NeuroImage 297 (2024) 120748

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Towards multicenter β-amyloid PET imaging in mouse models: A triple scanner head-to-head comparison

Johannes Gnörich [a,b,*], Mara Koehler [a], Karin Wind-Mark [a], Carolin Klaus [b], Artem Zatcepin [a,b], Giovanna Palumbo [a], Manvir Lalia [a], Laura Sebastian Monasor [b,c], Leonie Beyer [a], Florian Eckenweber [a], Maximilian Scheifele [a], Franz-Josef Gildehaus [a], Barbara von Ungern-Sternberg [a], Henryk Barthel [d], Osama Sabri [d], Peter Bartenstein [a,e], Jochen Herms [b,e,f], Sabina Tahirovic [b], Nicolai Franzmeier [e,g], Sibylle Ziegler [a], Matthias Brendel [a,b,f,*]

[a] Department of Nuclear Medicine, University Hospital, LMU Munich, Munich, Germany
[b] German Center for Neurodegenerative Diseases (DZNE) Munich, Munich, Germany
[c] Graduate School of Systemic Neuroscience, LMU Munich, Munich, Germany
[d] Department of Nuclear Medicine, University of Leipzig, Leipzig, Germany
[e] Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
[f] Center of Neuropathology and Prion Research, University of Munich, Munich Germany
[g] Institute for Stroke and Dementia Research, LMU Munich, Munich, Germany

ARTICLE INFO

Keywords:
None

ABSTRACT

Aim: β-amyloid (Aβ) small animal PET facilitates quantification of fibrillar amyloidosis in Alzheimer's disease (AD) mouse models. Thus, the methodology is receiving growing interest as a monitoring tool in preclinical drug trials. In this regard, harmonization of data from different scanners at multiple sites would allow the establishment large collaborative cohorts and may facilitate efficacy comparison of different treatments. Therefore, we objected to determine the level of agreement of Aβ-PET quantification by a head-to-head comparison of three different state-of-the-art small animal PET scanners, which could help pave the way for future multicenter studies.

Methods: Within a timeframe of 5 ± 2 weeks, transgenic APPPS1 ($n = 9$) and wild-type (WT) ($n = 8$) mice (age range: 13–16 months) were examined three times by Aβ-PET ([^{18}F]florbetaben) using a Siemens Inveon DPET, a MedisonanoScan PET/MR, and a MedisonanoScan PET/CT with harmonized reconstruction protocols. Cortex-to-white-matter 30–60 min p.i. standardized uptake value ratios ($SUVR_{CTX/WM}$) were calculated to compare binding differences, effect sizes (Cohen's d) and z-score values of APPPS1 relative to WT mice. Correlation coefficients (Pearson's r) were calculated for the agreement of individual SUVR between different scanners. Voxel-wise analysis was used to determine the agreement of spatial pathology patterns. For validation of PET imaging against the histological gold standard, individual SUVR values were subject to a correlation analysis with area occupancy of methoxy-X04 staining.

Results: All three small animal PET scanners yielded comparable group differences between APPPS1 and WT mice (Δ_{PET}=20.4 % ± 2.9 %, $\Delta_{PET/MR}$=18.4 % ± 4.5 %, $\Delta_{PET/CT}$=18.1 % ± 3.3 %). Voxel-wise analysis confirmed a high degree of congruency of the spatial pattern (Dice coefficient $(DC)_{PETvs.PET/MR}$=83.0 %, $DC_{PETvs.PET/CT}$=69.3 %, $DC_{PET/MRvs.PET/CT}$=81.9 %). Differences in the group level variance of the three scanners resulted in divergent z-scores (z_{PET}=11.5 ± 1.6; $z_{PET/MR}$=5.3 ± 1.3; $z_{PET/CT}$=3.4 ± 0.6) and effect sizes (d_{PET}=8.5, $d_{PET/MR}$=4.5, $d_{PET/CT}$=4.1). However, correlations at the individual mouse level were still strong between scanners ($r_{PETvs.PET/MR}$=0.96, $r_{PETvs.PET/CT}$=0.91, $r_{PET/MRvs.PET/CT}$=0.87; all $p \leq 0.0001$). Methoxy-X04 staining exhibited a significant

Abbreviations: AD, Alzheimer's disease; ANOVA, Analysis of Variance; Aβ, β-amyloid; CT, Computerized Tomographie; DC, Dice coefficients; FDA, Food and Drug Administration; ([^{18}F]FDG), 2-Deoxy-2-[^{18}F]fluoro-D-glucose; FWHM, Full Width at Half Maximum; MR, Magnetic resonance; OSEM-3D, ordered subset expectation maximization; PET, Positron emission tomography; SUV, Standardized uptake value; WT, Wildtype.

* Corresponding authors at: Marchinonistrasse 15, 81377 Munich, Germany.
E-mail addresses: Johannes.Gnoerich@med.uni-muenchen.de (J. Gnörich), Matthias.Brendel@med.uni-muenchen.de (M. Brendel).

https://doi.org/10.1016/j.neuroimage.2024.120748
Received 29 March 2024; Received in revised form 17 July 2024; Accepted 18 July 2024
Available online 26 July 2024

correlation across all three PET machines combined ($r = 0.76$, $p < 0.0001$) but also at individual level (PET: $r = 0.81$, $p = 0.026$; PET/MR: $r = 0.89$, $p = 0.0074$; PET/CT: $r = 0.93$, $p = 0.0028$).
Conclusions: Our comparison of standardized small animal Aβ-PET acquired by three different scanners substantiates the possibility of moving towards a multicentric approach in preclinical AD research. The alignment of image acquisition and analysis methods achieved good overall comparability between data sets. Nevertheless, differences in variance of sensitivity and specificity of different scanners may limit data interpretation at the individual mouse level and deserves methodological optimization.

1. Introduction

Alzheimer's disease (AD) accounts for a significant portion of neurodegenerative diseases and is a primary cause of dementia in the elderly. Neurodegenerative diseases, including AD, are currently incurable, and efforts are being made worldwide to explore therapeutic possibilities. Recently, on June 7, 2021, and January 6, 2023, the Food and Drug Administration (FDA) in the USA approved the first potential causal therapies, Aducanumab (Sevigny et al., 2016) and Lecanemab (van Dyck et al., 2023), respectively. These therapies are designed to eliminate Aβ protein deposits, which are, besides neurofibrillary tau-tangles, a fundamental component in the pathogenesis of Alzheimer's disease (Hardy and Higgins, 1992; Braak and Braak, 1991). Positron emission tomography (PET) for detecting β-amyloid (Aβ) has gained widespread acceptance as a valuable tool for both, identifying and quantifying amyloidosis in the brains of individuals suspected to have AD (Nelissen et al., 2009). Moreover, positivity in Aβ-PET is an important inclusion criterion in clinical trials of anti-amyloid therapy (Sabri et al., 2015) and functions as a biological endpoint in terms of drug efficacy within various trials (Sevigny et al., 2016). In basic science, Aβ-PET has proven to be a promising instrument for monitoring neuropathological changes in transgenic Aβ mouse models (Xia et al., 2022; Zhao et al., 2022; A Rominger et al., 2013). Within this context, molecular imaging with PET has been established in a diverse array of mouse models that either overexpress familial AD-related genes, mutant amyloid precursor protein (APP), or APP and Presenilin (PS) (first generation). Second generation AD mouse models employ knock-in modifications producing pathogenic Aβ such as Aβ42 without an overexpression APP (Watamura et al., 2022; Sasaguri et al., 2017). In the realm of Aβ-PET, both in larger interventional and clinical multicenter studies, it is a common scenario for Aβ-PET-imaging to be carried out using different scanners (Nordberg et al., 2013). However, this practice introduces additional disruptive factors, such as inter-reader variability and acquisition heterogeneity, which require adequate standardization (Klunk et al., 2015; Smith et al., 2023). Therefore, harmonizing imaging and reconstruction procedures becomes imperative to mitigate issues arising from the diverse acquisition settings. For instance, previously published preclinical imaging studies utilizing 2-Deoxy-2-[^{18}F]fluoro-D-glucose ([^{18}F]FDG) have identified various factors influencing uptake patterns, such as animal handling (Fueger et al., 2006), dietary state (Wong et al., 2011) and anesthesia (Tremoleda et al., 2012). Thus, a notable advantage of Aβ-PET imaging is its relative independence from metabolic alterations in the mice. However, several challenges considering Aβ-PET imaging have been reported (Syvänen et al., 2022). Among the frequently used agents, such as [^{18}F]florbetaben, high white matter retention and nonspecific binding in adjacent bone are common issues (Richards and Sabbagh, 2014; Wolk et al., 2012). Additionally, the presence of asymmetric plaque burden (Sacher et al., 2020) may account for the inter-animal variation observed in many studies (Syvänen et al., 2022).

Considering preclinical imaging studies however, only scarce evidence of studies conducted in a multicenter approach is available to date (Mannheim et al., 2019; Dirnagl and International, 2012). A comprehensive analysis to assess the comparability and feasibility of small animal Aβ-PET recordings when performed using different scanners in mice with amyloid pathology has not yet been conducted.

Hence, our primary objective was to gauge the level of agreement in Aβ-PET quantification by directly comparing the performance of three distinct state-of-the-art PET scanners. This endeavor is aimed at facilitating future multicenter investigations in this field.

2. Methods

2.1. Animals

Experiments were carried out with APPPS1–21 (APPPS1) mice, generated on a C57BL/6 J genetic background, which exhibit substantial fibrillar Aβ plaque pathology. This transgenic mouse model coexpresses mutated APP and presenilin 1 (PS) under a neuron-specific Thy1 promoter, leading to the onset of cerebral amyloidosis at the age of 6–8 weeks (Radde et al., 2006). Age- and sex- matched C57BL/6 J mice served as controls (WT).

2.2. Experimental setup and study design

Within a timeframe of 5 ± 2 weeks, all transgenic APPPS1 ($n = 9$, 5 female, 4 male) and wildtype (WT, $n = 8$, 4 female, 4 male) mice (age range: 13–16 months) were each consecutively examined in total of three times by Aβ-PET ([^{18}F]florbetaben) using a MedisonanoScan PET/MR (1st), a Siemens Inveon DPET (2nd), and a MedisonanoScan PET/CT (3rd) with harmonized reconstruction protocols (see below). All three scanners were available at the same site at LMU Munich. Randomization was not used to allocate experimental units because there was no intervention involved. Blinding was not performed during scanning but during image analysis, where an automatic co-registration step (Overhoff et al., 2016) ensured reader independence. After the final scan, brains were collected for subsequent analysis of brain pathology. Sample size calculation with previous data (Biechele et al., 2021) and α=0.05 revealed a required cohort of $n = 7$ APPPS1 and $n = 7$ WT to detect group differences at a power of 0.8 (G*Power 3.1.9.7, HHU Düsseldorf).

To account for animal and imaging based drop-outs $n = 10$ mice per genotype were included in the study. One APPPS1 mouse and two WT mice were excluded from the analysis a priori due to incomplete imaging data. Additionally, we conducted a sample size calculation to estimate the number of animals required for each specific scanner, as compared to a multicentric study in a sensitive and robust paradigm. This calculation was based on the variances determined in this study. We assumed a meaningful effect of 5 %, with an α level of 0.05 and a power of 0.95, aiming to detect sensitive differences for potential interventional trials in mouse models of Alzheimer's disease (AD). All animal experiments were approved by the local animal care committee of the Government of Oberbayern (Regierung Oberbayern) (AZ 55.2–1–54–2532–210–2015 and AZ 55.2–1–54–2531–162–13, 24,015, 106–17), overseen by a veterinarian and in compliance with the ARRIVE guidelines 2.0 (Supplement) and were carried out in accordance with the U.K. Animals (Scientific Procedures) Act, 1986 and associated guidelines, EU Directive 2010/63/EU for animal experiments. Animals were housed in a temperature- and humidity-controlled environment with a 12 h light–dark cycle, with free access to food (Ssniff Spezialdiäten GmbH) and water. Anesthesia prior to tracer application was induced with 3.0 % isoflurane and maintained during PET scanning with isoflurane 1.5 %

delivered via a mask at 3.5 L/min. Following the experiment, the mice were transferred to fresh temporary cages equipped with food, water, and heating mats for warming. The animals were returned to their home cages only after they had fully regained consciousness (Palumbo et al., 2023). [^{18}F]florbetaben (Neuraceq) precursor was provided by LIFE Molecular Imaging GmbH and the synthesis was performed as reported previously (A Rominger et al., 2013), yielding a specific activity of ~80 GBq/μmol and a radiochemical purity of 98 %.

2.3. Image acquisition and phantom scan

PET images were acquired by using a dedicated PET-scanner (Siemens Inveon DPET), a PET/MR (3T MedisonanoScan PET/MR scanner, MedisoLtd., Hungary) and a nanoscan PET/CT (MedisoLtd., Budapest, Hungary) with harmonized acquisition and reconstruction protocols to ensure comparable spatial resolution. For this purpose, spatial resolution among the scanner was determined using a high-resolution phantom in mouse size (Quality Assurance in Radiation and Medicine (QRM), Germany). The Micro Derenzo phantom used for this purpose has a total length of the active part of 32 mm, with the middle 12 mm long insert containing fillable hot-rod resolution sections with rod diameters of 0.6, 0.8, 1.0, 1.2, 1.5, and 2.0 mm (Omidvari et al., 2017). The phantom was filled with an [^{18}F]FDG solution, resulting in an activity of approximately 20 MBq at the start of the 1-hour acquisition. Subsequently, the phantom scan was reconstructed according to the default reconstruction protocols of each individual scanner. PMOD (V3.5, PMOD technologies, Basel, Switzerland) was used to delineate the different segments of the phantom and a straight line was placed in transaxial images of the phantom through the segments of 2.0 mm and 1.2 mm. Line profiles of respective sections were created and compared between all three scanner modalities. Higher peak-to-valley ratios, as a surrogate of resolution, were observed in both segments of the PET/MR and PET/CT compared to the PET-only device (Supplementary Figure 1A, C). Based on these findings for PET images derived from the PET/MR and PET/CT, a Gaussian filter (1.0 mm^3) was applied for harmonization of the full-width-at-half-maximum to achieve similar spatial resolution (Supplementary Figure 1), as previously performed in human multicenter studies (i.e. (Brendel et al., 2020)). In brief, animal scans consisted of a static single emission frame made in the interval 30–60 min (A Rominger et al., 2013) after bolus injection of approximately 15.2 ± 2.8 MBq (PET=13.5 ± 2.9 MBq; PET/MR=17.0 ± 2.6 MBq; PET/CT=14.9 ± 1.8 MBq) [^{18}F]florbetaben (in 150 μl saline) to a tail vein.

2.3.1. Siemens inveon dpet

Mice were placed in the aperture of the scanner a “four-bed mouse hotel”, followed by a 15 min transmission scan made using a rotating [^{57}Co] point source. The image reconstruction procedure was performed with a three-dimensional ordered subset expectation maximization (OSEM-3D) with 4 iterations and 12 subsets followed by a maximum-a-posteriori (MAP) algorithm with 32 iterations. All data were random, scatter-, attenuation-, and decay-corrected and processed with a zoom factor of 1.0 leading to a final voxel dimension of 0.78×0.78×0.80 mm^3.

2.3.2. 3T Medisonanoscan PET/MR

When conducting scans using the 3T MedisonanoScan PET/MR system, in accordance with previously published studies (Gnorich et al., 2023), we employed a single-mouse imaging chamber. A 15-minute anatomical T1 MR scan was performed before [^{18}F]florbetaben injection (head receive coil, matrix size 96×96×22, voxel size 0.24×0.24×0.80 mm^3, repetition time 677 ms, echo time 28.56 ms, flip angle 90°). PET list-mode data were reconstructed using a 3D iterative algorithm (Tera-Tomo 3D, MedisoLtd, Hungary) with the following parameters: matrix size 55×62×187 mm^3, voxel size 0.3 × 0.3 × 0.3 mm^3, 8 iterations, 6 subsets. Random, scatter, attenuation, and decay corrections were applied. The T1 image was used to create a body-air material map for scatter and attenuation correction.

2.3.3. nanoScan PET/CT

Similar to the Siemens Inveon DPET, mice were analyzed in a 4-mouse imaging chamber.The PET reconstruction procedure of scans performed by small animal MedisoPET/CT system also included an 3D iterative algorithm (Tera-Tomo 3D, MedisoLtd, Hungary) with 4 iterations, 6 subsets, decay correction, scatter correction, attenuation correction, dead time correction, and sensitivity normalization. An x-ray computed tomography scan was used for generating a material map for scatter and attenuation correction. The resulting PET images had 212 × 212 × 235 voxels of 0.4 × 0.4 × 0.4 mm^3. The resulting CT images had 972 × 972 × 930 voxels of 0.12 × 0.12 × 0.12 mm^3 (Bartos et al., 2022).

2.4. Image preprocessing and analysis

Aβ-PET preprocessing for all image data was performed by PMOD (V3.5, PMOD technologies, Basel, Switzerland). All analyses were standardized to ensure comparability of procedures (Overhoff et al., 2016). Unified spatial normalization (nonlinear warping, transient 0.6 mm^3 Gaussian smoothing of the input image, 16 iterations, frequency cutoff 3, no thresholding) of all original Aβ-PET images to the same previously established Aβ-PET template was performed, to ensure spatial alignment between all different studies. Cortex-to-white-matter standardized uptake value ratios ($SUVR_{(CTX/WM)}$) were calculated utilizing our a bilateral neocortical volume of interest (15 mm^3) (Biechele et al., 2021) in relation to our previously validated white matter reference region (Overhoff et al., 2016) to compare percentage difference, effect sizes (Cohen’s d) and z-score values relative to WT mice. For visualization purposes, PET scans from all three devices are depicted on a standard mouse T1w-MRI template in Ma-Benveniste-Mirrione space (Ma et al., 2005).

2.5. Methoxy-X04 staining

Brains used to carry out histological analysis were separated at the axial plane and one hemisphere was fixed in 4 % PFA at 4 °C overnight. For each animal ($n = 7$) three 50 μm thick slices, containing the frontal cortex were cut at a vibratome (VT1200S, Leica Biosystems, Germany). Free-floating sections were permeabilized with 0.03 % Trition X-100 in PBS at 4 °C overnight and subsequent blocked with 10 % NGS in PBS for 1 h at RT. To stain fibrillar Aβ plaques, methoxy-X04 (0.01 mg/ml 50 % Ethanol) was applied onto the slices for 15 min at RT. Excessive dye was removed by three times washing with PBS, before the slices were mounted on microscope slides with fluorescent mounting medium (Dako, Santa Clara, USA).

In the motor cortex, three-dimensional images were acquired with a LSM 900 confocal microscope (Zeiss, Oberkochen, Germany) using a 10x immersion objective. The excitation wavelength for methoxy-X04 was 405 nm and the emission was detected from 403 to 585 nm. Image were quantified with ImageJ/FIJI (Version v1.53j). To analyze fibrillar Aβ plaques, pictures were contrast/brightness adjusted and a fixed threshold was set to quantify the percentage area (area-%) of the methoxy-X04 stain. For quantification, the area-% of three images per animal was averaged.

2.6. Statistical analysis

Statistical tests were performed using SPSS (V27.0; IBM Corp., Chicago, IL, USA) and GraphPad Prism 9. $SUVR_{(CTX/WM)}$ values were compared between APPPS1 mice and WT mice with regards to mean values at the group level,%-differences, effect sizes (Cohen’s d) and individual z-scores for each PET system. Furthermore, a correlation analysis was performed between $SUVR_{(CTX/WM)}$ values derived from different PET systems in the whole cohort and in the APPPS1 cohort. We performed whole-brain voxel-wise comparisons of white matter scaled

SUVR images by statistical parametric mapping 12 (SPM) in all three scanner modalities (PET, PET/CT, PET/MR). SPM analyses were corrected for multiple comparisons by a false discovery rate (FDR) correction (Benjamini and Hochberg, 1995). An FDR-corrected p-value of 0.05 was set as the threshold for voxelwise analyses. We calculated two sample *t*-tests between each of the three different imaging modalities to obtain T-score maps. After binarization of T-score maps based on a significance threshold, dice coefficients (DC) were computed. For validation of PET imaging against the histological gold standard, individual white matter scaled SUVR values were subject to a correlation analysis with area occupancy of methoxy-X04 staining.

3. Results

3.1. Inter-scanner agreement of Aβ $SUVR_{CTX/WM}$ in APPPS1 and WT mice

A comparable mean $SUVR_{(CTX/WM)}$ was observed within all three scanners for both genotypes: APPPS1 (PET: 0.94; PET/MR: 0.94; PET/CT: 0.93) and WT (PET: 0.78; PET/MR: 0.80; PET/CT: 0.79) (Supplementary Fig. 2), yielding significant group differences between both mouse models across all three modalities (all $p < 0.0001$). No significant differences were observed in the intragroup comparisons (Supplementary Table 1). Fig. 1 shows the association between the three different scanner modalities of Aβ-PET $SUV_{(CTX/WM)}$ in predefined cortical VOIs for APPPS and WT mice. Including both abovementioned cohorts in the correlation analysis revealed a significant positive association between all three modalities: PET vs. PET/MR ($r = 0.96$, $p < 0.0001$), PET vs.

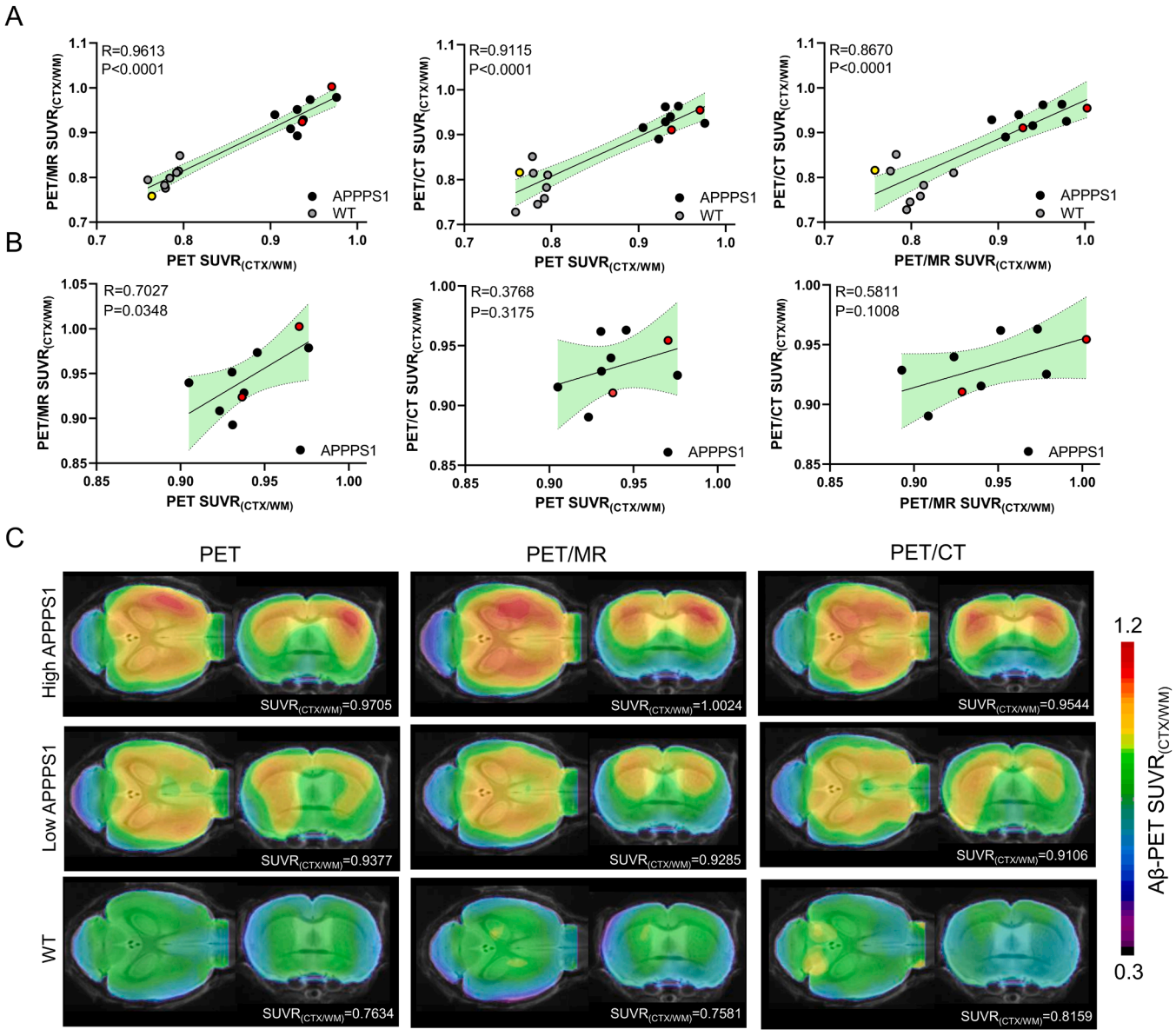

Fig. 1. (A) Correlation analysis between the [^{18}F]florbetaben $SUVR_{(CTX/WM)}$ values determined by the three different scanners in APPPS1 ($n = 9$) and WT mice ($n = 8$) aged 13–16 months. Scatter plots show strong correlations between all the modalities. (B) Dedicated APPPS1 analysis shows significant correlation between PET vs. PET/MR ($p < 0.05$), whereas PET vs. PET/CT and PET/MR vs. PET/CT did not meet the 5 % significance level. R indicates Pearson's coefficient of correlation. Red dots highlight two distinct APPPS1 mice, and yellow dots highlight one respective WT mouse visualized in panel C. (C) Representative [^{18}F]florbetaben PET SUVR images upon an MRI template of two transgenic APPPS1 mice compared to one WT mouse show similar $SUVR_{(CTX/WM)}$ values and visual image impressions across the three PET systems.

PET/CT ($r = 0.91$, $p < 0.0001$) and PET/MR vs. PET/CT ($r = 0.87$, $p < 0.0001$) (Fig. 1A). Sole consideration of transgenic APPPS1 mice, however, showed only significant association in the PET vs. PET/MR comparison ($r = 0.70$, $p = 0.035$), whereas PET vs. PET/CT ($r = 0.38$, $p = 0.32$) and PET/MR vs. PET/CT ($r = 0.58$, $p = 0.10$) correlations did not meet the significance threshold (Fig. 1B). Thus, we questioned the reason for lower $SUVR_{(CTX/WM)}$ correlations in the APPPS1 cohort. Differences in the group-level variance among the three scanners led to varying z-scores (z_{PET}=11.5 ± 1.6; $z_{PET/MR}$=5.3 ± 1.3; $z_{PET/CT}$=3.4 ± 0.6) (Fig. 2B) and effect sizes (d_{PET}=8.5; $d_{PET/MR}$=4.6; $d_{PET/CT}$=4.1), whereas%-differences between APPPS1 and WT mice were stable (Δ_{PET}=20.4 % ± 2.9 %, $\Delta_{PET/MR}$=18.4 % ± 4.5 %, $\Delta_{PET/CT}$=18.1 % ± 3.3 %) across the three small animal PET systems (Fig. 2A). Hence, the smaller range of $SUVR_{(CTX/WM)}$ values in the cohort of transgenic mice suffered from methodological variance in individual mice, whereas quantitative group differences were obtained at a similar level across PET machines. This variance was also observed in the sole correlation of WT mice where the significance level was only reached between PET and PET/MR ($r = 0.74$, $p = 0.037$) compared to PET vs. PET/CT ($r = 0.10$, $p = 0.82$) and PET/MR vs. PET/CT ($r = -0.23$, $p = 0.59$), also pointing on methodological but not animal specific variance in Aβ-PET SUVR of WT mice without target in brain. To determine the appropriate group sizes for a study using each of the employed scanners, we performed a sample size calculation based on the observed group variances specific to each scanner. Our sample size calculation determined that 10 animals (5 per group, $d = 2.7$) are required for the PET-only device, 24 animals (12 per group, $d = 1.6$) for the PET/MR, and 28 animals (14 per group, $d = 1.5$) for the PET/CT. This calculation assumes a meaningful effect size of 5 %, with a power of 0.95 and an alpha level of 0.05. On average, this equates to 20.7 animals across the three scanners. Conversely, a multicentric study that incorporates the variances from all aforementioned scanners, would require a total of 20 animals (10 per group, $d = 1.8$).

3.2. Comparable patterns of a voxel-wise APPPS1 vs. WT mice analysis across [^{18}F]florbetaben imaging with different PET imaging systems

We asked if regional information of small animal Aβ-PET is similar across PET systems and assessed spatial PET signal patterns using voxel-wise analysis via SPM. All three small animal PET scanners yielded comparable group differences between APPPS1 and WT mice as shown in Fig. 2A. Voxel-wise analysis further confirmed a high degree of congruency in the spatial patterns ($DC_{PETvs.PET/MR}$=83.0 %; $DC_{PETvs.PET/CT}$=69.3 %; $DC_{PET/MRvs.PET/CT}$=81.9 %) (Fig. 3).

3.3. Positive association between Aβ imaging with different PET systems and staining for fibrillar Aβ plaques

Assessed area coverage of fibrillar Aβ (methoxy-X04 positive; Fig. 4A) in the motor cortex of $n = 7$ APPPS1 mice ranged from 2.0 % to 5.2 % with an average of 3.6 %±1.0 % plaque load. A notable agreement with methoxy-X04 staining was revealed across all three PET machines combined ($r = 0.76$, $p < 0.0001$) (Fig. 4B). Correlation with SUVR values from each individual PET scanner revealed a comparable trend (PET: $r = 0.81$, $p = 0.026$; PET/MR: $r = 0.89$, $p = 0.0074$; PET/CT: $r = 0.93$, $p = 0.0028$).

Discussion

Here, we addressed the major question of multi-scanner acquisition with the idea of acquiring small animal Aβ-PET scans with different PET scanners by a head-to-head comparison of three different modalities (Siemens Inveon DPET, MedisonanoScan PET/MR, and MedisonanoScan PET/CT). We compared [^{18}F]florbetaben $SUVR_{CTX/WM}$ values among each device to determine the degree of congruency and to translate these findings to multicenter preclinical studies.

Our findings provide valuable insights into the comparability of different scanner modalities in assessing Aβ-PET tracer uptake in Aβ-rich cortical regions. In our first analysis, which included both APPPS1 and WT mice, we observed a strong and significant positive association

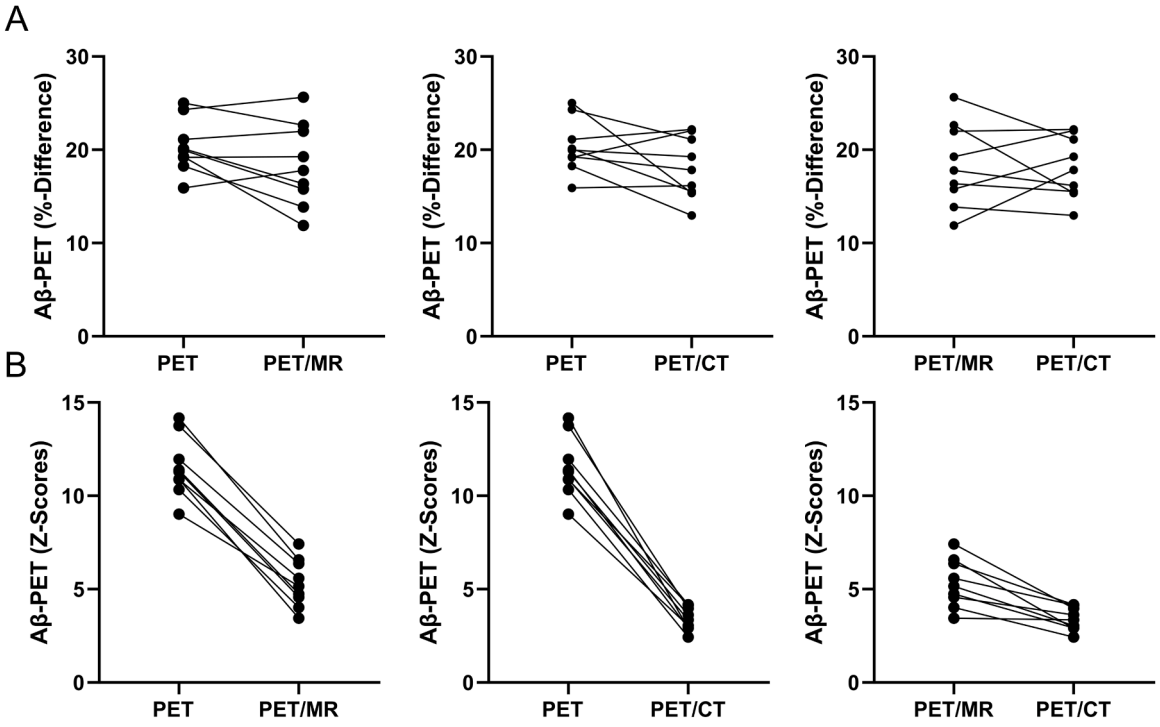

Fig. 2. Aβ percentage difference values (A) of individual APPPS1 mice ($n = 9$) SUVRs to WT mean coincide well between the different modalities in comparison to differing z-score values (B) due to the higher variance in PET/MR and PET/CT.

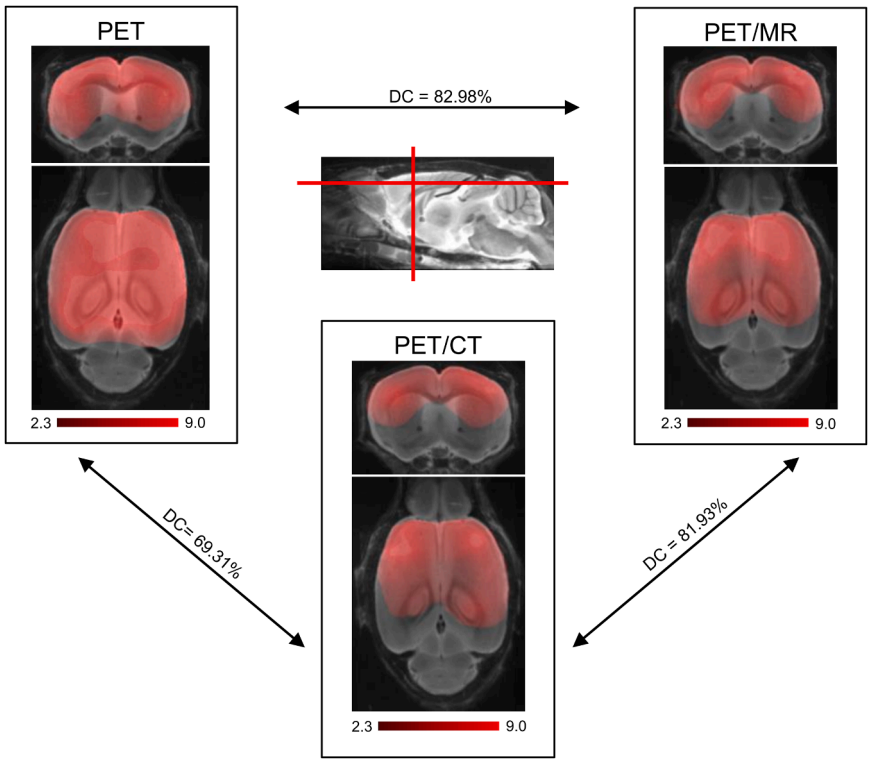

Fig. 3. Whole-brain voxel-wise group comparisons between APPPS1 mice ($n = 9$) versus WT ($n = 8$) based on white matter scaled [^{18}F]florbetaben images of three different PET scanners (PET, PET/CT, PET/MR). T-score maps show results of a two independent sample *t*-test, $p < 0.05$, FDR-corrected, $k > 20$ voxel, projected on MR imaging mouse atlas (gray scale). Dice coefficient (DC) calculations between the binarized T-score maps show very high spatial agreement in PET vs. PET/MR and PET/MR vs. PET/CT comparisons and high agreement in PET vs. PET/CT comparison.

between all three modalities. These results underscore the robustness and consistency of Aβ $SUVR_{CTX/WM}$ measurements across different scanner technologies when the method is used to distinguish Aβ-positive from Aβ-negative tissue, also highlighting the potential utility of multisite imaging in preclinical research. However, when we restricted our analysis to the transgenic APPPS1 cohort, we found that only the PET vs. PET/MR comparison showed a significant positive association due to smaller group variances. In contrast, PET vs. PET/CT and PET/MR vs. PET/CT did not reach the 5 % significance threshold. These findings suggest that while PET and PET/MR remain consistent in measuring [^{18}F]florbetaben $SUVR_{CTX/WM}$ in transgenic mice, there may be some variations when using PET/CT, possibly due to differences in precision and consequently sensitivity and specificity warranting further investigation and standardization.

Notably, the mean $SUVR_{CTX/WM}$ uptake values were comparable across all three scanners for both WT and APPPS1 groups, further supporting the agreement observed in our correlation analysis. This consistency in mean values adds strength to the argument for the robustness of these imaging modalities.

With regard to the choice of the quantification method, $SUVR_{CTX/WM}$ ratio demonstrated superior performance in Aβ detection through PET imaging in transgenic mice compared to other reference regions, making it a valuable choice in preclinical imaging studies (Overhoff et al., 2016). Especially in studies of longitudinal design or preclinical interventional trials, where precise monitoring of individual animals is essential, white matter intensity scaling along with automated spatial normalization can improve the quantification and comparability of [^{18}F] florbetaben uptake (Overhoff et al., 2016).

Next, we investigated voxel-wise [^{18}F]florbetaben $SUVR_{CTX/WM}$, comparing APPPS1 to WT mice across three different scanners. Notably, all three scanners consistently revealed comparable group differences between APPPS1 and WT mice, all yielding highly significant p-values. These findings indicate the robustness and reliability of the imaging modalities in detecting differences in the PET readout at the group level. Furthermore, the voxel-wise analysis emphasized the high degree of congruency in the spatial patterns of the group differences between the scanners. The Dice coefficient calculations demonstrated substantial agreement between PET vs. PET/MR and PET/MR vs. PET/CT, while the agreement was slightly lower in the PET vs. PET/CT comparison. This observation underscores the consistency in the distribution of the PET signal group differences across different modalities, enhancing confidence in their reliability for research purposes.

However, it is important to note that despite the consistency in group differences and spatial patterns, there were differences in group-level variance across the three scanners. This variation resulted in different z-scores and effect sizes, with Inveon DPET achieving the highest z-score and MediosPET/CT achieving the lowest z-scores. The application of a 3D Gaussian filter to approximate spatial resolution between the two more modern devices (PET/MR and PET/CT) compared to the PET-only device could have contributed to the observed variations. Furthermore, variations in z-scores and effect sizes among the scanners suggest differences in sensitivity and specificity, which should be considered in study cohort size estimation and/or in the interpretation of results when using different imaging technologies. In contrast, the%-differences between APPPS1 and WT mice remained relatively stable across the scanners, which again highlights the comparability of small animal Aβ-PET quantification at the group level. Additionally, we conducted a sample size calculation to determine the number of animals required for each specific scanner, compared to a multicentric study. This calculation was based on the variances observed in this study with specified statistical parameters. Our findings indicated that the different PET scanners required varying group sizes, reflecting the need to account for scanner-specific variances. On average, the total number of animals required across the three scanners was slightly higher than that for a multicentric study, which is driven by the highest variance among the scanners (PET/CT > PET/MR > PET). Importantly, the average SUVr

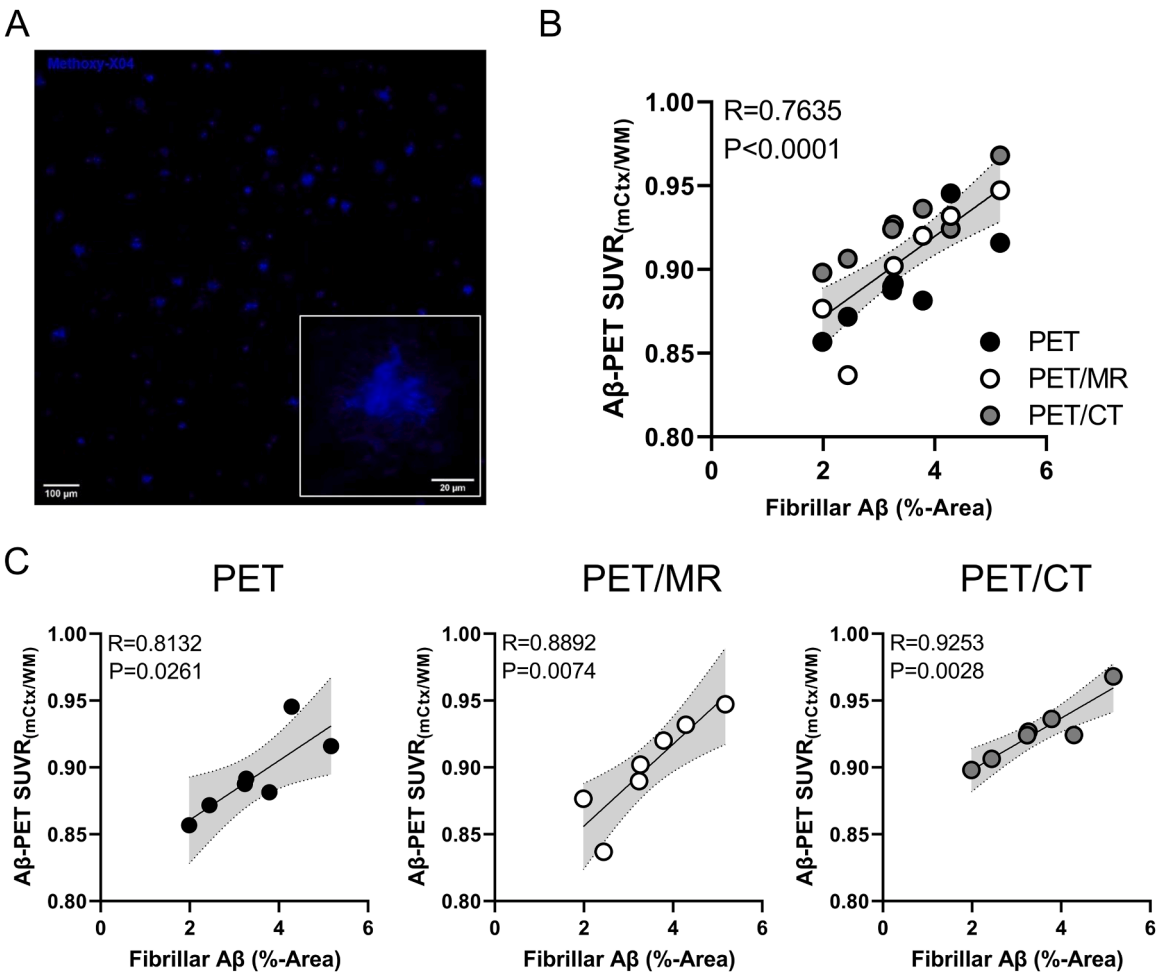

Fig. 4. Association of methoxy-X04 positive Aβ plaque load in the motor cortex with quantitative [^{18}F]florbetaben PET $SUVR_{(mCTX/WM)}$ acquired in the motor cortex by different PET systems. (A) Representative image of methoxy-X04 staining. (B,C) Correlation analysis was conducted on APPPS1 mice ($n = 7$) using data from all three PET scanners combined, as well as from each PET scanner separately.

values across all WT scans were not significantly different, resulting in comparable numbers of animals required for both multicentric and single-site settings.

In multiple analyses involving amyloidosis mouse models, a robust concordance was observed between Aβ-PET imaging and various histological or immunohistochemical markers for Aβ (Xia et al., 2022; A Rominger et al., 2013). In keeping with these literature data, we found an overall positive association between both methods. Establishing this association is crucial for using in vivo PET imaging as a surrogate for the ex vivo diagnostic gold standard, also in small animals.

There are several limitations in this study. Besides the relatively limited sample size of both cohorts (APPPS1: $n = 9$; WT: $n = 8$), immunohistochemistry was only performed in $n = 7$ APPPS1 mice. Also, differences in PET scanner sensitivity, spatial resolution and distinct attenuation corrections (PET/MR and PET/CT, and Co-57 retractable point source for DPET) (Constantinescu and Mukherjee, 2009) could have influenced our findings. In brief, the PET-only scanner utilizes an external radioactive source of known activity. Correction values are measured with an initial blank scan without an object in the field of view (FOV). In a second transmission scan, an attenuating object is added to the FOV and compared to the blank scan. Another approach is the use of coupled systems for the correction of emission data. MRI provides information on proton density and the relaxation properties of spins in tissue but does not provide information on photon attenuation. Hence, generating attenuation maps with PET/MRI hybrid devices is more challenging and must be done indirectly, such as through atlas-based approaches (Chen and An, 2017). In contrast, the CT component provides direct information on the X-ray attenuation properties of tissues and is therefore suitable for generating attenuation coefficients for 511 keV radiation through bilinear transformation. In this context, a previous preclinical study has highlighted the limitations and challenges of current multicentric PET/CT imaging, proposing standardized protocols to ensure reproducibility across different sites (McDougald et al., 2020). Given these differences in attenuation correction across systems, dedicated future studies with a technical focus will further improve harmonization.

Importantly, our genuine goal was not to compare the characteristics of the three different scanners but rather to assess the feasibility and execution of a multicenter preclinical study under default conditions in which each device operates with its pre-set standard settings. Nevertheless, we conducted a preliminary study with the Micro Derenzo Phantom to identify necessary adjustments for achieving similar spatial resolution across the scanners. Given the lower resolution of the DPET scanner (Constantinescu and Mukherjee, 2009; Visser et al., 2009; Nagy et al., 2013), we applied 3D Gaussian smoothing to all PET images from the PET/MR and PET/CT with a FWHM of 1 mm, taking into account the loss of spatial specificity. Thus, we conducted PET acquisitions on both APPPS1 and WT mice, ensuring uniform housing conditions, harmonized PET scan protocols, and image reconstruction parameters. This approach effectively mitigated potential methodological bias.

Multicentric studies facilitate the inclusion of larger and more diverse cohorts for investigation. This increase in sample size may

enhance the statistical power of the study by providing more precise estimates of the effect size, thereby improving the ability to detect true differences between groups and enhancing the generalizability of the findings. However, we acknowledge that multicentric studies can also introduce additional variability due to differences in equipment, procedures, and environments across centers. This increased variability might necessitate a larger overall sample size to achieve the same statistical power, potentially conflicting with the 3Rs principle, particularly the Reduction aspect (Russell et al., 1959). In our study, the sample size calculation for a multicentric approach, which incorporated variances from all three scanners, resulted in a requirement of 20 animals (10 per group). Although this number is ultimately constrained by the least precise scanner, it is comparable to the average number required for individual scanners (20.7 animals), indicating that a multicentric design does not necessarily require more animals in this case. Still the single use of the most robust system would have resulted in less animals, which would cover the Reduction principle best. In summary, it is crucial to standardize protocols across centers to minimize variability and optimize the use of animals to comply with the 3Rs principle.

To achieve reliable and comparable results across different centers, there must be consensus on imaging parameters, such as radiotracers (Son et al., 2018), mouse model variability (Jankowsky and Zheng, 2017), acquisition times, and imaging processing methods. Achieving this level of standardization is challenging but crucial for data quality and interpretation. This issue recently gained paramount significance (Mannheim et al., 2019), especially in the context of both prospective and retrospective multicenter studies, which have remained largely unexplored in the realm of preclinical imaging in mouse models of AD.

This study sheds light on the feasibility of multimodal preclinical Aβ imaging in a mouse model of AD. Overall, our assessment of standardized small animal Aβ-PET scans provides support for the potential transition to a multicenter approach in preclinical AD research. The harmonization of image acquisition and analysis techniques has yielded a high level of comparability across datasets. However, the above-discussed variations in scanner properties may pose challenges when comparing data at the individual mouse level, underscoring the need for further methodological refinement.

CRediT authorship contribution statement

Johannes Gnörich: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Mara Koehler:** Methodology, Investigation, Formal analysis. **Karin Wind-Mark:** Methodology, Investigation. **Carolin Klaus:** Investigation, Formal analysis. **Artem Zatcepin:** Software, Methodology, Formal analysis. **Giovanna Palumbo:** Resources, Project administration. **Manvir Lalia:** Methodology, Investigation. **Laura Sebastian Monasor:** Resources, Methodology. **Leonie Beyer:** Validation, Supervision. **Florian Eckenweber:** Investigation, Formal analysis. **Maximilian Scheifele:** Methodology, Formal analysis. **Franz-Josef Gildehaus:** Validation, Supervision, Resources. **Barbara von Ungern-Sternberg:** Supervision, Resources, Project administration. **Henryk Barthel:** Writing – review & editing, Supervision. **Osama Sabri:** Supervision. **Peter Bartenstein:** Validation, Supervision. **Jochen Herms:** Writing – review & editing, Supervision, Conceptualization. **Sabina Tahirovic:** Writing – review & editing, Supervision, Project administration. **Nicolai Franzmeier:** Writing – review & editing, Visualization, Validation, Supervision, Software, Methodology. **Sibylle Ziegler:** Writing – review & editing, Validation, Supervision, Software, Conceptualization. **Matthias Brendel:** Writing – review & editing, Validation, Supervision, Resources, Conceptualization.

Declaration of competing interest

L.B. is an Novartis Radiopharmaceuticals employee, unrelated to this work. M.B. received speaker honoraria from GE healthcare, Roche and Life Molecular Imaging and is an advisor of Life Molecular Imaging. All other authors do not report a conflict of interest.

Data availability

Imaging data are available in nifti format and can be transferred per request by the corresponding author. Statistics and Graphs are available per request as GraphPad Prism Data.

Acknowledgement

We thank Mathias Jucker (Hertie-Institute for Clinical Brain Research, University of Tübingen, Germany) for providing the APPPS1 mice and LIFE Molecular Imaging GmbH for providing [18F]florbetaben (Neuraceq) precursor.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2024.120748.

References

Sevigny, J., Chiao, P., Bussiere, T., Weinreb, P.H., Williams, L., Maier, M., et al., 2016. The antibody aducanumab reduces Abeta plaques in Alzheimer's disease. Nature 537 (7618), 50–56.

van Dyck, C.H., Swanson, C.J., Aisen, P., Bateman, R.J., Chen, C., Gee, M., et al., 2023. Lecanemab in Early Alzheimer's Disease. N. Engl. J. Med. 388 (1), 9–21.

Hardy, J.A., Higgins, G.A., 1992. Alzheimer's disease: the amyloid cascade hypothesis. Science (1979) 256 (5054), 184–185.

Braak, H., Braak, E., 1991. Demonstration of amyloid deposits and neurofibrillary changes in whole brain sections. Brain Pathol. 1 (3), 213–216.

Nelissen N., Laere K.V., Thurfjell L., Owenius R., Vandenbulcke M., Koole M., et al. Phase 1 Study of the Pittsburgh Compound B Derivative 18F-Flutemetamol in Healthy Volunteers and Patients with Probable Alzheimer Disease. 2009;50(8):1251–9.

Sabri O., Sabbagh M.N., Seibyl J., Barthel H., Akatsu H., Ouchi Y., et al. Florbetaben PET imaging to detect amyloid beta plaques in Alzheimer's disease: phase 3 study. 2015; 11(8):964–74.

Xia, D., Lianoglou, S., Sandmann, T., Calvert, M., Suh, J.H., Thomsen, E., et al., 2022. Novel App knock-in mouse model shows key features of amyloid pathology and reveals profound metabolic dysregulation of microglia. Mol. Neurodegener. 17 (1), 41.

Zhao, P., Xu, Y., Jiang, L., Fan, X., Li, L., Li, X., et al., 2022. A tetravalent TREM2 agonistic antibody reduced amyloid pathology in a mouse model of Alzheimer's disease. Sci. Transl. Med. 14 (661), eabq0095.

Rominger, A., Brendel, M., Burgold, S., Keppler, K., Baumann, K., Xiong, G., et al., 2013a. Longitudinal assessment of cerebral β-amyloid deposition in mice overexpressing Swedish mutant β-amyloid precursor protein using 18F-florbetaben PET. J. Nucl. Med. 54 (7), 1127–1134.

Watamura, N., Sato, K., Saido, T.C., 2022. Mouse models of Alzheimer's disease for preclinical research. Neurochem. Int. 158, 105361.

Sasaguri, H., Nilsson, P., Hashimoto, S., Nagata, K., Saito, T., De Strooper, B., et al., 2017. APP mouse models for Alzheimer's disease preclinical studies. EMBO J. 36 (17), 2473–2487.

Nordberg, A., Carter, S.F., Rinne, J., Drzezga, A., Brooks, D.J., Vandenberghe, R., et al., 2013. A European multicentre PET study of fibrillar amyloid in Alzheimer's disease. Eur. J. Nucl. Med. Mol. ImAging 40 (1), 104–114.

Klunk, W.E., Koeppe, R.A., Price, J.C., Benzinger, T.L., Devous Sr., M.D., Jagust, W.J., et al., 2015. The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. Alzheimers. Dement. 11 (1), 1–15 e1-4.

Smith, A.M., Obuchowski, N.A., Foster, N.L., Klein, G., Mozley, P.D., Lammertsma, A.A., et al., 2023. The RSNA QIBA Profile for Amyloid PET as an Imaging Biomarker for Cerebral Amyloid Quantification. J. Nuclear Med. 64 (2), 294–303.

Fueger B.J., Czernin J., Hildebrandt I., Tran C., Halpern B.S., Stout D., et al. Impact of Animal Handling on the Results of 18F-FDG PET Studies in Mice. 2006;47(6): 999–1006.

Wong K.-.P., Sha W., Zhang X., Huang S.-.C. Effects of Administration Route, Dietary Condition, and Blood Glucose Level on Kinetics and Uptake of 18F-FDG in Mice. 2011;52(5):800–7.

Tremoleda, J.L., Kerton, A., Gsell, W., 2012. Anaesthesia and physiological monitoring during in vivo imaging of laboratory rodents: considerations on experimental outcomes and animal welfare. EJNMMI. Res. 2 (1), 44.
Syvänen, S., Meier, S.R., Roshanbin, S., Xiong, M., Faresjö, R., Gustavsson, T., et al., 2022. PET Imaging in Preclinical Anti-Aβ Drug Development. Pharmaceutical Research. 39 (7), 1481–1496.
Richards, D., Sabbagh, M.N., 2014. Florbetaben for PET Imaging of Beta-Amyloid Plaques in the Brain. Neurol. Ther. 3 (2), 79–88.
Wolk, D.A., Zhang, Z., Boudhar, S., Clark, C.M., Pontecorvo, M.J., Arnold, S.E., 2012. Amyloid imaging in Alzheimer's disease: comparison of florbetapir and Pittsburgh compound-B positron emission tomography. J. Neurol. Neurosurg. Psychiatry 83 (9), 923–926.
Sacher, C., Blume, T., Beyer, L., Biechele, G., Sauerbeck, J., Eckenweber, F., et al., 2020. Asymmetry of Fibrillar Plaque Burden in Amyloid Mouse Models. J. Nuclear Med. 61 (12), 1825–1831.
Mannheim J.G., Mamach M., Reder S., Traxl A., Mucha N., Disselhorst J.A., et al. Reproducibility and Comparability of Preclinical PET Imaging Data: a Multicenter Small-Animal PET Study. 2019;60(10):1483–91.
Dirnagl, U., International, Fisher M., 2012. Multicenter Randomized Preclinical Trials in Translational Stroke Research: it's Time to Act. J. Cerebral Blood Flow Metabolism 32 (6), 933–935.
Radde, R., Bolmont, T., Kaeser, S.A., Coomaraswamy, J., Lindau, D., Stoltze, L., et al., 2006. Abeta42-driven cerebral amyloidosis in transgenic mice reveals early and robust pathology. EMBo Rep. 7 (9), 940–946.
Overhoff, F., Brendel, M., Jaworska, A., Korzhova, V., Delker, A., Probst, F., et al., 2016. Automated Spatial Brain Normalization and Hindbrain White Matter Reference Tissue Give Improved [(18)F]-Florbetaben PET Quantitation in Alzheimer's Model Mice. Front. Neurosci. 10, 45.
Biechele, G., Sebastian Monasor, L., Wind, K., Blume, T., Parhizkar, S., Arzberger, T., et al., 2021. Glitter in the Darkness? Non-fibrillar beta-amyloid Plaque Components Significantly Impact the beta-amyloid PET Signal in Mouse Models of Alzheimer's Disease. J. Nucl. Med.
Palumbo, G., Kunze, L.H., Oos, R., Wind-Mark, K., Lindner, S., von Ungern-Sternberg, B., et al., 2023. Longitudinal Studies on Alzheimer Disease Mouse Models with Multiple Tracer PET/CT: application of Reduction and Refinement Principles in Daily Practice to Safeguard Animal Welfare during Progressive Aging. Animals (Basel). 13 (11).
Rominger, A., Brendel, M., Burgold, S., Keppler, K., Baumann, K., Xiong, G., et al., 2013b. Longitudinal assessment of cerebral beta-amyloid deposition in mice overexpressing Swedish mutant beta-amyloid precursor protein using 18F-florbetaben PET. J. Nucl. Med. 54 (7), 1127–1134.
Omidvari, N., Cabello, J., Topping, G., Schneider, F.R., Paul, S., Schwaiger, M., et al., 2017. PET performance evaluation of MADPET4: a small animal PET insert for a 7 T MRI scanner. Physics in Medicine & Biology 62 (22), 8671.
Brendel, M., Barthel, H., van Eimeren, T., Marek, K., Beyer, L., Song, M., et al., 2020. Assessment of 18F-PI-2620 as a Biomarker in Progressive Supranuclear Palsy. JAMa Neurol. 77 (11), 1408–1419.
Gnorich, J., Reifschneider, A., Wind, K., Zatcepin, A., Kunte, S.T., Beumers, P., et al., 2023. Depletion and activation of microglia impact metabolic connectivity of the mouse brain. J. Neuroinflammation. 20 (1), 47.
Bartos L.M., Kirchleitner S.V., Blobner J., Wind K., Kunze L.H., Holzgreve A., et al. 18 kDa translocator protein positron emission tomography facilitates early and robust tumor detection in the immunocompetent SB28 glioblastoma mouse model. 2022;9.
Ma, Y., Hof, P.R., Grant, S.C., Blackband, S.J., Bennett, R., Slatest, L., et al., 2005. A three-dimensional digital atlas database of the adult C57BL/6J mouse brain by magnetic resonance microscopy. Neuroscience 135 (4), 1203–1215.
Benjamini Y., Hochberg Y. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. 1995;57(1):289–300.
Constantinescu, C.C., Mukherjee, J., 2009. Performance evaluation of an Inveon PET preclinical scanner. Phys. Med. Biol. 54 (9), 2885–2899.
Chen, Y., An, H., 2017. Attenuation Correction of PET/MR Imaging. Magn. Reson. ImAging Clin. N. Am. 25 (2), 245–255.
McDougald, W., Vanhove, C., Lehnert, A., Lewellen, B., Wright, J., Mingarelli, M., et al., 2020. Standardization of Preclinical PET/CT Imaging to Improve Quantitative Accuracy, Precision, and Reproducibility: a Multicenter Study. J. Nuclear Med. 61 (3), 461–468.
Visser, E.P., Disselhorst, J.A., Brom, M., Laverman, P., Gotthardt, M., Oyen, W.J.G., et al., 2009. Spatial Resolution and Sensitivity of the Inveon Small-Animal PET Scanner. J. Nuclear Med. 50 (1), 139–147.
Nagy, K., Tóth, M., Major, P., Patay, G., Egri, G., Häggkvist, J., et al., 2013. Performance Evaluation of the Small-Animal nanoScan PET/MRI System. J. Nuclear Med. 54 (10), 1825–1832.
Russell W.M.S., Burch R.L., Hume C.W. The principles of humane experimental technique: methuen London; 1959.
Son, H.J., Jeong, Y.J., Yoon, H.J., Lee, S.Y., Choi, G.E., Park, J.A., et al., 2018. Assessment of brain beta-amyloid deposition in transgenic mouse models of Alzheimer's disease with PET imaging agents (18)F-flutemetamol and (18)F-florbetaben. BMC. Neurosci. 19 (1), 45.
Jankowsky, J.L., Zheng, H., 2017. Practical considerations for choosing a mouse model of Alzheimer's disease. Mol. Neurodegener. 12 (1), 89.