-}0}) / (^{PiB}SUVR_{AD\text{-}100} - {}^{PiB}SUVR_{YC\text{-}0})$$

Where:
$^{PiB}SUVR_{IND}$ = individual's PiB SUVR value
$^{PiB}SUVR_{YC\text{-}0}$ = MEAN YC-0 PiB SUVR value
$^{PiB}SUVR_{AD\text{-}100}$ = MEAN AD-100 PiB SUVR value

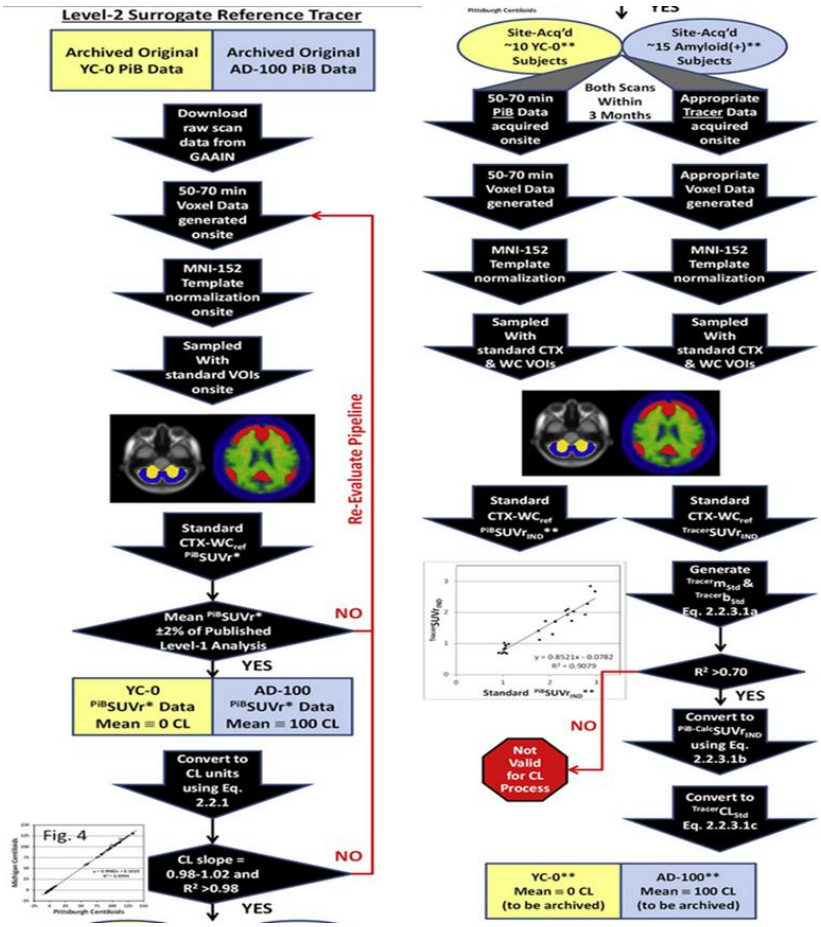


Figure 26 Centiloid Process for the Validation of Process Pipeline (left) and Application to Tracer (right) Ref: [Klunk et al. 2015]

The second part of the process is the application of the methods to other tracers. Site-derived PET and MR images for both PiB and other 18F-labeled tracers obtained in the same subject can be analysed using same concept of VOI application, with the added step of scaling the tracer SUVR values to PiB-equivalent SUVRs for conversion to Centiloid. The correlation between the PiB and tracer SUVRs is used to calculate the following equations:

PiB-Calculated SUVRs: $^{PiB\text{-}Calc}SUVR_{IND} = (^{Tracer}SUVR_{IND} - {}^{Tracer}b_{Std}) / {}^{Tracer}m_{Std}$

Where:
$^{Tracer}SUVR_{IND}$ = individual's TRACER SUVr value
$^{Tracer}b_{Std}$ = intercept value (AD & HC subjects)
$^{Tracer}m_{Std}$ = slope value (AD & HC subjects)
$^{PiB\text{-}Calc}SUVR_{IND}$ = converted $^{Tracer}SUVR_{IND}$ value into PiB value

Centiloids are then calculated using the following equation:

Centiloid Calculation: $^{Tracer}CL_{Std} = 100X (^{PiB\text{-}Calc}SUVR_{IND} - {}^{PiB}SUVR_{YC\text{-}0}) / (^{PiB}SUVR_{AD\text{-}100} - {}^{PiB}SUVR_{YC\text{-}0})$

Where:
$^{Tracer}Cl_{Std}$ = TRACER Centiloid value
Following these methods groups have published Centiloid equations for 3 tracers:

Flutemetamol CL[2] = (121.42*SUVR-Flute) - 121.16

Florbetapir Centiloids[3] = $183 \times SUVr_{Avid} - 177$

Florbetaben Centiloids[4]: CL units = $153.4 \times SUVR_{FBB} - 154.9$

8 OVERALL CONCLUSIONS

AMYPAD has presented in this document a wealth of data to demonstrate that the Centiloid metric is a credible, versatile and robust measure of brain amyloid burden. As the use of amyloid PET becomes more omnipresent in the Alzheimers field it is important that methodologies that are well utilised in both clinical development and academic research are fully characterised prior to their use in clinical routine. We present here three distinct categories of work where the Centiloid measure is:

a) tested for performance robustness

b) assessed cross sectionally different clinical scenarios as an initial measure of brain amyloid burden and

c) measured for longitudinal follow up where our AMYPAD studies have provided us with the opportunity to perform the both the baseline and follow up scans.

In addition to support this application we have reviewed in detail and produced our own state of the art review (published in the European Journal of Nuclear Medicine and Molecular Imaging) to cover the ever-increasing literature in the Centiloid field (now over 90 papers as of September 2022) which demonstrate the universality of the methodology across multiple clinical cohorts and clinical development programs. We have worked closely with software vendors to support their own implementation of the Centiloid measure in quantitative amyloid tools which aim to support the routine physician in his measurement of brain amyloid burden. The Centiloid measure provides a consistent and accurate means to assess brain amyloid across tracers, sites, pipelines etc with the ultimate aim to better serve AD patients with consistent decision making around diagnostic and management decisions. AMYPAD has also been invited to contribute it's findings to further the understanding of the use of the Centiloid metric in a meeting convened by FDA in November 2022 which is dedicated to the assessment of quantitative amyloid PET methodology.

9 QUESTIONS FOR SAWP

Question 1

Does EMA agree it valuable to have a single universal metric that is tracer independent to measure amyloid burden in the brain?

Q1: AMYPAD Position:

Determination of brain amyloid load by determining SUVr in specific brain regions is currently the most widely used and established metric (Kinahan & Fletcher, 2010), having been implemented in several recent trials to assess treatment efficacy (Doody et al., 2013, Doody et al.,2014; Honig et al., 2018; Liu et al., 2015; Ostrowitzki et al., 2017; Relkin et al., 2017; Salloway et al., 2014). However, accurate measurement and cut-off values are highly dependent on the chosen tracer, reference region and delineation method (Klunk et al., 2015; S. M. Landau et al., 2014; Tryputsen et al., 2015), which challenges the pooling of multi-centre SUVr data across tracers (Kolinger et al., 2021). In addition, there is high variability in longitudinal results (Susan M. Landau et al., 2015; Tryputsen et al., 2015), which limits the power in detecting genuine biological differences. Therefore, as the use of different amyloid PET tracers has grown in both clinical and research settings, there has been a concurrent need for inter-tracer standardization of the SUVr metric in both multi-centre collaborations as well as for the general understanding of measures of amyloid burden in routine settings where potentially diagnostic use may increase with the introduction of novel therapies.

Question 2

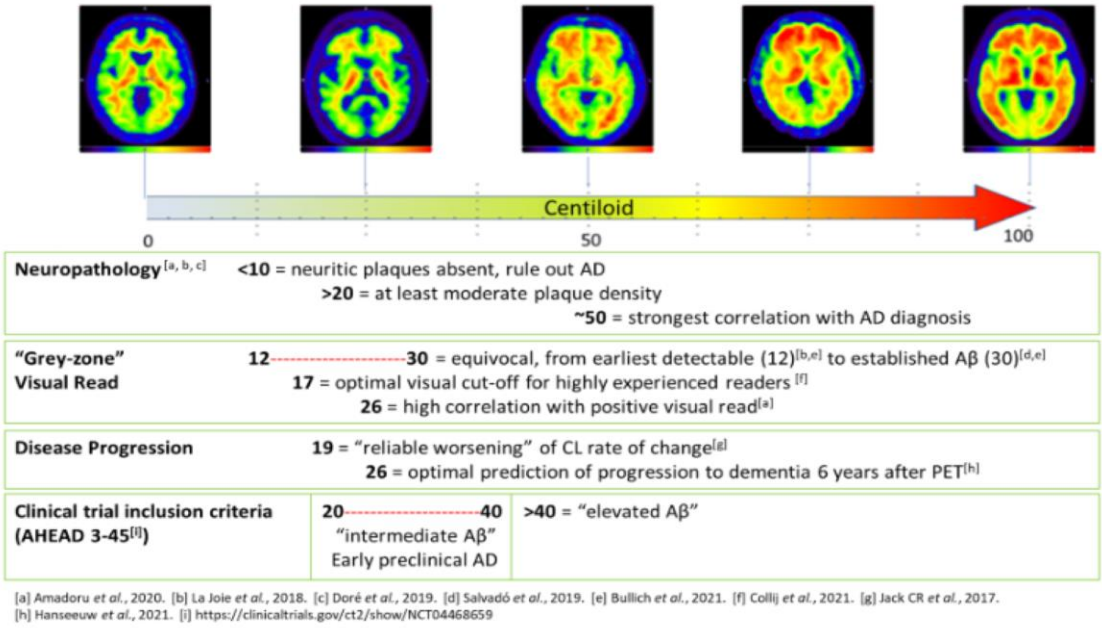
Does EMA agree that the Centiloid metric is suitable for measuring amyloid burden in the brain?

Q2: AMYPAD Position:

In order to address the limitations mentioned above in Question 1, the CL scale was developed by Klunk et al 2015. The main aims of the CL scale were to: (i) simplify and expedite direct comparison of Aβ PET results across sites and studies; (ii) outline the earliest thresholds for amyloid positivity and define the range of positivity in AD; (iii) robustly quantify longitudinal change; and (iv) facilitate inter-tracer comparisons (Klunk et al., 2015). Since then, several studies have tested the scale's validity and used it to improve the harmonisation and standardisation of Aβ PET quantification across tracers,

scanners, and analytical implementations (Battle et al., 2018; Bourgeat et al., 2021, Bourgeat et al.,2018; Cho, Choe, Kim, et al., 2020; Cho, Choe, Park, et al., 2020; Leuzy et al., 2016; Navitsky et al., 2018; Rowe et al., 2017, Rowe et al., 2016; Schwarz et al., 2018; Su et al., 2018, Su et al., 2019; Tudorascu et al., 2018; Yun et al., 2017).

One of the key advantages of a 'universal' metric of amyloid burden is generalisation of quantitative thresholds across tracers and pipeline implementations. Universal cut-off or threshold values to denote amyloid status can be applied alongside visual reads and in longitudinal multi-centre studies to facilitate inter-centre and inter-tracer comparisons. Various cut-offs established in the literature are summarised in the figure below.



[a] Amadoru *et al.*, 2020. [b] La Joie *et al.*, 2018. [c] Doré *et al.*, 2019. [d] Salvadó *et al.*, 2019. [e] Bullich *et al.*, 2021. [f] Collij *et al.*, 2021. [g] Jack CR *et al.*, 2017. [h] Hanseeuw *et al.*, 2021. [i] https://clinicaltrials.gov/ct2/show/NCT04468659

Figure 27 ***Summary of the various CL thresholds established in the literature and in use for clinical trial inclusion***

Question 3

Does EMA concur that the Centiloid measure has been sufficiently characterised for use in both research and clinical applications?

Q3: AMYPAD Position

AMYPAD considers that the Centiloid metric is a robust methodology for the measurement of both early and established amyloid in the brain and AMYPAD has generated significant evidence to demonstrate the robustness of the methodology by multiple means highlighted in Chapter 6 of this BQO dossier.

AMYPAD believes it has adequately demonstrated that the Centiloid measure can be considered as a valid tracer independent means to assess amyloid burden in the brain across the full dynamic range of the measure and in longitudinal settings. We have shown in multiple cohorts that the anchor point in the negative scans is repeatedly observed at CL zero as predicted when the methodology was originally developed. Additionally, we have demonstrated that the Centiloid metric can be assessed in multiple pipeline analysis variants where different volumes of interest/reference regions/native vs standard space etc. The CL outcome measure remains steady indicating that during use across multiple sites and reconstruction settings the value (and hence utility) will remain constant. The test-retest variability of the Cl measure is low and the minor variations (<5CL units) that are seen across pipeline designs are approximately similar to this variability.

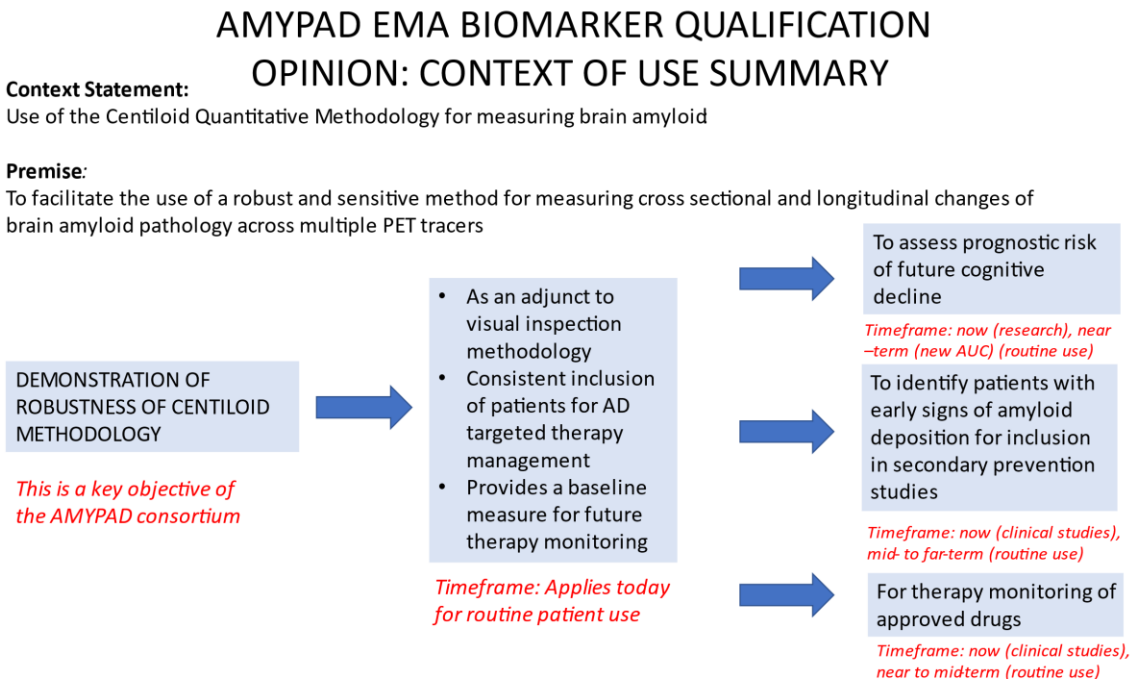
Question 4:

Does EMA concur that the body of evidence provided by AMYPAD supports the diverse utility of the Centiloid metric as a means for example to (i) support the current visual inspection of tracers as an adjunctive tool, (ii) for the consistent inclusion of patients for AD targeted therapies and (iii) to provide a potential baseline measure for future therapy monitoring/follow up scanning as indicated in the context of use summary?

Q4: AMYPAD Position.

AMYPAD has presented strong evidence to demonstrate the value and wide utility of the Centiloid measure in both clinical and research (ie earlier) populations across the AD continuum. We show data from 4 independent cohorts (AMYPAD DPMS and PNHS as well as IDEAS and ABIDE) which confirm that the negative scans are anchored at the CL=0 point and the positive scans range from 40 to 80+ depending upon the clinical-subtype examined. Data from IDEAs (including a large population of florbetapir scans) used the CL measure to harmonise quantitative values across the three approved tracers so as to measure the consistency of local visual inspection results relative to quantitation.

Hence, we consider the CL metric is a valuable means to supplement visual read methodology and it has the added value of being a continuous measure with the potential for further applications beyond supporting an initial dichotomous decision of diagnosis. The potential for future therapy monitoring of any possible approved amyloid targeted therapies relies upon both the baseline and follow up scans being performed in a consistent and reliable way to ensure optimal patient outcome.



In combination with the current body of literature AMYPAD has demonstrated the utility of the CL measure to go beyond the ability to support an initial confirmation of the amyloid status of a patient seeking therapy or a subject being included into a clinical trial.

The CL can measure earlier deposition of pathological amyloid in the CERAD 'sparse' band and therefore could be a valuable tool for targeting preclinical AD individuals who are clinically normal but have pathological signs indicating that the disease process is underway. Multiple secondary prevention studies are now underway to examine this concept.

To support this premise,the CL approach has been validated against neuropathology (Amadoru et al., 2020; La Joie et al., 2019) where CL<10 correlates with absence of neuritic plaques, CL>20 specified at least moderate plaque density, and >50 CL best confirmed both neuropathological and clinicopathological evidence of AD.

Predictive models using the CL scale have also been developed for calculating rate of cognitive decline in cognitively normal subjects (Farrell et al., 2018, Farrell et al., 2021; van der Kall et al., 2021). In addition, Hanseeuw et al., 2021 found that a CL threshold of 26 in memory clinic patients optimally predicts progression to dementia 6 years after PET (Hanseeuw et al., 2021). Work continues in various global cohorts (including AMYPAD) to consolidate these observations relating to prognostic value.

In clinical trial settings, quantification may also be used to identify the optimal window for therapeutic intervention (Bischof & Jacobs, 2019). This is illustrated by the AHEAD 3-45 Study, which requires participants to have specific levels of amyloid pathology, either 'intermediate' (20-40 CL) or 'elevated' (>40 CL), signifying the added value beyond binary classifications (Aisen et al., 2020). As widely reported the CL scale has been used in clinical trial settings to track therapy response measure (Bateman et al., 2020; Klein et al., 2021, ; Klein et al., 2019; Mintun et al., 2021; Roberts et al., 2021; Salloway et al., 2021), determine strategies for reducing AD prevention trial sample sizes (Lopes Alves et al., 2021) and improve patient selection for trials (Knopman et al., 2021; Weiner et al., 2017) and could assist in treatment endpoint decisions (Lopes Alves et al., 2020).

10 REFERENCES (available upon request)

Aisen, P. S., Zhou, J., Irizarry, M. C., Kramer, L. D., Swanson, C. J., Dhadda, S., … Sperling, R. A. (2020). AHEAD 3-45 study design: A global study to evaluate the efficacy and safety of treatment with BAN2401 for 216 weeks in preclinical Alzheimer's disease with intermediate amyloid (A3 trial) and elevated amyloid (A45 trial). *Alzheimer's & Dementia*, *16*(S9), e044511. https://doi.org/10.1002/alz.044511

Altomare, D., Collij, L., Caprioglio, C., Scheltens, P., van Berckel, B.N.M., … Frisoni, G.B., AMYPAD Consortium (2022). Description of a European memory clinic cohort undergoing amyloid-PET: The AMYPAD Diagnostic and Patient Management Study. *Alzheimers Dement*. doi: 10.1002/alz.12696.

Amadoru, S., Doré, V., McLean, C. A., Hinton, F., Shepherd, C. E., Halliday, G. M., … Rowe, C. C. (2020). Comparison of amyloid PET measured in Centiloid units with neuropathological findings in Alzheimer's disease. Alzheimer's Research and Therapy, 12(1), 1–8. https://doi.org/10.1186/s13195-020-00587-5

Barthel H, Sabri O. (2017). Clinical Use and Utility of Amyloid Imaging. J Nucl Med. Nov;58(11):1711-1717

Bateman, R. J., Aschenbrenner, A. J., Benzinger, T. L. S., Clifford, D., Coalier, K., Cruchaga, C., … Kerchner, G. A. (2020). Overview of dominantly inherited AD and top-line DIAN-TU results of solanezumab and gantenerumab. Alzheimer's & Dementia, 16(S9), e041129. https://doi.org/10.1002/alz.041129

Battle, M. R., Pillay, L. C., Lowe, V. J., Knopman, D., Kemp, B., Rowe, C. C., … Buckley, C. J. (2018). Centiloid scaling for quantification of brain amyloid with [18 F]flutemetamol using multiple processing methods. EJNMMI Research, 8(1), 1–11. https://doi.org/10.1186/s13550-018-0456-7

Biogen Inc (2021). Prescriber Information for Aduhelm-aducanumab injection, solution (BLA 761178)

Bischof, G. N., & Jacobs, H. I. L. (2019). Subthreshold amyloid and its biological and clinical meaning: Long way ahead. Neurology, 93(2), 72–79. https://doi.org/10.1212/WNL.0000000000007747

Bourgeat, P., Doré, V., Doecke, J., Ames, D., Masters, C. L., Rowe, C. C., … Villemagne, V. L. (2021). Non-negative matrix factorisation improves Centiloid robustness in longitudinal studies. NeuroImage, 226.

https://doi.org/10.1016/j.neuroimage.2020.117593

Bourgeat, P., Doré, V., Fripp, J., Ames, D., Masters, C. L., Salvado, O., … Rowe, C. C. (2018). Implementing the centiloid transformation for 11C-PiB and β-amyloid 18F-PET tracers using CapAIBL. NeuroImage, 183, 387–393. https://doi.org/10.1016/j.neuroimage.2018.08.044

Buckley, C. J., Foley, C., Battle, M., Grecchi, E., Farrar, G., Gispert, J., … Modat, M. (2019). AmyPype: An automated system to quantify AMYPAD's [18F]flutemetamol and [18F]florbetaben images including regional SUVR and Centiloid analysis. European Journal of Nuclear Medicine and Molecular Imaging, 46(1), S323–S324. Retrieved from https://www.embase.com/search/results?subaction=viewrecord&id=L629699161&from=export%0Ahttp://dx.doi.org/10.1007/s00259-019-04486-2

Bullich, S., Roé-Vellvé, N., Marquié, M., Landau, S. M., Barthel, H., Villemagne, V. L., … Boada, M. (2021). Early detection of amyloid load using 18F-florbetaben PET. Alzheimer's Research and Therapy, 13(1). https://doi.org/10.1186/s13195-021-00807-6

Cho, S. H., Choe, Y. S., Kim, H. J., Jang, H., Kim, Y., Kim, S. E., … Seo, S. W. (2020). A new Centiloid method for 18F-florbetaben and 18F-flutemetamol PET without conversion to PiB. European Journal of Nuclear Medicine and Molecular Imaging, 47(8), 1938–1948. https://doi.org/10.1007/s00259-019-04596-x

Cho, S. H., Choe, Y. S., Park, S., Kim, Y. J., Kim, H. J., Jang, H., … Seo, S. W. (2020). Appropriate reference region selection of 18F-florbetaben and 18F-flutemetamol beta-amyloid PET expressed in Centiloid. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-70978-z

Clark, C. M., Pontecorvo, M. J., Beach, T. G., Bedell, B. J., Coleman, R. E., Doraiswamy, P. M., … Skovronsky, D. M. (2012). Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: A prospective cohort study. The Lancet Neurology, 11(8), 669–678. https://doi.org/10.1016/S1474-4422(12)70142-4

Collij, L. E., Salvadó, G., Shekari, M., Lopes Alves, I., Reimand, J., Wink, A. M., … Gispert, J. D. (2021). Visual assessment of [18F]flutemetamol PET images can detect early amyloid pathology and grade its extent. European Journal of Nuclear Medicine and Molecular Imaging, 48(7), 2169–2182. https://doi.org/10.1007/s00259-

020-05174-2

de Wilde, A., Ossenkoppele, R., Pelkmans, W., Bouwman, F., Groot, C., van Maurik, I., … van der Flier, W.M. (2019). Assessment of the appropriate use criteria for amyloid PET in an unselected memory clinic cohort: The ABIDE project. *Alzheimer's & Dementia*, 15: 1458-1467. https://doi.org/10.1016/j.jalz.2019.07.003

Doody, R. S., Raman, R., Farlow, M., Iwatsubo, T., Vellas, B., Joffe, S., … Mohs, R. (2013). A Phase 3 Trial of Semagacestat for Treatment of Alzheimer's Disease. New England Journal of Medicine, 369(4), 341–350. https://doi.org/10.1056/nejmoa1210951

Doody, R. S., Thomas, R. G., Farlow, M., Iwatsubo, T., Vellas, B., Joffe, S., … Mohs, R. (2014). Phase 3 Trials of Solanezumab for Mild-to-Moderate Alzheimer's Disease. New England Journal of Medicine, 370(4), 311–321. https://doi.org/10.1056/nejmoa1312889

Doré, V., Bullich, S., Rowe, C.C., Bourgeat, P., Konate, S., … De Santi, S. (2019). Comparison of 18F-florbetaben quantification results using the standard Centiloid, MR-based, and MR-less CapAIBL® approaches: Validation against histopathology. *Alzheimers Dement*. 15(6):807-816. doi: 10.1016/j.jalz.2019.02.005.

Fantoni, E.R., Chalkidou, A., O' Brien, J.T., Farrar, G., Hammers, A. (2018). A Systematic Review and Aggregated Analysis on the Impact of Amyloid PET Brain Imaging on the Diagnosis, Diagnostic Confidence, and Management of Patients being Evaluated for Alzheimer's Disease. *J Alzheimers Dis*. 63(2):783-796. doi: 10.3233/JAD-171093.

Farrell, M. E., Chen, X., Rundle, M. M., Chan, M. Y., Wig, G. S., & Park, D. C. (2018). Regional amyloid accumulation and cognitive decline in initially amyloid-negative adults. Neurology, 91(19), E1809–E1821. https://doi.org/10.1212/WNL.0000000000006469

Farrell, M. E., Jiang, S., Schultz, A. P., Properzi, M. J., Price, J. C., Becker, J. A., … Buckley, R. F. (2021). Defining the Lowest Threshold for Amyloid-PET to Predict Future Cognitive Decline and Amyloid Accumulation. Neurology, 96(4), e619–e631. https://doi.org/10.1212/WNL.0000000000011214

Grabowski TJ, Frank RJ, Szumski NR, Brown CK, Damasio, (2000). Validation of partial tissue segmentation of single-channel magnetic resonance images of the brain. Neuroimage,12(6):640-56. doi: 10.1006/nimg.2000.0649.

Hanseeuw, B. J., Malotaux, V., Dricot, L., Quenon, L., Sznajer, Y., Cerman, J., … Lhommel, R. (2021). Defining a Centiloid scale threshold predicting long-term progression to dementia in patients attending the memory clinic: an [18F] flutemetamol amyloid PET study. European Journal of Nuclear Medicine and Molecular Imaging, 48(1), 302–310. https://doi.org/10.1007/s00259-020-04942-4

Honig, L. S., Vellas, B., Woodward, M., Boada, M., Bullock, R., Borrie, M., … Siemers, E. (2018). Trial of Solanezumab for Mild Dementia Due to Alzheimer's Disease. New England Journal of Medicine, 378(4), 321–330. https://doi.org/10.1056/nejmoa1705971

Iaccarino, L., La Joie, R., Koeppe, R., Siegel, B.A., Hillner, B.E., … Rabinovici, G.D., Alzheimer's Disease Neuroimaging Initiative. (2022) rPOP: Robust PET-only processing of community acquired heterogeneous amyloid-PET data. *Neuroimage*. 1;246:118775. doi: 10.1016/j.neuroimage.2021.118775.

Jack, C. R., Wiste, H. J., Weigand, S. D., Therneau, T. M., Knopman, D. S., Lowe, V., … Petersen, R. C. (2017). Age-specific and sex-specific prevalence of cerebral β-amyloidosis, tauopathy, and neurodegeneration in cognitively unimpaired individuals aged 50–95 years: a cross-sectional study. The Lancet Neurology, 16(6), 435–444. https://doi.org/10.1016/S1474-4422(17)30077-7

Jack, C. R., Wiste, H. J., Weigand, S. D., Therneau, T. M., Lowe, V. J., Knopman, D. S., … Petersen, R. C. (2017). Defining imaging biomarker cut points for brain aging and Alzheimer's disease. Alzheimer's and Dementia, 13(3), 205–216. https://doi.org/10.1016/j.jalz.2016.08.005

Jack, C.R. Jr., Wiste H.J., Botha, H., Weigand, S.D., Therneau, T.M., … Petersen, R.C. (2019). The bivariate distribution of amyloid-β and tau: relationship with established neurocognitive clinical syndromes, *Brain*, 142:10, 3230–3242. https://doi.org/10.1093/brain/awz268

Johnson KA, Minoshima S, Bohnen NI, Donohoe KJ, Foster NL, Herscovitch P, Karlawish JH, Rowe CC, Carrillo MC, Hartley DM, Hedrick S, Pappas V, Thies WH.J (2013). Appropriate use criteria for amyloid PET: a report of the Amyloid Imaging Task Force, the Society of Nuclear Medicine and Molecular Imaging, and the Alzheimer's Association. Nucl Med. 2013 Mar;54(3):476-90. doi: 10.2967/jnumed.113.120618.

Kim, Y., Rosenberg, P., Oh, E. (2018). A Review of Diagnostic Impact of Amyloid

Positron Emission Tomography Imaging in Clinical Practice. *Dement Geriatr Cogn Disord*. 46(3-4):154-167. doi: 10.1159/000492151.

Kinahan, P. E., & Fletcher, J. W. (2010). Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. Seminars in Ultrasound, CT and MRI, 31(6), 496–505. https://doi.org/10.1053/j.sult.2010.10.001

Klein, G., Delmar, P., Kerchner, G. A., Hofmann, C., Abi-Saab, D., Davis, A., … Doody, R. (2021). Thirty-Six-Month Amyloid Positron Emission Tomography Results Show Continued Reduction in Amyloid Burden with Subcutaneous Gantenerumab. Journal of Prevention of Alzheimer's Disease, 8(1), 3–6. https://doi.org/10.14283/jpad.2020.68

Klein, G., Delmar, P., Voyle, N., Rehal, S., Hofmann, C., Abi-Saab, D., … Doody, R. (2019). Gantenerumab reduces amyloid-β plaques in patients with prodromal to moderate Alzheimer's disease: A PET substudy interim analysis. Alzheimer's Research and Therapy, 11(1). https://doi.org/10.1186/s13195-019-0559-z

Klunk, W. E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D. P., … Långström, B. (2004). Imaging Brain Amyloid in Alzheimer's Disease with Pittsburgh Compound-B. Annals of Neurology, 55(3), 306–319. https://doi.org/10.1002/ana.20009

Klunk, W. E., Koeppe, R. A., Price, J. C., Benzinger, T. L., Devous, M. D., Jagust, W. J., … Mintun, M. A. (2015). The Centiloid project: Standardizing quantitative amyloid plaque estimation by PET. Alzheimer's and Dementia, 11(1), 1-15.e4. https://doi.org/10.1016/j.jalz.2014.07.003

Klunk, W.E. (2011). Amyloid imaging as a biomarker for cerebral β-amyloidosis and risk prediction for Alzheimer dementia. *Neurobiol Aging*. 32:S20-36. doi: 10.1016/j.neurobiolaging.2011.09.006.

Knopman, D. S., Lundt, E. S., Therneau, T. M., Albertson, S. M., Gunter, J. L., Senjem, M. L., … Jack, C. R. (2021). Association of Initial β-Amyloid Levels with Subsequent Flortaucipir Positron Emission Tomography Changes in Persons without Cognitive Impairment. JAMA Neurology, 78(2), 217–228. https://doi.org/10.1001/jamaneurol.2020.3921

Kolinger, G. D., García, D. V., Willemsen, A. T. M., Reesink, F. E., de Jong, B. M., Dierckx, R. A. J. O., … Boellaard, R. (2021). Amyloid burden quantification

depends on PET and MR image processing methodology. PLoS ONE, 16(3 March 2021), e0248122. https://doi.org/10.1371/journal.pone.0248122

Krishnadas, N., Villemagne, V. L., Doré, V., & Rowe, C. C. (2021). Advances in Brain Amyloid Imaging. Seminars in Nuclear Medicine, 51(3), 241–252. https://doi.org/10.1053/j.semnuclmed.2020.12.005

La Joie, R., Ayakta, N., Seeley, W. W., Borys, E., Boxer, A. L., DeCarli, C., … Rabinovici, G. D. (2019). Multisite study of the relationships between antemortem [11 C]PIB-PET Centiloid values and postmortem measures of Alzheimer's disease neuropathology. Alzheimer's and Dementia, 15(2), 205–216. https://doi.org/10.1016/j.jalz.2018.09.001

Laidlaw, D.H., Fleischer, K.W., Barr, A.H. (1998). Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms. *IEEE transactions on medical imaging*. 17(1):74-86.

Landau, S. M., Thomas, B. A., Thurfjell, L., Schmidt, M., Margolin, R., Mintun, M., … Jagust, W. J. (2014). Amyloid PET imaging in Alzheimer's disease: A comparison of three radiotracers. European Journal of Nuclear Medicine and Molecular Imaging, 41(7), 1398–1407. https://doi.org/10.1007/s00259-014-2753-3

Landau, Susan M., Fero, A., Baker, S. L., Koeppe, R., Mintun, M., Chen, K., … Jagust, W. J. (2015). Measurement of longitudinal β-amyloid change with 18F-florbetapir PET and standardized uptake value ratios. Journal of Nuclear Medicine, 56(4), 567–574. https://doi.org/10.2967/jnumed.114.148981

Leuzy, A., Chiotis, K., Hasselbalch, S. G., Rinne, J. O., De Mendonça, A., Otto, M., … Nordberg, A. (2016). Pittsburgh compound B imaging and cerebrospinal fluid amyloid-β in a multicentre European memory clinic study. Brain, 139(9), 2540–2553. https://doi.org/10.1093/brain/aww160

Liu, E., Schmidt, M. E., Margolin, R., Sperling, R., Koeppe, R., Mason, N. S., … Brashear, H. R. (2015). Amyloid-β 11C-PiB-PET imaging results from 2 randomized bapineuzumab phase 3 AD trials. Neurology, 85(8), 692–700. https://doi.org/10.1212/WNL.0000000000001877

Lopes Alves, I., Collij, L. E., Altomare, D., Frisoni, G. B., Saint-Aubert, L., Payoux, P., … Molinuevo, J. L. (2020). Quantitative amyloid PET in Alzheimer's disease: the AMYPAD prognostic and natural history study. Alzheimer's and Dementia, 16(5), 750–758. https://doi.org/10.1002/alz.12069

Lopes Alves, I., Heeman, F., Collij, L. E., Salvadó, G., Tolboom, N., Vilor-Tejedor, N., … Gispert, J. D. (2021). Strategies to reduce sample sizes in Alzheimer's disease primary and secondary prevention trials using longitudinal amyloid PET imaging. Alzheimer's Research and Therapy, 13(1). https://doi.org/10.1186/s13195-021-00819-2

Lynch, S.Y., Irizarry, M.C., Dhadda, S., Li, D.J., Kanekiyo, M., Swanson, C.J. (2021). Baseline characteristics for CLARITY AD: A phase 3 placebo-controlled, double-blind, parallel-group, 18-month study evaluating lecanemab (ban2401) in early Alzheimer's disease. *Alzheimer's Dement*., 17: e054331. https://doi.org/10.1002/alz.054331

Matsuda, H., Ito, K., Ishii, K., Shimosegawa, E., Okazawa, H., Mishina, M., … Sato, N. (2021). Quantitative Evaluation of 18F-Flutemetamol PET in Patients With Cognitive Impairment and Suspected Alzheimer's Disease: A Multicenter Study. Frontiers in Neurology, 11. https://doi.org/10.3389/fneur.2020.578753

Milà-Alomà, M., Salvadó, G., Shekari, M., Grau-Rivera, O., Sala-Vila, A., Sánchez-Benavides, G., … Molinuevo, J. L. (2021). Comparative Analysis of Different Definitions of Amyloid-β Positivity to Detect Early Downstream Pathophysiological Alterations in Preclinical Alzheimer. Journal of Prevention of Alzheimer's Disease, 8(1), 68–77. https://doi.org/10.14283/jpad.2020.51

Milà-Alomà, Marta, Shekari, M., Salvadó, G., Gispert, J. D., Arenaza-Urquijo, E. M., Operto, G., … Molinuevo, J. L. (2021). Cognitively unimpaired individuals with a low burden of Aβ pathology have a distinct CSF biomarker profile. Alzheimer's Research and Therapy, 13(1), 1–12. https://doi.org/10.1186/s13195-021-00863-y

Mintun, M. A., Lo, A. C., Duggan Evans, C., Wessels, A. M., Ardayfio, P. A., Andersen, S. W., … Skovronsky, D. M. (2021). Donanemab in Early Alzheimer's Disease. New England Journal of Medicine, 384(18), 1691–1704. https://doi.org/10.1056/nejmoa2100708

Navitsky, M., Joshi, A. D., Kennedy, I., Klunk, W. E., Rowe, C. C., Wong, D. F., … Devous, M. D. (2018). Standardization of amyloid quantitation with florbetapir standardized uptake value ratios to the Centiloid scale. Alzheimer's and Dementia, 14(12), 1565–1571. https://doi.org/10.1016/j.jalz.2018.06.1353

Nordberg, A., Carter, S.F., Rinne, J. (2013). A European multicentre PET study of fibrillar amyloid in Alzheimer's disease. *Eur J Nucl Med Mol Imaging*. 40, 104–114. https://doi.org/10.1007/s00259-012-2237-2.

Ostrowitzki, S., Lasser, R. A., Dorflinger, E., Scheltens, P., Barkhof, F., Nikolcheva, T., … Fontoura, P. (2017). A phase III randomized trial of gantenerumab in prodromal Alzheimer's disease. Alzheimer's Research and Therapy, 9(1). https://doi.org/10.1186/s13195-017-0318-y

Pemberton HG, Collij LE, Heeman F, Bollack A, Shekari M, Salvadó G, Alves IL, Garcia DV, Battle M, Buckley C, Stephens AW, Bullich S, Garibotto V, Barkhof F, Gispert JD, Farrar G; AMYPAD consortium. (2022). Quantification of amyloid PET for future clinical use: a state-of-the-art review. Eur J Nucl Med Mol Imaging, 49(10):3508-3528. doi: 10.1007/s00259-022-05784-y

Piper, J., A. Nelson, and A. Javorek. Evaluation of a Quantitative Method for Florbetaben (FBB) PET Using SUVR. in EANM. 2014

Rabinovici, G.D., Gatsonis, C., Apgar, C., Chaudhary, K., Gareen, I., … Carrillo, M.C. Association of Amyloid Positron Emission Tomography With Subsequent Change in Clinical Management Among Medicare Beneficiaries With Mild Cognitive Impairment or Dementia (2019). *JAMA*. 2;321(13):1286-1294. doi: 10.1001/jama.2019.2000.

Relkin, N. R., Thomas, R. G., Rissman, R. A., Brewer, J. B., Rafii, M. S., Van Dyck, C. H., … Aisen, P. S. (2017). A phase 3 trial of IV immunoglobulin for Alzheimer disease. Neurology, 88(18), 1768–1775. https://doi.org/10.1212/WNL.0000000000003904

Roberts, C., Kaplow, J., Giroux, M., Krause, S., & Kanekiyo, M. (2021). Amyloid and APOE Status of Screened Subjects in the Elenbecestat MissionAD Phase 3 Program. Journal of Prevention of Alzheimer's Disease, 8(2), 218–223. https://doi.org/10.14283/jpad.2021.4

Rowe, C. C., Doré, V., Jones, G., Baxendale, D., Mulligan, R. S., Bullich, S., … Villemagne, V. L. (2017). 18F-Florbetaben PET beta-amyloid binding expressed in Centiloids. European Journal of Nuclear Medicine and Molecular Imaging, 44(12), 2053–2059. https://doi.org/10.1007/s00259-017-3749-6

Rowe, C. C., Jones, G., Dore, V., Pejoska, S., Margison, L., Mulligan, R. S., … Villemagne, V. L. (2016). Standardized Expression of 18F-NAV4694 and 11C-PiB b-amyloid PET results with the centiloid scale. Journal of Nuclear Medicine, 57(8), 1233–1237. https://doi.org/10.2967/jnumed.115.171595

Rowe, C. C., Pejoska, S., Mulligan, R. S., Jones, G., Chan, J. G., Svensson, S., …

Villemagne, V. L. (2013). Head-to-head comparison of11C-PiB and18F-AZD4694 (NAV4694) for β-amyloid imaging in aging and dementia. Journal of Nuclear Medicine, 54(6), 880–886. https://doi.org/10.2967/jnumed.112.114785

Royse, S. K., Minhas, D. S., Lopresti, B. J., Murphy, A., Ward, T., Koeppe, R. A., … Landau, S. M. (2021). Validation of amyloid PET positivity thresholds in centiloids: a multisite PET study approach. Alzheimer's Research and Therapy, 13(1), 1–10. https://doi.org/10.1186/s13195-021-00836-1

Ruan, S., Fadili, J., Xue, J.H., Bloyet, D. (2000). Brain tissue classification based on a mixel model and Markov random field models. *First International Conference On Image and Graphics*. 369-372.

Sabri, O., Sabbagh, M. N., Seibyl, J., Barthel, H., Akatsu, H., Ouchi, Y., … Schulz-Schaeffer, W. J. (2015). Florbetaben PET imaging to detect amyloid beta plaques in Alzheimer's disease: Phase 3 study. Alzheimer's and Dementia, 11(8), 964–974. https://doi.org/10.1016/j.jalz.2015.02.004

Salloway, S., Farlow, M., McDade, E., Clifford, D. B., Wang, G., Llibre-Guerra, J. J., … van Dyck, C. H. (2021). A trial of gantenerumab or solanezumab in dominantly inherited Alzheimer's disease. Nature Medicine, 27(7), 1187–1196. https://doi.org/10.1038/s41591-021-01369-8

Salloway, S., Gamez, J. E., Singh, U., Sadowsky, C. H., Villena, T., Sabbagh, M. N., … Curtis, C. (2017). Performance of [18F]flutemetamol amyloid imaging against the neuritic plaque component of CERAD and the current (2012) NIA-AA recommendations for the neuropathologic diagnosis of Alzheimer's disease. Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring, 9, 25–34. https://doi.org/10.1016/j.dadm.2017.06.001

Salloway, S., Sperling, R., Fox, N. C., Blennow, K., Klunk, W., Raskind, M., … Brashear, H. R. (2014). Two Phase 3 Trials of Bapineuzumab in Mild-to-Moderate Alzheimer's Disease. New England Journal of Medicine, 370(4), 322–333. https://doi.org/10.1056/nejmoa1304839

Salvadó, G., Milà-Alomà, M., Shekari, M., Minguillon, C., Fauria, K., Niñerola-Baizán, A., … Gispert, J. D. (2021). Cerebral amyloid-β load is associated with neurodegeneration and gliosis: Mediation by p-tau and interactions with risk factors early in the Alzheimer's continuum. Alzheimer's and Dementia, 17(5), 788–800. https://doi.org/10.1002/alz.12245

Salvadó, G., Molinuevo, J. L., Brugulat-Serrat, A., Falcon, C., Grau-Rivera, O., Suárez-Calvet, M., … Gispert, J. D. (2019). Centiloid cut-off values for optimal agreement between PET and CSF core AD biomarkers. Alzheimer's Research and Therapy, 11(1), 1–12. https://doi.org/10.1186/s13195-019-0478-z

Santago, P., Gage, H.D. (1995). Statistical models of partial volume effect. IEEE Trans Image Process. 4(11):1531-40. doi: 10.1109/83.469934.

Santago P, Gage H.D.(1993) Quantification of MR brain images by mixture density and partial volume modelling. IEEE Trans Med Imaging, 12(3):566-74. doi: 10.1109/42.241885.

Schwarz, C. G., Tosakulwong, N., Senjem, M. L., Gunter, J. L., Therneau, T. M., Vemuri, P., … Jack, C. R. (2018). Considerations for performing level-2 centiloid transformations for amyloid PET SUVR values. Scientific Reports, 8(1). https://doi.org/10.1038/s41598-018-25459-9

Sevigny, J., Chiao, P., Bussière, T., Weinreb, P.H., Williams, L., … Sandrock, A. (2016). The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. *Nature*. 1;537(7618):50-6. doi: 10.1038/nature19323.

Sevigny, J., Suhy, J., Chiao, P., Chen, T., Klein, G., … Barakos, J. (2016). Amyloid PET Screening for Enrichment of Early-Stage Alzheimer Disease Clinical Trials: Experience in a Phase 1b Clinical Trial. *Alzheimer Dis Assoc Disord*. 30(1):1-7. doi: 10.1097/WAD.0000000000000144.

Shcherbinin, S., Evans, C.D., Lu, M., Andersen, S.W., Pontecorvo, M.J., … Sims, J.R. (2022). Association of Amyloid Reduction After Donanemab Treatment With Tau Pathology and Clinical Outcomes: The TRAILBLAZER-ALZ Randomized Clinical Trial. *JAMA Neurol*. doi: 10.1001/jamaneurol.2022.2793.

Shea, Y.F., Barker, W., Greig-Gusto, M.T., Loewenstein, D.A., Duara, R., DeKosky, S.T. (2018). Impact of Amyloid PET Imaging in the Memory Clinic: A Systematic Review and Meta-Analysis. *J Alzheimers Dis*. 64(1):323-335. doi: 10.3233/JAD-180239.

Stephenson, D., Hill, D., Cedarbaum, J.M., Tome, M., Vamvakas, S., … Critical Path for Parkinson's Consortium. (2019). The Qualification of an Enrichment Biomarker for Clinical Trials Targeting Early Stages of Parkinson's Disease. *J Parkinsons Dis*. 9(3):553-563. doi: 10.3233/JPD-191648.

Su, Y., Flores, S., Hornbeck, R. C., Speidel, B., Vlassenko, A. G., Gordon, B. A., …

Benzinger, T. L. S. (2018). Utilizing the Centiloid scale in cross-sectional and longitudinal PiB PET studies. NeuroImage: Clinical, 19, 406–416. https://doi.org/10.1016/j.nicl.2018.04.022

Su, Y., Flores, S., Wang, G., Hornbeck, R. C., Speidel, B., Joseph-Mathurin, N., … Benzinger, T. L. S. (2019). Comparison of Pittsburgh compound B and florbetapir in cross-sectional and longitudinal studies. Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring, 11, 180–190. https://doi.org/10.1016/j.dadm.2018.12.008

Swanson, C.J., Zhang, Y., Dhadda, S. (2021). A randomized, double-blind, phase 2b proof-of-concept clinical trial in early Alzheimer's disease with lecanemab, an anti-Aβ protofibril antibody. *Alz Res Therapy* 13, 80. https://doi.org/10.1186/s13195-021-00813-8

Tryputsen, V., Dibernardo, A., Samtani, M., Novak, G. P., Narayan, V. A., & Raghavan, N. (2015). Optimizing regions-of-interest composites for capturing treatment effects on brain amyloid in clinical trials. Journal of Alzheimer's Disease, 43(3), 809–821. https://doi.org/10.3233/JAD-131979

Tudorascu, D. L., Minhas, D. S., Lao, P. J., Betthauser, T. J., Yu, Z., Laymon, C. M., … Cohen, A. D. (2018). The use of Centiloids for applying [11C]PiB classification cutoffs across region-of-interest delineation methods. Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring, 10, 332–339. https://doi.org/10.1016/j.dadm.2018.03.006

van der Kall, L. M., Truong, T., Burnham, S. C., Doré, V., Mulligan, R. S., Bozinovski, S., … Rowe, C. C. (2021). Association of β-Amyloid Level, Clinical Progression, and Longitudinal Cognitive Change in Normal Older Individuals. Neurology, 96(5), e662–e670. https://doi.org/10.1212/WNL.0000000000011222

Villemagne, V. L., Barkhof, F., Garibotto, V., Landau, S. M., Nordberg, A., & van Berckel, B. N. M. (2021). Molecular imaging approaches in dementia. Radiology, 298(3), 517–530. https://doi.org/10.1148/radiol.2020200028

Meyer PF, Pichet Binette A, Gonneaud J, Breitner JCS, Villeneuve S.(2020). Characterization of Alzheimer Disease Biomarker Discrepancies Using Cerebrospinal Fluid Phosphorylated Tau and AV1451 Positron Emission Tomography. JAMA Neurol., 77(4):508-516. doi: 10.1001/jamaneurol.2019.4749

Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., …

Trojanowski, J. Q. (2017). The Alzheimer's Disease Neuroimaging Initiative 3: Continued innovation for clinical trial improvement. Alzheimer's and Dementia, 13(5), 561–571. https://doi.org/10.1016/j.jalz.2016.10.006

Yun, H. J., Moon, S. H., Kim, H. J., Lockhart, S. N., Choe, Y. S., Lee, K. H., … Seo, S. W. (2017). Centiloid method evaluation for amyloid PET of subcortical vascular dementia. Scientific Reports, 7(1). https://doi.org/10.1038/s41598-017-16236-1