Alzheimer's & Dementia 11 (2015) 1050-1068

The influence of biological and technical factors on quantitative analysis of amyloid PET: Points to consider and recommendations for controlling variability in longitudinal data

Mark E. Schmidt[a,*], Ping Chiao[b], Gregory Klein[c], Dawn Matthews[d], Lennart Thurfjell[e], Patricia E. Cole[f], Richard Margolin[g], Susan Landau[h], Norman L. Foster[i], N. Scott Mason[j], Susan De Santi[k], Joyce Suhy[c], Robert A. Koeppe[l], William Jagust[h], for the Alzheimer's Disease Neuroimaging Initiative

[a]*Janssen Pharmaceutica, Beerse, Belgium*
[b]*Biogen IDEC, Cambridge, MA, USA*
[c]*Synarc, Newark, CA, USA*
[d]*ADM Diagnostics LLC, Chicago, IL, USA*
[e]*GE Healthcare, Uppsala, Sweden*
[f]*Takeda Pharmaceuticals, Deerfield, IL, USA*
[g]*CereSpir, Inc., New York, NY, USA*
[h]*Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA, USA*
[i]*Division of Cognitive Neurology, University of Utah, Salt Lake City, UT, USA*
[j]*Department of Radiology, University of Pittsburgh, Pittsburgh, PA, USA*
[k]*Piramal Imaging, Boston, MA, USA*
[l]*Division of Nuclear Medicine, University of Michigan, Ann Arbor, MI, USA*

Abstract In vivo imaging of amyloid burden with positron emission tomography (PET) provides a means for studying the pathophysiology of Alzheimer's and related diseases. Measurement of subtle changes in amyloid burden requires quantitative analysis of image data. Reliable quantitative analysis of amyloid PET scans acquired at multiple sites and over time requires rigorous standardization of acquisition protocols, subject management, tracer administration, image quality control, and image processing and analysis methods. We review critical points in the acquisition and analysis of amyloid PET, identify ways in which technical factors can contribute to measurement variability, and suggest methods for mitigating these sources of noise. Improved quantitative accuracy could reduce the sample size necessary to detect intervention effects when amyloid PET is used as a treatment end point and allow more reliable interpretation of change in amyloid burden and its relationship to clinical course.

Keywords: Amyloid; Positron emission tomography; Alzheimer's disease; ADNI; Quantitative analysis; Within subject variability; Multi-site trials

1. Introduction

The development of noninvasive methods for detecting amyloid plaques in human brain using positron emission tomography (PET) has made remarkable contributions to our understanding of the pathophysiology of Alzheimer's disease (AD) and related conditions, and it has facilitated more accurate and confident diagnosis. PET imaging with

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

*Corresponding author. Tel.: +32-14-60-6399; Fax: +32-1460-5353.
E-mail address: mschmid4@its.jnj.com

http://dx.doi.org/10.1016/j.jalz.2014.09.004

tracers for fibrillar amyloid-β (Aβ) (termed amyloid PET) has been widely adopted as a means of identifying and following amyloid burden by pharmaceutical companies conducting interventional trials in AD, academic research centers, and consortia studying the evolution of AD. In this setting, it has facilitated testing of a number of candidate disease-modifying treatments.

Clinical use of amyloid PET commonly relies on visual interpretation (reading) of scans, in which the signal intensity of "target-rich" brain regions such as frontal cortex is contrasted to that of "target-poor" regions such as subcortical white matter. Visual reads can be used for assessing the likelihood of significant fibrillar Aβ burden in the brain by correlating imaging findings with postmortem plaque density. The standardization and qualification of visual interpretation has been a major undertaking and is described elsewhere [1,2]. Here, we will focus on the use of quantitative measurement of amyloid PET, which requires a number of practices that are not necessary for visual interpretation.

Quantitative analysis is necessary for longitudinal observational studies and intervention trials when change in amyloid burden measured by PET serves as a treatment end point. Treatments intended to remove Aβ may have modest effects on the amyloid PET signal that are not apparent by visual comparison of scans. Detecting the effect of drugs designed to modify Aβ deposition, such as secretase inhibitors, will require identifying treatment effects against the background rate of signal change, which may be slight and affected by genetic risk factors and age. Testing for PET signal differences in studies including placebo and multiple active dose arms requires averaging across subject responses in each treatment group. Phase 2 and 3 AD trials are typically large, requiring participation by multiple imaging centers, which adds the complexity of using different scanner models that may use different image construction methods. Site practices commonly differ with regard to dose calibration and tracer administration, scanner and analytic equipment quality control (QC), and subject management after dose administration and during the scan. These factors can increase measurement noise, undermining the quality of the image data and the reliability of the results of the analyses, and consequently reducing the power to detect treatment differences. The impact of differences between sites, scanners, acquisition protocols, and subject management on image data quality is not unique to quantitative amyloid PET. These topics and more have been the focus for the Quantitative Imaging Biomarkers Alliance (QIBA) in the Radiological Society of North America and the European Association of Nuclear Medicine (EANM) to improve the value and reliability of quantitative fluorodeoxyglucose (FDG) PET/computed tomography (CT) for oncology trials [3,4]. Our goal was to bring these issues to the attention of clinical and pharmaceutical researchers considering quantitative amyloid PET as an end point so that they can be considered when acquiring and analyzing the data. To illustrate the point, inspection of individual trajectories of the ^{11}C-PiB PET cortical signals from the Alzheimer's Disease Neuroimaging Initiative (ADNI), one of the first efforts to use the method at multiple sites over time with a standard acquisition protocol [5], reveals most subjects having modest differences from scan to scan (Fig. 1). Nevertheless, several longitudinal changes follow patterns that are biologically implausible and exceed the 3% to 10% variability expected from test-retest studies of ^{11}C-PiB reported from single sites [6,7]. An important concept in quantitative amyloid PET that will underlie discussion in this article is that of amyloid positivity. In the most widely used method, cortical amyloid signal (reflecting amyloid plaque deposition) is related to a reference region that is believed not to accumulate amyloid (reference), producing a measurement known as the standard uptake value ratio (SUVR) and is further described later.

In this article, we will consider the impact of two sources of measurement variability: biological and technical. Technical sources can be further divided into factors occurring at the PET center (scan acquisition and scanner instrumentation) and post–image acquisition factors (image processing and data analysis). Beyond describing these factors and assessing their impact, we aim to identify ways to optimize quantitative amyloid PET imaging for multisite trials, to provide guidance for mitigating the most common sources of error, and to support the qualification of amyloid PET imaging as a biomarker in interventional trials.

2. Biological factors

Biological factors fall into two groups: factors that generate differences between subjects and factors affecting measurements within subjects during or between scans. The former group is relevant for understanding disease biology and progression, and it includes variables that can be considered when enrolling subjects for trials or stratifying analyses. Understanding the second group is critical for reliable measurement over the course of longitudinal studies.

2.1. Between-subject factors

Aging is a major risk factor for the development of AD reflecting the many years over which the pathophysiology evolves including accumulation of insoluble amyloid deposits. Amyloid plaques are first noted at autopsy between the fourth and fifth decade of life [8]. Longitudinal and cross-sectional amyloid PET studies that have included young and middle-aged healthy adults have identified subjects with significant cortical signal somewhat later, beginning in the sixth decade of life [9,10]. The earlier appearance at autopsy likely reflects the years of accumulation of dense amyloid plaques necessary to achieve a minimum density and binding volume for signal detection with PET, as well as the threshold selected for amyloid positivity. The incidence of healthy adults

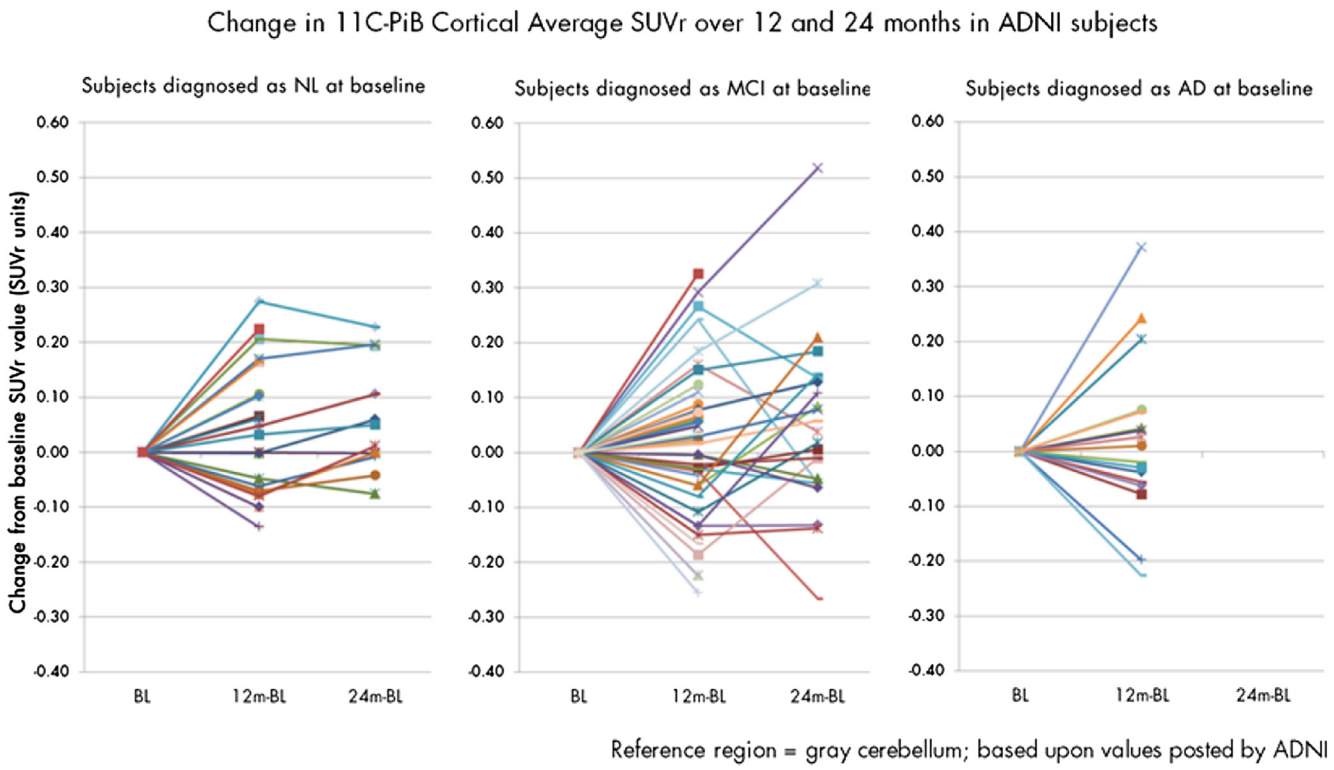


Fig. 1. Longitudinal changes in ^{11}C-PiB SUVR in ADNI-1. The change in cortical average SUVr from the first ^{11}C-PiB scan is plotted for each of 83 subjects from ADNI-1 (20 NL, 45 MCI, and 18 AD based on diagnosis at the time of first ^{11}C-PiB scan). Unique colors and symbols have been assigned to each subject in the three diagnostic groups to facilitate following individual trajectories. Cortical average was calculated using the mean of values for anterior cingulate, frontal cortex, lateral temporal cortex, parietal cortex, and precuneus cortex, normalized to gray cerebellum (source: adni_pibpetsuvr_2011_07_04). Despite a small mean change for each group and each interval relative to baseline (0.02–0.06, 1%–4%), variance is high (range −0.27 to 0.52, −13% to 29%) and trajectories change direction for several subjects. For up-to-date information on ADNI and how to access ADNI data, see www.adni-info.org.

becoming "PiB positive" in a longitudinal study conducted by Washington University, which included subjects from 45 to 86 years old, was estimated to be 3% per year, and the risk of conversion to amyloid positive appeared to rise with advancing age [9]. The annual incidence rate of becoming PiB positive in a recent study from the Mayo Clinic, evaluating subjects on average 10 years older than those evaluated by Washington University, was estimated to be 13% per year [11]. The prevalence of subjects who were florbetapir positive in their eighth decade in the study by Fleisher et al. [10] was 40%, double the prevalence of subjects in their seventh decade. Aging is also associated with susceptibility to genetic risk factors and emergence of other risk factors for AD such as cerebrovascular disease. How these other factors interact with measurement of amyloid burden on PET is considered later.

Genetic background can strongly influence amyloid burden. The amyloid precursor protein (APP) and presenilin mutations in families with dominantly inherited early-onset AD, and the overexpression of APP resulting from trisomy 21 in Down syndrome, are associated with increased proteolysis of APP (APP mutations) or reduced production of shorter, more soluble protein fragments in favor of longer, more amyloidogenic fragments (presenilin mutations). As in sporadic AD, significant brain amyloid burden appears years before symptoms of dementia [12,13] although accumulation of amyloid deposits may start earlier than in sporadic AD. The location of the mutation may influence the timing, rate, and location of accumulation (reviewed in the study by Karran et al. [14]). Amyloid PET signal in striatal regions has been reported to be relatively higher early in the disease course of familial AD compared with that in the sporadic AD [12,15]. Carrying one or more *APOE* ε4 alleles is the major risk gene associated with sporadic AD and lowers the age of onset compared with noncarriers [16,17]. Cross-sectional amyloid PET imaging studies including healthy elders, subjects with mild cognitive impairment (MCI), and subjects with probable AD have reported a higher frequency of "positive" scans in *APOE* ε4 carriers [9,18,19]. Modeling of florbetapir cortical SUVRs by age indicated that significant amyloid burden could occur as early as 58 years in *APOE* ε4 carriers, 20 years before the first noncarrier reached threshold in their sample [20], suggesting that carriers begin accumulating amyloid earlier than noncarriers. An impact on the rate of deposition is suggested by a longitudinal study with PiB imaging in well-characterized subjects with probable AD, which reported a gene-dose effect on change in cortical PiB signal over 2 years

[21]. This was a small study with a very limited number of homozygotes, and a change of signal might also occur as a consequence of longer opportunity for posttranslational modification of plaques if deposition begins earlier in carriers, resulting in greater binding capacity for PiB. Cumulatively, amyloid PET data stratified by carrier status are consistent with the idea that the *APOE* ε4 allele can accelerate the onset of AD, although carrier status does not appear to affect duration of disease [22].

Gender is another possible genetic contributor to amyloid deposition. An age- and gender-dependent interaction of *APOE* ε4 and plaque density has been reported at autopsy [23] although to date gender effects on amyloid burden and change have not been identified. Vascular disease is a common antecedent to cognitive impairment and dementia [24], and cerebrovascular pathology has been observed more frequently at autopsy in AD than other neurodegenerative disorders. Cerebrovascular disease has been proposed to increase the risk of dementia in AD, especially in early stages of disease [25]. Unmedicated hypertension was found to be correlated with higher amyloid burden measured by florbetapir in cognitively healthy elderly subjects carrying one or more *APOE* ε4 alleles [26]. Arterial stiffness was explored as a risk factor for brain amyloid burden and change over 2 years in a sample of nondemented elderly using ^{11}C-PiB PET. Peripheral versus central arterial stiffness contributed independently to amyloid burden and change [27]. No interaction was seen with *APOE* ε4 carrier status although these subjects were on average 10 years older than those in the study testing the role of hypertension, suggesting an interaction between vascular and genetic risk factors, so that the impact of vascular factors on brain amyloid PET measures may be age dependent.

A particular vascular pathology that occurs in AD is cerebral amyloid angiopathy (CAA). Many subjects with AD develop significant CAA in addition to parenchymal plaques. Current tracers do not discriminate the two types of amyloid deposition, and nondemented subjects with probable CAA have been reported to manifest significant brain amyloid burden with ^{11}C-PiB PET, notably having higher signal in the occipital cortex relative to other cortical regions compared with AD subjects [28]. The extent to which CAA contributes to amyloid signal poses a problem for interpretation of signal in intervention trials, for example, testing whether amyloid-targeting treatments affect vascular amyloid deposits or whether CAA load could predict the emergence of amyloid-related imaging abnormalities during treatment [29,30].

Demonstrating an effect on amyloid PET signal may depend on the type of amyloid deposited. The in vivo cortical signal detected with current tracers has been linked to dense plaque burden postmortem both quantitatively [31] and semiquantitatively [32]. Indeed, one subject has been reported in whom a ^{11}C-PiB PET scan acquired 3 years before death failed to detect significant diffuse amyloid burden observed at autopsy, in the setting of low cerebrospinal fluid (CSF) $A\beta_{42}$ and mild dementia of the AD type [33]. On the other hand, in vivo PiB signal has been reported in association with postmortem findings of high diffuse plaque burden [34]. Moreover, retention of PET tracer is seen in the striatum early in the course of disease in familial AD due to certain mutations, and this region is known to harbor large numbers of diffuse plaques [35,36]. In addition, cerebellar cortex manifests large numbers of diffuse plaques postmortem in late-stage disease and in some types of familial AD, leading to some concern about relying on this area as a reference region [37].

Lifetime education and occupational achievement can modulate the onset of cognitive decline and AD dementia [38]. Higher education appears to delay clinical impairment in the face of disease progression, potentially by calling on neural reserve or other compensatory mechanisms [39]. Whether cognitive activity could mitigate disease progression has been explored in a cross-sectional study of amyloid burden in healthy elderly subjects. Those subjects with the highest level of cognitive engagement had a significantly lower cortical amyloid burden, measured with ^{11}C-PiB, than those with lower levels of engagement [40].

2.2. Within-subject factors

A limited number of biological factors may contribute to within-subject amyloid PET signal variability. These can be broadly divided into factors that can affect the tissue signal and factors that can affect tracer delivery. Progressive cortical atrophy has the potential to alter the tissue signal over the course of AD trials, as loss of tissue volume can result in a drop in the amyloid signal in the absence of change in amyloid burden. This phenomenon, termed partial volume effect (PVE), can alter serial PET measurements if the regions of interest (ROIs) are defined on the baseline scan and are then applied to all other subsequent scans. How this can be addressed by partial volume correction is discussed later. Possible reduction of the amyloid PET signal by a drug treatment that competes with tracer binding also needs to be considered. The binding site(s) for thioflavin- and stilbene-derivative amyloid tracers is not fully understood although binding pockets are hypothesized to occur across strands of Aβ within the β-sheeted fibrils [41]. In the phase 2 study testing the effect of the amyloid-targeting monoclonal antibody bapineuzumab on brain amyloid using ^{11}C-PiB PET, competition studies using antibody concentrations well in excess of predicted tissue levels showed no difference in binding of ^{3}H-PiB to Aβ fibrils [42]. The epitope for bapineuzumab is the amine terminus of Aβ; whether treatments targeting other regions of amyloid peptide or fibril aggregation compete with amyloid PET ligand binding may need to be confirmed. Interference with drug metabolism, which is frequently found in pharmacology and test drug exposures, can change rapidly over the course of a clinical trial if concomitant medications that are metabolic enzyme inducers or inhibitors are introduced. PET

ligands can be substrates for these enzymes, and inhibition of metabolic clearance can significantly affect the "input function" of the tracer or radiolabeled metabolites [43]. Currently available amyloid PET ligands are metabolized very rapidly (eg, see the article by Wong et al. [44]). Whether metabolism of PiB is susceptible to interference by inhibition of metabolism has been tested. Metabolism occurs via several metabolic pathways rendering clearance of PiB less susceptible to inhibition by a particular metabolic enzyme inhibitor [45]. Whether other PET ligands are similarly cleared via multiple pathways and whether alteration in metabolic clearance could occur as a result of drug interaction or liver disease have not been reported. Delivery to and clearance of the tracer from brain is highly dependent on blood flow, which differs across tissue types [46] and is significantly reduced in AD, particularly in gray matter regions known to be susceptible to neurodegeneration and atrophy [47]. Unlike with FDG PET, a subject's mental state is not thought to affect amyloid PET results although testing for effects of major differences in attention and arousal, such as sedation, has not been conducted in humans. The critical dependence on blood flow for delivery of the tracer to the brain underlies the need for strict control over tracer uptake and emission data acquisition and different approaches to data collection outlined later.

3. Acquisition factors

PET image acquisition involves a series of technical processes, and a thorough description of how images are generated is beyond the scope of this review. Nonetheless, a brief summary may help provide an orientation to how acquisition factors influence quantitative analysis of amyloid PET. Further information on PET data acquisition can be found in recent textbooks [48].

The basis of PET signal generation lies in a γ photon, two of which are produced by mutual annihilation after a positron emitted from the PET isotope collides with a neighboring electron. The photons travel away from the annihilation site in the opposite directions and can be detected by a pair of detectors located opposite one another in the PET camera. This defines what is called a "line of response" (LOR) through the part of the body being imaged and indicates that a PET radioisotope was somewhere along the line. Gamma photons continue to be produced from the radiolabeled tracer in the tissue generating LORs in all directions. Detectors will also record γ photons that have ricocheted off their original LOR after colliding with electrons during their transit away from positron annihilation (scatter) and other stray events. Contrast in the image depends on more signals emanating from the radiotracer bound to the target of interest than radiotracer retained in target-poor areas or from radiolabeled metabolites and reducing noise such as scattered counts. The first form of data is the recording of the detected events in each LOR. These data are represented in a sinogram according to the spatial orientation of the LORs to facilitate subsequent computations. Emission data can be collected in two or three dimensions (2D or 3D); the latter has greater sensitivity and is available on most modern PET scanners. Data can be recorded continuously (list mode) or over discrete intervals resulting in "frames" (from time frames) of data. Collection can begin when the tracer is injected and followed for an extended period to allow full characterization and modeling of tracer kinetics (dynamic scans) and is more fully described later, although for diagnostic scans, collection usually begins long after tracer injection. Data are typically collected over 5 minute periods for up to 10 or 20 minutes to record enough counts resulting in two to four frames of PET data (static scans). These frames can be inspected separately for evidence of movement during the scan, and acceptable frames are then combined to form the complete emission data file. These emission data are corrected for signal loss or attenuation by bone and other tissues calculated from a transmission scan: a scan generated by an external positron source such as a rod containing ^{68}Ge germanium rotated around the bore of PET-only cameras, or more commonly by a scan generated by x-ray in PET/CT cameras. The emission scan is also corrected for noise resulting from scattered and random counts (i.e., counts that are not representative of paired photons from a single annihilation or "true events"), signals missed by the detectors when they are busy recording counts (dead time), and corrections for sensitivity differences between the detectors (obtained from "normalization" scans performed as part of camera QC). The corrected emission data are then reconstructed to form an image that is a quantitative estimate of the spatial distribution of the radioisotope. All these corrections occur on the sinogram, and accurate image reconstruction is critically important for quantification purposes. Reconstruction occurs by using either filtered back projection (FBP) or iterative algorithms such as ordered-subsets expectation maximization (OSEM). Each of these processing steps can involve different assumptions and algorithms unique to the scanner type, and how these can impact measurements in amyloid PET is considered later.

3.1. Scan acquisition

Although the measure of interest in amyloid PET is amyloid burden, the PET signal is a composite of radioactivity from tracer concentrations in multiple "compartments": the blood into which the tracer was infused and from which it will be cleared, a "free" tissue compartment in which the tracer has crossed the blood-brain barrier but is not bound, the specific amyloid-binding compartment of interest, and a nonspecific binding compartment [49]. Two primary approaches are available for scan acquisition, each with advantages and limitations.

The "gold standard" for quantitative PET image analysis is full kinetic modeling with arterial blood sampling, in which the scan is collected as a series of contiguous time

frames from tracer injection to past a point at which equilibrium is reached. The time course of radioactivity signal at any point within the brain forms a "time-activity curve." This includes an initial period of a few minutes during which the signal largely reflects tracer influx from plasma into tissue, highly correlated with regional cerebral blood flow rate [50,51]. A nonequilibrium stage follows during which specific binding approaches a plateau, followed by a period of relative equilibrium. The time to reach equilibrium varies with the particular tracer, the region of accumulation, and amyloid load [44,52,53]. With the plasma input concentration known, a set of differential equations characterizing the tracer influx and efflux from each compartment are solved to achieve a best fit to the acquired curve. To eliminate the discomfort and logistical issues associated with arterial blood sampling, simplified models have been developed that use ratios of time-activity curves in target regions to those in a region of tissue in which specific binding is negligible (reference region) without requiring blood data. Amyloid burden is expressed as the quantity distribution volume ratio (DVR) [9,49]. Although some models have demonstrated better robustness, noise, and bias characteristics than others [54], by using information from the time of injection until tracer activity in tissue approaches relative equilibrium, all are able to dissociate PET signal contributions from blood flow and rate of tracer clearance. This can define and thereby minimize erroneous impact of these factors on amyloid burden estimates.

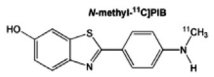
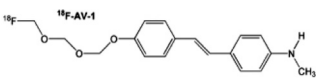
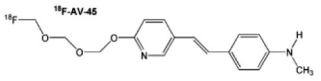
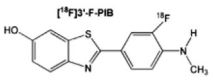
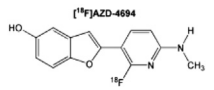
Although dynamic scanning is the gold standard for quantitative measurements, as we noted previously, the most widely used approach to amyloid PET is the use of the average activity measured over late time frame, "pseudo-equilibrium" phase scans, in which a ratio of activity in tissue regions known to accumulate high amyloid burden to that in a reference region with presumably no dense fibrillar amyloid (SUVR) is used to estimate amyloid burden. This minimizes the time that the patient must spend in the scanner, reduces acquisition and analysis complexity, and does not require a scanner capable of acquiring dynamic data and is by far the most practical. A high correlation has been established between SUVR and DVR measures, with both resulting in significant discrimination between AD patients and normal controls [44,53,55] and in similar rates of amyloid accumulation in longitudinal studies [21,56]. A limitation is that the SUVR approach, which does not use early frame information, does not distinguish contributions of blood flow and clearance. One effect is that SUVRs may overestimate amyloid burden, a bias that increases with later time windows, because of a lack of true equilibrium between plasma and tissue as the tracer clears from plasma [57]. If the tracer delivery and clearance is similar within a subject group, this bias should result in a simple scalar difference in empirically generated thresholds for SUVR versus DVR values [58]. However, differences across subject groups and longitudinal changes in blood flow and/or clearance within subjects are more difficult to address. A study in AD, MCI, and healthy controls using ^{11}C-PiB showed that although longitudinal reductions in late frame SUVR values were observed in AD subjects, there was little or no change when dynamic modeling was used. The decline in SUVR values in AD subjects paralleled a decrease in the uptake rate of the tracer from blood into brain: detected by the "k1" rate constant derived from kinetic modeling but not from SUVR calculations. Thus, apparent decreases in SUVR values may be driven by reductions in cerebral blood flow in AD [58], which is a consideration for longitudinal studies following disease progression. In the following discussion, we assume SUVR will be used as the quantitative measure for amyloid PET studies.

Other approaches have been proposed with the goal of improving accuracy. These include a bolus plus infusion approach that provides true equilibrium between plasma and tissue in late time frames [57] but presents implementation challenges and a regional equilibrium correction that reduces the bias in late time frame estimations [59]. Regardless of the acquisition and analysis approach used, however, amyloid burden measurement can only be valid if the raw PET signal is without artifact.

3.1.1. Radiotracer characteristics

Several radioligands have been developed for imaging of fibrillar amyloid in the brain. The chemical structure and standard late time frame acquisition periods for ^{11}C-PiB [55] and the ^{18}F fluorine–labeled tracers florbetaben [18,60], florbetapir [61], flutemetamol [53,62], and NAV4694 [63,64] are provided in Table 1. The use of ^{11}C carbon as the radiolabel, with a 20-minute half-life versus ^{18}F fluorine with a half-life of 110 minutes makes a huge difference for multisite trials. Radiolabeling with ^{11}C tracers requires on-site production, whereas radiolabeling with ^{18}F allows for regional distribution and the potential for scanning more than one subject per tracer production driving the adoption of ^{18}F tracers for intervention studies. Preclinical profiling and autopsy studies have provided evidence that cerebral cortical retention of these tracers in vivo correlates with the burden of neuritic plaque load seen in the same brain regions postmortem [31,32,34] and at biopsy [65]. They rapidly enter the brain by passive diffusion across the blood-brain barrier, and the first-pass extraction fraction is likely high, resulting in delivery being as highly dependent on blood flow as described previously. An important source of variability in amyloid measurement comes from differences between tracers in white tissue residence time and uptake in gray matter. High-affinity binding of the tracer to fibrillar amyloid results in much slower clearance from gray matter bearing amyloid plaque. Clearance from white matter depends on the absence of high-affinity binding targets such as amyloid plaque but also the degree to which nonspecific low-affinity binding occurs; for example, more lipophilic tracers can have a larger distribution volume in highly lipid white matter. White matter may also have

Table 1
Amyloid PET tracers currently available for research or commercial use

Tracer	Structure	Recommended dose	Recommended acquisition time post injection (p.i.) of tracer from the literature/allowable range in the FDA label
^{11}C-PIB		500 MBq	50–70 min p.i.
Florbetaben		300 MBq	90–110 min/15–20 min scan beginning 45–130 min p.i.
Florbetapir		370 MBq	50–70 min/10 min scan beginning 30–50 min p.i.
Flutemetamol		185 MBq	90–110 min/20 min scan beginning 90+ min p.i.
NAV4694		200 MBq	40–70 min p.i.

Abbreviations: PET, positron emission tomography; FDA, Food and Drug Administration.

NOTE. Subtle differences in the chemical structure can alter the clearance rate from the brain, for example the introduction of a fluorine atom results in flutemetamol being more slowly cleared than PiB while introduction of a single nitrogen atom results in florbetapir being more rapidly cleared than florbetaben.

specific binding targets and can show at least regional ^{11}C-PiB signal reductions associated with white matter disease [66]. As a consequence of their unique physical and chemical characteristics, gray versus white matter uptake and retention varies between tracers [67]. For those tracers that have been compared "head to head," the retention in cerebellar gray matter is highest for ^{11}C-PiB, followed by flutemetamol and then florbetapir, whereas white matter retention is highest for flutemetamol followed by ^{11}C-PiB and then florbetapir [68]. Retention in gray and white matter has been reported to be equivalent between ^{11}C-PiB and NAV4694 [64]. Whether differences in tissue retention among the tracers translate into material differences in detecting presence or change of an amyloid signal has not been determined. Approaches to reconcile these differences have been proposed [68] and are further discussed in the analysis section.

Independent of the tracer selected, variability in tracer composition and activity can impact measured signal, and QC is therefore essential. Generation of a good imaging signal depends on production of the radiotracer with sufficient radioactivity, purity, and specific activity, and a number of specifications need to be met to ensure the safety and reliability of the dose. In the United States, the guidance for production of marketed PET radiotracers is the PET Drugs Current Good Manufacturing Practice 21 CFR Part 212 [69]. Investigational radiotracers produced under an Investigational New Drug application can follow either Part 212 or United States Pharmacopeia Chapter <823> [70]. There are similar guidance documents for production of radiopharmaceuticals in the European Union [71,72]. Trial sponsors may need to ensure these standards are met during the trial through expert review of the production records and quality assurance checks. Failure to meet these standards can result in loss of authorization to release the tracer for clinical studies until the deficiencies have been corrected, resulting in a loss or delay in getting scans.

3.1.2. *Patient motion*

Variability due to patient motion is likely one of the largest contributing factors to measurement noise in amyloid PET studies. There are two components to motion-induced error. First, mismatch between the PET transmission or CT attenuation and emission data can induce attenuation-correction artifacts usually seen near the outer boundaries of the brain. The attenuation properties of the head are mostly uniform, so some types of rotational motion often do not dramatically affect the result of attenuation correction. Translations can be more troublesome. Quantification errors in average pixel counts for cortical ROIs have been estimated to be 10% for translations of 5 mm and >20% to 40% for translations >10 mm [73]. An example of a motion-induced attenuation-correction artifact is seen in Fig. 2.

Second, motion during the emission scan itself will result in misaligned or blurry images so that the brain tissue sampled in a ROI at the early portion of the acquisition will not correspond to tissue sampled during later phases of the scan. Anecdotally, the motion seen between dynamic 5-minute time frames in typical amyloid scans from the ADNI study ranges from translations and rotations of <0.5 mm and 0.5°, respectively, in well-controlled subjects, to ≥5 mm translation and 5° rotation in more difficult cases. These are motion estimates relative to the center of the

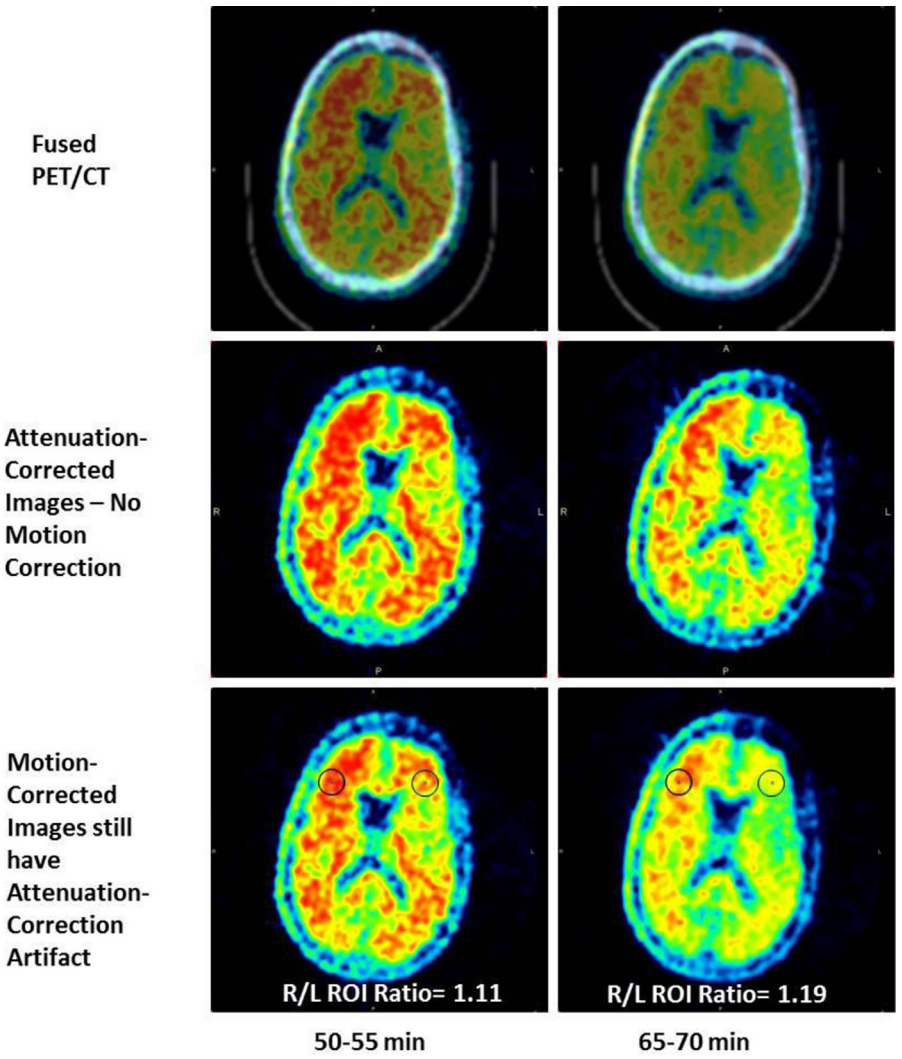


Fig. 2. Subject motion between acquisition frames. Examples of a PET/CT scan with significant motion and artifacts due to attenuation correction. Fused data set (top) shows moderate PET/CT misregistration already apparent during the first 5-minute frame (left column), and severe errors by the fourth frame (right column). In this case (middle row), the correction biases are visually evident even after motion correction by registering all time frames to the first (bottom row), the attenuation-correction error can induce large differences in quantification, as demonstrated by the visible difference in signal on the right and left sides of the brain and reflected in the ratios of regional uptake in simple circular ROIs on the bottom of the images.

imaging volume, so actual motion in the peripheral cortex could be considerably higher, and variability in non-ADNI studies could also be greater. Qualitatively, one can infer that because ROIs are usually specified on thin gray matter regions and because uptake differences between gray and white matter can be large, uncompensated motion error of this extent could add significantly to measurement noise. Few studies of this effect have been published. One from the Japanese Alzheimer's Disease Initiative reported changes in ROI values in the cortex of $>10\%$ due to motion correction in the severest cases [74] although the frequency and impact of more typical movements during scans are not well understood.

Methods to combat motion start with good subject management before the scan. It is essential for site technical staff to be aware of appropriate placement in the scanner and help subjects get into as comfortable a position as possible before the start of the scan to minimize movement and emphasize the importance of remaining as still as possible during acquisition. Multiple methods including thermoplastic masks, adjustable pillows, applying straps across the forehead, and tape across the forehead and chin are used by sites. There has been no standardization or systematic comparison of the efficacy of these methods. Minor movements that occur during dynamic scans can be managed by dividing the acquisition into separate short time frames ($\leq$5 minutes) that may be retrospectively realigned or omitted from the analysis [5]. Monitoring of subject motion during image acquisition is also important, at a minimum using visual observation and noting if excessive motion occurs to provide information during image QC. Automated systems for motion detection and correction have been

developed but have not been standardized and unfortunately are not typically used in clinical trials. Severe motion that induces attenuation-correction artifacts is difficult to address but image QC should attempt to identify such cases as described in the analysis section.

3.1.3. Variability due to uptake and scan time differences

Although late-phase imaging captures the brain uptake characteristics in a state of approximate equilibrium, variances in the emission start time after tracer injection or in the length of scanning time can introduce variability in SUVRs. Emission start time is frequently delayed from a specified target time because of difficulties while positioning patients or simply because of inattentiveness at the imaging site. A significantly delayed start time can also occur in cases where same-day rescans are required because of technical issues in the first attempt at a scan. Differences in the length of scanning time can occur if time frames are omitted because of motion or if the imaging protocol in a follow-up scan was not identical to baseline settings. As an example of potential differences in SUVR values due to timing differences, Nelissen et al. [53] showed a progressive increase from about 2.0 to 2.5 in mean frontal cortex SUVR values computed in 5-minute intervals between 85 and 170 minutes, representing a 25% difference. Similar, but more subtle, effects are seen qualitatively in increasing the scan lengths from 5 to 20 minutes. Therefore, it is important to match imaging times for individual subjects' baseline and follow-up visits to avoid unnecessary measurement variability. If a delay in start time occurs at baseline, then this should be documented as the target start time for follow-up visits. The standard suggested by the QIBA for follow-up FDG SUV measurements is $+10/-5$ minutes of the target time [3]. A similar time window appears suitable for longitudinal amyloid studies. Practical experience shows that with proper training, most sites can comply with targeted start times to within ± 5 minutes.

4. Instrumentation factors

4.1. Spatial resolution effects

Native spatial resolutions of scanners range from just over 2-mm full-width half maximum (FWHM) for research scanners such as the Siemens ECAT HRRT to >7-mm FWHM for older, but still commonly used PET scanners. In practice, postreconstruction Gaussian filtering further reduces image resolution to 7- to 12-mm FWHM [75]. This introduces a PVE, where pixel values from a region are blurred into and combined with those from neighboring regions. The phenomenon is fairly easy to describe for "hot spot imaging" typical of tumor quantification, where a recovery coefficient can be defined as the ratio of the observed activity concentration to the true concentration. Typically, the recovery coefficient is <100% for objects <30 mm in size [76]. The PVE for amyloid images is more difficult to describe. White matter tracer uptake is nonspecific and has roughly the same activity in normal and diseased subjects for a given tracer. As noted previously, gray matter retention is lower than white matter uptake for several currently available tracers in regions devoid of amyloid and is approximately equal or greater than white matter uptake in amyloid-positive regions. The PVE tends to blur high activity from white matter into gray matter for normal subjects (thereby increasing gray matter SUVR values), while having less of an effect for diseased subjects because of the relative uniformity of the uptake in gray and white matter. Similarly, PVE can blur low activity from CSF, which increases in volume surrounding gray matter as atrophy occurs. Therefore, the PVE at different spatial resolutions causes different effects for normal compared with diseased subjects. An estimated -7% to 5% shift in SUVR values has been estimated for differences in spatial resolution of 9-mm FWHM [77].

Joshi et al. [78] proposed a method for spatial resolution compensation based on calibrated Hoffman phantom scan data acquired at all imaging sites in a clinical trial. This "resolution matching" approach smooths all data to the effective resolution of the lowest resolution scanner in the study. The technique was adopted by the ADNI study, where all images were smoothed to an effective resolution of 8-mm FWHM [79]. This approach has intuitive advantages for cross-sectional quantitative measurements and qualitative visual reads used for enrollment criteria. All data appear much more uniform to the readers or to the quantitative method. However, it has a disadvantage that potentially useful high-resolution signals from the best scanners in the study are smoothed away. This can be mitigated by specifying minimum scanner performance standards for a site to be qualified to participate. Moreover, reconciliation of resolution differences does not address what the authors identify as "low-frequency" noise sources such as the scatter correction methods used by different scanners. The value of spatial smoothing for longitudinal studies where within-subject change is the primary end point is debatable because the nonspecific activity in the white matter of each subject is believed to remain stable during treatment, so the spillover of activity from the white matter may remain the same between baseline and follow-up. If this is the case, the PVE may not significantly contribute to change from baseline in SUVR values.

Partial volume correction is another potential method for dealing with variability due to the PVE [80–83]. Partial volume correction methods attempt to recover the true value of the PET signal by estimating and adjusting for spillover contributions from neighboring white, CSF, or other gray matter tissue. Some techniques use only the PET image information and knowledge of the scanner's resolution, whereas other methods also make use of the structural gray, white, and CSF information available from the subject's corresponding magnetic resonance image (MRI) [84]. In support of PVE correction, its application was found to increase the accuracy of SUVR measurement

in AD simulation phantom studies and in patients, recovering up to 56% of lost signal due to PVEs and increasing group discrimination [82]. In a separate study, signal recovery ranged from 20% to 40% and cross-sectional discrimination between normal and AD groups was increased despite increased variance [83]. In both cases, the researchers strongly recommended incorporation of PVE correction for clinical studies. However, other studies have not found a difference when correcting for PVEs [85] or have found that increased variance due to not consistently acquiring MRIs on the same scanner over the course of a study offsets benefits of increased accuracy [56]. Increased variability is of particular concern in longitudinal studies, where accumulation rates and measurable changes are very small. Although a consensus has not been reached on when and which correction to use, there is agreement that PVE correction is highly method dependent and that its use must be applied consistently and with caution, including recognition that positivity threshold values will change with the method applied and that reference MRIs must be acquired on the same scanner to be reliable when using MRI-based PVE corrections.

4.2. Attenuation correction

Most PET scanners currently available to clinical trials are integrated PET/CT systems, yet it is still common for a multicenter study to also include PET-only systems. Nakamoto et al. [86] compared quantitative differences between CT and germanium-based (for PET-only systems) attenuation correction in various organs including the brain and found increases of about 5% in both the temporal lobe and cerebellum for the CT-based method. Therefore, it is recommended to scan the subject on the same scanner throughout the study and not to interchangeably use a PET/CT and PET-only (or PET/MRI) system on baseline and follow-up scans of the same subject. Direct comparisons of the effect of different attenuation correction methods on amyloid SUVR quantities have not been reported in the literature.

Another aspect of CT-based attenuation correction that may have a small effect on quantification is the CT settings such as tube voltage, tube current, slice thickness, and reconstruction kernel. Although in principle, diagnostic-quality CT scans using high tube current and voltages should produce attenuation-corrected PET data with less noise than with low-dose CT protocols, phantom studies have shown that differences compared with low-dose CT parameters yield at most a 2% variation to PET noise [87,88]. Therefore, to minimize radiation exposure, a low-dose CT scan is suitable for amyloid studies with a tube voltage of approximately 120 kVp, a tube current of 10 to 30 mA, and slice thickness approximately the same as the native PET slice thickness (1.22–4.25 mm depending on the scanner) [87,88]. To minimize variability, the same CT settings should be used for baseline and follow-up visits, and it is recommended to set up all PET/CT scanners in the study with similar CT parameters.

More recently, attenuation techniques are becoming available that use mathematical modeling instead of data sets acquired from CT or positron emission sources. The Chang algorithm [89] is one attenuation modeling method, and many different analytical attenuation correction methods have been proposed for PET/MRI systems. Because of differences in the cerebellar regions compared with CT methods for the mathematical techniques [90] and for PET/MRI methods [91], these two attenuation correction methods need to be validated against CT and external positron source methods before use in amyloid quantitative analyses. Attenuation correction for PET/MRI is a highly dynamic area of research and may be highly vendor specific. Likewise, the scatter correction techniques used by PET/CT and PET/MRI differ. Cumulatively, the differences in attenuation and scatter correction in current generation PET/MRI platforms could introduce large quantitative differences relative to measurements with PET and PET/CT systems. Further validation of PET/MRI is needed before it can be recommended in quantitative amyloid PET multicenter studies.

4.3. Acquisition mode, scatter correction

Selections of 3D versus 2D acquisition modes and scatter correction are interrelated factors in the acquisition process. PET data may be collected in the 3D mode using coincidence detection between all rings or 2D mode where coincidence detection is only between nearby detector rings. A few scanners are only capable of 2D acquisition, but most modern scanners operate only in 3D mode for brain imaging. The 3D mode offers much improved sensitivity, but data are affected much more by scattered events. Scatter correction methods differ between PET vendors and system models, and quantitative biases in SUVR calculations due to different or incorrect scatter correction have been reported [92,93]. Moreover, correction is not uniform, for example, differing between inferior planes encompassing the cerebellum and superior planes encompassing cortical regions. Because of this phenomenon, consistent patient positioning in the axial center of the scanner field of view (FOV) is important for minimizing longitudinal variability due to scatter correction differences.

4.4. Reconstruction methods

The most common reconstruction methods are FBP or OSEM. More recent scanners offer methods to further improve signal to noise using time of flight (ToF) or spatial resolution using resolution recovery techniques incorporating a spatially variant point spread function. Iterative techniques have become the standard for qualitative visual analysis because of their improved noise properties over FBP. Likewise, iterative techniques have been used in

most multicenter quantitative analyses of amyloid data published recently. Doot et al. [94] has found relatively small differences due to analytical and iterative techniques in quantitative analysis of phantom data, provided smoothing levels were matched and a reasonably large number of iterations and subjects were used to ensure algorithm convergence. Resolution recovery and ToF techniques are relatively new, not available on all scanners, and the potential improvement on amyloid imaging is not fully understood at this point, so these methods are not currently implemented in a multicenter clinical study.

4.5. Scanner field of view, sensitivity, and noise equivalent counts

Axial FOVs for modern PET scanners range from approximately 140 mm to >200 mm. Scanners with shorter axial FOV run the risk of subjects not being positioned well, resulting in truncation of the inferior portions of the brain such as cerebellum, or inclusion in the noisy planes at the edge of the FOV. Scanners with longer axial FOVs offer advantages in two ways for amyloid imaging. First, the overall scanner sensitivity is increased because of the increased number of crossdetector planes in a 3D acquisition. Second, most subjects can be positioned so that important reference regions such as the cerebellum are not imaged in the outer planes, where sensitivity is lowest.

Overall sensitivity is another scanner property that could cause measurement variability. Ideally, an imaging protocol could be designed to obtain equivalent counts from scanners with low sensitivity by increasing acquisition time. However, scanner sensitivities can differ by more than a factor of two, making it impractical to harmonize scanner sensitivities. Although some studies have shown no quantitative bias in SUVR for low versus high statistics amyloid acquisitions [53,61], the observations that SUV estimates of FDG uptake tend to increase with decreased signal noise [95] and that visual analyses are qualitatively more difficult to assess with more noise could be adequate reasons for rejecting a scanner in a clinical trial with a small axial FOV or low sensitivity.

4.6. Scanner qualification and QC issues

Current best practice for PET imaging QC in a multicenter study requires acquisition and analysis of a phantom data set at each imaging center before the start of the study as a qualification step to verify adequate image uniformity, signal to noise, and sensitivity. The phantom scans are also used to calibrate the spatial resolution properties of the scanner. Throughout the course of a trial, scanner quality must be maintained, and this is generally achievable through manufacturer-recommended daily QC. Long-term global drifts in scanner sensitivity are not likely to affect SUVR measurements as any change should apply uniformly to target and reference regions. More important are QC issues that could regionally change scanner sensitivity or uniformity, such as streak artifacts caused by bad detector blocks or axial nonuniformities due to normalization errors. A detailed description of QC procedures is beyond the scope of this document, but daily automated and visual checks of these issues at the site and central core laboratory checks for these issues as data are collected are crucial to maintaining high-quality data. Scanner upgrades or hardware repairs may require requalification to verify that the scanner is performing according to the recorded baseline performance.

It is highly advisable to acquire all longitudinal scans for a subject using the same scanner, with the same acquisition and reconstruction parameters, because longitudinal follow-up using a different scanner can introduce substantial variability in SUVR measures. In general, differences between scanners within the same vendor are less than differences between vendors due to software differences; however, before allowing longitudinal data acquired from any form of different scanner vendor or model, a phantom validation study should be done to evaluate differences. Unfortunately, very little data exist to estimate the potential SUVR variability likely to be induced by a scanner change. The approach using Hoffman phantom data to calibrate for spatial resolution differences has been adopted by ADNI and many pharmaceutical studies; however, the authors of this technique acknowledge that modeling differences due to scatter, attenuation correction, and differing brain sizes is difficult and that it may not be possible to calibrate scanner differences using phantom data alone [78]. In most cases, it is likely that correction factors cannot be found to compensate for all differences between different scanners used on the same subject, and additional scans should not be collected for a subject in an intervention trial after a scanner change.

5. Image processing and analysis factors

Transforming reconstructed PET image data to a quantitative measure of amyloid burden involves several steps, and each of these can significantly affect the values derived from the image. The first step is a QC inspection of the image files to identify issues that may have arisen during acquisition. This is followed by image transformation and measurement, where QC of the results of the various image postprocessing steps is also essential. Because the final values are so contingent on the methods used, all the steps followed and software used should be described to ensure the reproducibility and reliability of the results.

6. Image QC factors

A summary of quality issues that can be used as the basis of data QC is presented in Table 2. Verification of compliance with the acquisition protocol can be checked by reviewing information reported in the "header" file. This file is attached to the image file and contains data fields identifying

Table 2
PET image quality control checks

Category	Quality item
Protocol adherence	Injected dose
	Injected volume
	Scan initiation time after injection
	Scan duration
	Emission acquisition (3D vs. 2D)
	Number of time frames
	Duration of each frame
	Transmission scan performed before or after emission scan
	File format (e.g., DICOM)
	Subject orientation
	Voxel size
	Decay correction performed
	Reconstruction parameters per protocol (method, iterations if OSEM, filtering, other)
Subject positioning and motion	Anatomic inclusion (complete cerebellum, top of head)
	Immersion in noise at edge of scanner axial field of view
	Interframe motion
	Motion between emission and transmission scan
	Adequate counts
Image processing	Consistent orientation (including left-right designation) for all scans
	Proper alignment between PET scans of different time points
	Proper alignment of MRI to PET scan (when MRI available for transformation to template space or transformation of VOIs to native scan)
	Goodness of fit between spatially normalized scan and template
Image analysis	Fit of VOIs on subject scan
	Consistency in boundary definition of VOIs and reference region across subjects
	If adjustment made for tissue truncation from scanner FOV, consistency in use of adjusted VOI or reference region for all scans within subject
	Consistency across SUVRs obtained using different reference regions
	Longitudinal stability of the SUVRs of regions not expected to accumulate amyloid (e.g., pons, subcortical white matter)

Abbreviations: PET, positron emission tomography; DICOM, Digital Imaging and Communication in Medicine; OSEM, ordered-subsets expectation maximization; MRI, magnetic resonance imaging; FOV, field of view.

scan acquisition parameters and should be checked for compliance with predefined acceptable values. The format most commonly used in clinical trials is Digital Imaging and Communication in Medicine (DICOM). The output data from some scanners are in a proprietary format that must be converted to DICOM, and these files may not include details from the original header file that must be obtained from the imaging site. Proper image orientation including left-right convention should be verified and corrected if necessary, as this will affect subsequent processing and measurement.

Visual inspection by an expert can reveal a number of possible issues [96]. Incomplete brain volumes due to positioning of the subject at the edge or out of the FOV can be identified. This most frequently occurs in the inferior slices and is found in approximately 4% of ADNI PiB scans. This can hamper reliable measurement, especially if the reference region used for SUVR calculation is cerebellar gray matter. The reconstructed emission images will preferably have been provided as subdivisions of the entire scan, commonly in 5-minute blocks or frames. Subject motion can be evaluated by comparing the head position from frame to frame, and the amount of translation and rotation can be calculated. Although the frames can be realigned, excessive movement may indicate that significant motion-related noise is embedded within the frame. Selected frames may be removed although resulting in a tradeoff between signal reduction and retained noise. Possible misalignment between the emission and transmission scan should also be checked by looking for asymmetry in the periphery and skull in the attenuation-corrected image. Unless data are available for re-reconstruction, this noise is embedded in the image and a likely reason for exclusion. Aligned frames with adequate subject stability are then summed or averaged into a single scan for further processing. To manage the volume of larger studies, automated software can be used to automate much of the image QC. The software can also be used to generate an audit trail as part of 21 CFR Part 11 compliance.

7. Image transformation factors

Registering or aligning images is necessary when more than one scan has been acquired from a subject and measurements need to be standardized across different sample times and when images from different modalities, such as PET and MRI, are used together. Spatial normalization, whereby image data from a subject are mapped to a reference brain space or atlas or a mathematic transformation, is calculated to map the reference atlas to each individual image, allows

the same sampling method such as ROIs to be used across subjects and across studies, and allows development of automated analysis pipelines. A number of tools have been developed for registration and mapping and will generate different results [97–99]. For example, an adaptive PET-based alignment approach has been shown to result in more consistent fitting of the cerebellum reference region than a version of SPM MR-based spatial normalization [99]. A comparison of different spatial normalization software packages showed varying degrees of fit or overlap with predefined anatomic regions, and this metric was used to rank tool performance [97]. Registration can fail or alignment errors occur when brain slices are missing or signal occurs outside the brain. In all cases, verification of alignment and proper fit of ROIs through visual inspection, which can be further aided through automated detection of out-of-range values, is critical, and method selection will impact SUVRs.

8. Measurement region selection and definition factors

8.1. Reference region selection

One of the most critical factors affecting amyloid measurement is the reference region. The normalization of all ROIs to this reference to obtain SUVRs means that variability in this region will impact all other measures for the scan. For example, a low reference region value due to artifact can result in a high SUVR value and falsely elevated estimate of amyloid burden. In longitudinal studies, where changes in amyloid burden may be modest, variability in the reference region signal measurement can generate significant shifts in the SUVR value. If unaddressed, the sensitivity of SUVR-based analyses may be reduced leading to erroneous conclusions regarding disease progression treatment effect.

Selection of the reference region can depend on a number of considerations including the tracer used, comparison with other studies, reliability in region delineation, and the subjects to be studied. The cerebellar cortex was selected for the first clinical studies with ^{11}C-PiB PET because of its reported lack of Congo red and thioflavin-S–stained plaques [100–102], presumed stability over time, and because clearance of the tracer from cerebellar gray is more similar to its clearance from the cerebral gray matter target regions than from white matter [52]. As noted previously, cerebellar gray ROIs are highly vulnerable to noise due to its low signal level, proximity to the edge of the scanner FOV where noise and truncation can occur, influence by scatter, and susceptibility to intrusion of higher white cerebellar signal or lower CSF signal into region boundaries if subject motion occurs. Cerebellar cortex may also accumulate plaques in late-stage disease [103] and in presenilin 1 mutation carriers [104]. Whole cerebellum has been used as a reference for florbetapir scans [1,61] and by including white matter may have higher signal intensity, less susceptibility to noise, and may include tissue less vulnerable to edge and truncation effects. Whole cerebellum is the preferred reference method for the Centiloid project (see the following section). However, the whole cerebellum can still be affected by the variability of cerebellar gray. The pons (and/or brainstem), which has high retention of tracer as a white matter structure but is devoid of amyloid deposits until very late-stage disease [37], has been used in analysis of flutemetamol scans [105]. Advantages of the pons include higher signal, less vulnerability to edge noise and truncation, and superior statistical power in some comparisons. The initial delivery of tracer to the pons measured in dynamic flutemetamol scans is very similar to cerebellar gray matter and much higher than the white matter region centrum semiovale, suggesting more comparable blood flow [105]. However, the pons can introduce variability because of its small volume and sensitivity to head motion. Subcortical white matter does not accumulate amyloid [31], and the centrum semiovale has been used as a reference region [44]. Such a region could have technical advantages by being located in the same plane as cortical target regions and therefore less susceptible to differences in scatter correction between superior and inferior planes. Nonetheless, delineation of the region has not been standardized and has not been commonly used. Furthermore, just as there can be spillover of signal from white matter into gray matter regions, cortical signal can spillover into adjacent white matter, confounding precise assignment of the measured signal to a particular tissue type. Combinations of reference regions including pons, whole cerebellum, and subcortical white matter have been proposed to offset individual region variability although formal comparisons in performance of these different reference regions in longitudinal studies have not yet been published.

8.2. Regions of interest

Selecting the regions of the brain for analysis is critical and depends on the hypotheses being testing. There are many methods for defining brain regions, including both manual and automated methods. These are discussed in further detail in the Supplementary Material. Studies have typically evaluated amyloid burden using cortical frontal, lateral temporal, lateral parietal, anterior cingulate, and posterior cingulate/precuneus regions, based on findings that these regions show the greatest tracer retention in AD patients relative to normal controls [106,107] and consistent with pathologic studies of amyloid accumulation [37,101]. A cortical average SUVR is often calculated using these regions for a simplified measure. Occipital cortex is sometimes included [108] but can add variance because of heterogeneity in amyloid accumulation [106] and possible association with CAA [28]. As clinical trials move into earlier disease stages, specific regions may provide a more sensitive measure than a cortical average [109,110].

Additional variability comes from the amount of a given anatomic region included for measurement (entire structure vs. most amyloid vulnerable subregions) and methods to reduce the influence of atrophy and white matter spillover. One approach is to erode the boundaries of each ROI from neighboring tissue [111], which can be done in combination with individualized gray segment masking. This may be a less variable though approximate approach than PVE correction, which is highly method dependent [82].

9. Standardization factors

9.1. Amyloid threshold selection

The definition of amyloid positivity as a single cutoff value (such as 1.5 for ^{11}C-PiB) has been appealing to characterize diagnostic groups [112] and provide a basis for individual enrollment in a clinical trial or analysis. However, the relevance of a cutoff depends on the tracer, image acquisition and analysis methods, population discrimination approach, and clinical objective. Thresholds have been proposed based on ROC separation between diagnostic groups [113], comparison with autopsy findings [1], and use of healthy young adults to define an amyloid-negative standard [10].

9.2. Reconciliation across measurement methods

Differences in tracer binding properties, choice of uptake units, and measurement methodologies result in different metrics and thresholds for quantification of tracer retention. The desire to be able to compare amyloid retention across studies and tracers has created a need for standardization of tracer retention in the brain. It was proposed at the 2012 Alzheimer's Association International Conference that quantitative amyloid imaging metrics could be standardized by fitting the outcome of each tracer/analysis method to a scale with units from zero to 100 called "centiloids." The Centiloid Project working group will establish a PiB PET reference scale using standard analysis of PiB scans performed on two groups to define the scale anchor points: a group of healthy control subjects aged <45 years (who should have no amyloid) as the lower anchor (zero point) for the scale and typical AD patients whose mean SUVR value will define the upper anchor (100 points) for the scale. The second phase of the project outlines a method for subsequently scaling any alternate method of PiB analysis or any other tracer and associated analysis method to this scale. To allow research laboratories to map their data and methods to the reference Centiloid scale, a set of PET and MRI scans along with details of the analysis methods used by the working group will be made available on a publically accessible server [114].

10. Discussion

Amyloid PET has the potential to significantly assist the development of disease-modifying treatments for AD. Quantitative measurement of amyloid PET, if sufficiently sensitive and reliable, can be used to confirm that treatments targeting fibrillar Aβ are able to get into the brain, are able to fulfill their mechanism of action, and can help define clinically relevant and safe doses. Tremendous strides in developing quantitative measurement of amyloid PET have been made by ADNI and related research consortia, imaging companies, and academic laboratories. The first standardized protocols, guidance on image QC, and analysis methods owe their existence to these efforts. Nonetheless, for quantitative amyloid PET to be robust and reliable, these have to be regarded as foundational and not final steps. The ongoing testing of methods for CSF biomarkers and volumetric MRI illustrate the work necessary to reduce variability due to measurement error and noise and support their qualification by health authorities as biomarkers of disease progression [115-117]. Similar efforts for quantitative PET are needed.

As we have outlined, numerous factors impact the values derived from amyloid PET and can be a source of measurement error. We have summarized those we regard as key factors, how they can impact data quality, and mitigation strategies in a table in the Supplementary Material and highlight them here.

Several biological factors influence the incidence and interpretation of the amyloid PET signal although most are primarily relevant for study design and subject selection/stratification. Impact on longitudinal change in amyloid PET signal once factors such as age, genetic background, comorbid conditions have been controlled is minimal, although effects on tracer distribution by drug-drug interaction and blood flow remain. Of these, blood flow is potentially the more variable and is best addressed through how the data are acquired; namely, the collection of dynamic versus late emission "static" scans. The lengthy time required for acquiring dynamic data renders this impractical for multisite phase 2 and 3 trials; however, dynamic scans could be considered for focused, proof-of-principle phase 2 studies conducted at a limited number of research sites.

The most critical family of factors is acquisition. For reliable quantitative amyloid PET data, good-quality acquisition is essential and this begins with how sites are selected and qualified and should include phantom testing and critical review of phantom data. Once qualified, sites need to demonstrate accurate and sensitive scanner performance and provide evidence of regular equipment QC to ensure data consistency. Examples of this include the standardization for FDG PET by QIBA and the site accreditation program developed for FDG PET/CT in cancer trials by the EANM Research Ltd [4,90]. The latter program includes a mechanism for site accreditation, standards for phantom testing, and direction for quarterly calibration QC testing and annual image QC testing, and accreditation is maintained by continuing to meet specifications. A critical lesson of standardization efforts by QIBA and the EANM is that defining best practices for quantitative imaging is a

precompetitive effort that allows engagement of imaging service providers, academic experts in imaging, and customers. As we have learned in ADNI, such efforts are beyond the reach of a single company and benefit from the collaboration of partners. Proprietary solutions may not be in anyone's interest, whereas shared investment may provide funds for methods work and testing that might otherwise not be done. Thus, next steps for optimizing quantitative amyloid PET could include outlining the scope, governance, and processes for reviewing topics such as site qualification, QC, and acquisition protocols to support standardizing.

Subject positioning in the scanner and subject movement are major and common sources of error. Some of the methods for head restraint may work well; others may exacerbate movement because of discomfort or excessive restraint. Determining the optimal approach could be easily tested by having subjects lay in the scanner with the different restraint methods and tracking movements under each method with a video camera. In the event that movement does occur, having software for registering emission frames with the transmission scan before reconstruction could reduce error due to embedded transmission/emission misregistration. Because transmission/emission motion correction needs to be carried out in projection space before reconstruction and because reconstruction implementations are not openly available and vary widely between scanners, this most likely needs to be a vendor-driven fix.

For image analysis, fundamental questions remain around the reference region. The optimal reference region may differ between cross-sectional studies, where differences between subjects are the focus, and longitudinal studies, where subtle differences within subjects are the end point. Cerebellar gray matter performs well for distinguishing between healthy subjects and those with AD or at risk and may be an acceptable reference tissue for many studies. However, the vulnerability of this region to accumulation of amyloid in mutations carriers and late-stage AD, its position close to the edge of the FOV, intrinsically low signal and errors in scatter correction, and potential errors in segmentation and coregistration may not make this the best region for longitudinal studies. Systematic comparison and analysis of within-subject variability using different reference regions such as whole cerebellum, pons, and subcortical white matter are needed. Ideally, these analyses should include at least three time points to allow detection of implausible trajectories (e.g., decreases followed by increases in excess of test-retest variability) and determining whether an alternative reference region reduces within-subject variability. PVE correction is another question that bedevils image analysis in AD and could also benefit from more systematic comparisons in longitudinal data sets. To resolve these questions, publically available image data sets such as ADNI are essential. Such data sets allow head-to-head comparison of methods and use of healthy control data as reference. This is planned to be a feature of the Centiloid project and is an argument for making image data from intervention trials available, such as amyloid PET scans acquired on placebo-treated subjects.

With the movement in intervention trials to earlier stages of disease, such as the Anti-Amyloid treatment in Asymptomatic Alzheimer's (A4) study (http://a4study.org/), it is time to revisit threshold values for amyloid burden that were established for discriminating between AD and healthy subjects. Lower or region-specific cutoffs may be appropriate for studies designed to enroll subjects very early in disease course, determine whether there is sufficient amyloid to detect removal, or to evaluate "emerging" subjects who may initially have one amyloid-positive region rather than a positive cortical average SUVR. Questions such as these can benefit from histology data from autopsies of subjects with antemortem amyloid PET scans as "ground truth," and we encourage companies and research groups to include autopsy protocols and consents in trials whenever possible. It is also important to ensure industry awareness that SUVR values used to determine trial inclusion or to support diagnosis, depends on the tracer, acquisition protocol, and processing methods that were used to generate the value.

Finally, acquisition and reconstruction protocols, missing or censored data including QC thresholds for exclusion, image analysis software, scanner or scanner software changes over the course of the study, and statistical analysis plans should be fully reported to allow for reproducing the results of a study, consonant with the emerging publishing standards for reproducible research.

These recommendations are for the immediate and near future of quantitative amyloid imaging in support of research in AD and related dementias. Future methods will likely include ligands for tau, inflammation markers, and functional measures, and each of these methods will confront the same issues regarding variability of data collected in multisite studies over time. Developing a robust approach to quantitative measurement of tau PET may be especially important, as autopsy data suggest that the signal distribution may evince more regionally distinct patterns than amyloid PET [118]. It is our hope that developing best practices for quantitative amyloid PET may provide a pathway for optimizing these newer methods as they too are adopted as clinical research tools.

Acknowledgments

Data collection and sharing for some of the analyses in this manuscript was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai

Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jalz.2014.09.004.

References

[1] Clark CM, Schneider JA, Bedell BJ, Beach TG, Bilker WB, Mintun MA, et al. Use of florbetapir-PET for imaging b-amyloid pathology. JAMA 2011;305:275–83.

[2] Available from: http://www.amyvid.com/Pages/reader-training-program.aspx.

[3] (QIBA), Q.I.B.A., QIBA Profile. FDG-PET/CT as an imaging biomarker. 3. Measuring response to cancer therapy. 2013.

[4] Boellaard R, O'Doherty M, Weber W, Mottaghy F, Lonsdale M, Stroobants S, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. Eur J Nucl Med Mol Imaging 2010;37:181–200.

[5] Jagust WJ, Bandy D, Chen K, Foster NL, Landau SM, Mathis CA, et al. The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core. Alzheimers Dementia 2010;6:221–9.

[6] Tolboom N, Yaqub M, Boellaard R, Luurtsema G, Windhorst A, Scheltens P, et al. Test-retest variability of quantitative [11C]PIB studies in Alzheimer's disease. Eur J Nucl Med Mol Imaging 2009;36:1629–38.

[7] Scheinin NM, Aalto S, Koikkalainen J, Lotjonen J, Karrasch M, Kemppainen N, et al. Follow-up of [11C]PIB uptake and brain volume in patients with Alzheimer disease and controls. Neurology 2009;73:1186–92.

[8] Braak H, Thal DR, Ghebremedhin E, Del Tredici K. Stages of the pathologic process in Alzheimer disease: age categories from 1 to 100 years. J Neuropathol Exp Neurol 2011;70:960–9.

[9] Vlassenko AG, Mintun MA, Xiong C, Sheline YI, Goate AM, Benzinger TL, et al. Amyloid-beta plaque growth in cognitively normal adults: longitudinal PIB data. Ann Neurol 2011;70:857–61.

[10] Fleisher AS, Chen K, Liu X, Ayutyanont N, Roontiva A, Thiyyagura P, et al. Apolipoprotein E epsilon4 and age effects on florbetapir positron emission tomography in healthy aging and Alzheimer disease. Neurobiol Aging 2013;34:1–12.

[11] Jack CR Jr, Wiste HJ, Weigand SD, Knopman DS, Lowe V, Vemuri P, et al. Amyloid-first and neurodegeneration-first profiles characterize incident amyloid PET positivity. Neurology 2013;81:1732–40.

[12] Bateman RJ, Xiong C, Benzinger TLS, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med 2012;367:795–804.

[13] Handen BL, Cohen AD, Channamalappa U, Bulova P, Cannon SA, Cohen WI, et al. Imaging brain amyloid in nondemented young adults with Down syndrome using Pittsburgh compound B. Alzheimers Dement 2012;8:496–501.

[14] Karran E, Mercken M, Strooper BD. The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. Nat Rev Drug Discov 2011;10:698–712.

[15] Villemagne VL, Ataka S, Mizuno T, Brooks WS, Wada Y, Kondo M, et al. High striatal amyloid {beta}-peptide deposition across different autosomal Alzheimer disease mutation types. Arch Neurol 2009;66:1537–44.

[16] Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 1993;261:921–3.

[17] Meyer MR, Tschanz JT, Norton MC, Welsh-Bohmer KA, Steffens DC, Wyse BW, et al. APOE genotype predicts when—not whether—one is predisposed to develop Alzheimer disease. Nat Genet 1998;19:321–2.

[18] Barthel H, Gertz HJ, Dresel S, Peters O, Bartenstein P, Buerger K, et al. Cerebral amyloid-[beta] PET with florbetaben (18F) in patients with Alzheimer's disease and healthy controls: a multicentre phase 2 diagnostic study. Lancet Neurol 2011;10:424–35.

[19] Reiman EM, Chen K, Liu X, Bandy D, Yu M, Lee W, et al. Fibrillar amyloid-ß burden in cognitively normal people at 3 levels of genetic risk for Alzheimer's disease. PNAS 2009;106:6820–5.

[20] Fleisher AS, Chen K, Liu X, Roontiva A, Thiyyagura P, Ayutyanont N, et al. Using positron emission tomography and florbetapir F 18 to image cortical amyloid in patients with mild cognitive impairment or dementia due to Alzheimer disease. Arch Neurol 2011;68:1404–11.

[21] Grimmer T, Tholen S, Yousefi BH, Alexopoulos P, Förschler A, Förstl H, et al. Progression of cerebral amyloid load is associated with the apolipoprotein E [epsilon]4 genotype in Alzheimer's disease. Biol Psychiatry 2010;68:879–84.

[22] Allan CL, Ebmeier KP. The influence of ApoE4 on clinical progression of dementia: a meta-analysis. Int J Geriatr Psychiatry 2011;26:520–6.

[23] Ghebremedhin E, Schultz C, Thal DR, Rüb U, Ohm TG, Braak E, et al. Gender and age modify the association between APOE and AD-related neuropathology. Neurology 2001;56:1696–701.

[24] Gorelick PB, Scuteri A, Black SE, Decarli C, Greenberg SM, Iadecola C, et al. Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2011;42:2672–713.

[25] Toledo JB, Arnold SE, Raible K, Brettschneider J, Xie SX, Grossman M, et al. Contribution of cerebrovascular disease in autopsy confirmed neurodegenerative disease cases in the National Alzheimer's Coordinating Centre. Brain 2013;136(Pt 9):2697–706.

[26] Rodrigue KM, Rieck JR, Kennedy KM, Devous MD, Sr, Diaz-Arrastia R, et al. Risk factors for β-amyloid deposition in healthy aging: vascular and genetic effects. JAMA Neurol 2013;70:600–6.

[27] Hughes TM, Kuller LH, Barinas-Mitchell EM. Arterial stiffness and β-amyloid progression in nondemented elderly adults. JAMA Neurol 2014;71:562–8.

[28] Johnson KA, Gregas M, Becker JA, Kinnecom C, Salat DH, Moran EK, et al. Imaging of amyloid burden and distribution in cerebral amyloid angiopathy. Ann Neurol 2007;62:229–34.

[29] Sperling RA, Jack CR, Black SE, Frosch MP, Greenberg SM, Hyman BT, et al. Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: recommendations from the Alzheimer's Association Research Roundtable Workgroup. Alzheimers Dementia 2011;7:367–85.

[30] Ly JV, Donnan GA, Villemagne VL, Zavala JA, Ma H, O'Keefe G, et al. 11C-PIB binding is increased in patients with cerebral amyloid angiopathy-related hemorrhage. Neurology 2010;74:487–93.
[31] Ikonomovic MD, Klunk WE, Abrahamson EE, Mathis CA, Price JC, Tsopelas ND, et al. Post-mortem correlates of in vivo PiB-PET amyloid imaging in a typical case of Alzheimer's disease. Brain 2008; 131:1630–45.
[32] Clark CM, Pontecorvo MJ, Beach TG, Bedell BJ, Coleman RE, Doraiswamy PM, et al. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-beta plaques: a prospective cohort study. Lancet Neurol 2012; 11:669–78.
[33] Cairns NJ, Ikonomovic MD, Benzinger T, Storandt M, Fagan AM, Shah AR, et al. Absence of Pittsburgh compound B detection of cerebral amyloid {beta} in a patient with clinical, cognitive, and cerebrospinal fluid markers of Alzheimer disease: a case report. Arch Neurol 2009;66:1557–62.
[34] Sojkova J, Driscoll I, Iacono D, Zhou Y, Codispoti KE, Kraut MA, et al. In vivo fibrillar beta-amyloid detected using [11C]PiB positron emission tomography and neuropathologic assessment in older adults. Arch Neurol 2011;68:232–40.
[35] Klunk WE, Price JC, Mathis CA, Tsopelas ND, Lopresti BJ, Ziolko SK, et al. Amyloid deposition begins in the striatum of presenilin-1 mutation carriers from two unrelated pedigrees. J Neurosci 2007;27:6174–84.
[36] Villemagne VL, Ataka S, Mizuno T, Brooks WS, Wada Y, Kondo M, et al. High striatal amyloid beta-peptide deposition across different autosomal Alzheimer disease mutation types. Arch Neurol 2009; 66:1537–44.
[37] Thal DR, Rub U, Orantes M, Braak H. Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology 2002;58:1791–800.
[38] Stern Y. Cognitive reserve in ageing and Alzheimer's disease. Lancet Neurol 2012;11:1006–12.
[39] Morbelli S, Perneczky R, Drzezga A, Frisoni GB, Caroli A, van Berckel BNM, et al. Metabolic networks underlying cognitive reserve in prodromal Alzheimer disease: a European Alzheimer disease consortium project. J Nucl Med 2013;54:894–902.
[40] Landau SM, Marks SM, Mormino EC, Rabinovici GD, Oh H, O'Neil JP, et al. Association of lifetime cognitive engagement and low beta-amyloid deposition. Arch Neurol 2012;69:623–9.
[41] Groenning M. Binding mode of Thioflavin T and other molecular probes in the context of amyloid fibrils—current status. J Chem Biol 2010;3:1–18.
[42] Rinne JO, Brooks DJ, Rossor MN, Fox NC, Bullock R, Klunk WE, et al. 11C-PiB PET assessment of change in fibrillar amyloid-ß load in patients with Alzheimer's disease treated with bapineuzumab: a phase 2, double-blind, placebo-controlled, ascending-dose study. Lancet Neurol 2010;9:363–72.
[43] Ryu YH, Liow JS, Zoghbi S, Fujita M, Collins J, Tipre D, et al. Disulfiram inhibits defluorination of (18)F-FCWAY, reduces bone radioactivity, and enhances visualization of radioligand binding to serotonin 5-HT1A receptors in human brain. J Nucl Med 2007;48:1154–61.
[44] Wong DF, Rosenberg PB, Zhou Y, Kumar A, Raymont V, Ravert HT, et al. In vivo imaging of amyloid deposition in Alzheimer disease using the radioligand 18F-AV-45 (florbetapir [corrected] F 18). J Nucl Med 2010;51:913–20.
[45] Van Vlaslaer A, Mortishire-Smith RJ, Mackie C, Langlois X, Schmidt ME. Profiling of hepatic clearance pathways of Pittsburgh compound B and human liver cytochrome p450 phenotyping. EJNMMI Res 2013;3:10.
[46] Herzog H, Seitz RJ, Tellmann L, Rota Kops E, Julicher F, Schlaug G, et al. Quantitation of regional cerebral blood flow with 15O-butanol and positron emission tomography in humans. J Cereb Blood Flow Metab 1996;16:645–9.
[47] Alegret M, Cuberas-Borrós G, Vinyes-Junqué G, Espinosa A, Valero S, Hernández I, et al. A two-year follow-up of cognitive deficits and brain perfusion in mild cognitive impairment and mild Alzheimer's disease. J Alzheimers Dis 2012;30:109–20.
[48] Saha GB. Basics of PET imaging: physics, chemistry and regulations. S.S.B. Media, Editor. Springer; 2005.
[49] Morris ED, Endres CJ, Schmidt KC, Christian BT, Muzic RF, Fisher RE. Kinetic modeling in positron emission tomography. In: Wernick MN, Aarsvold JN, eds. Emission tomography: the fundamentals of PET and SPECT. Academic Press; 2004. p. 499–540.
[50] Gjedde A, Aanerud J, Braendgaard H, Rodell AB. Blood-brain transfer of Pittsburgh compound B in humans. Front Aging Neurosci 2013;5:70.
[51] Blomquist G, Engler H, Nordberg A, Ringheim A, Wall A, Forsberg A, et al. Unidirectional influx and net accumulation of PIB. Open Neuroimag J 2008;2:114–25.
[52] Price JC, Klunk WE, Lopresti BJ, Lu X, Hoge JA, Ziolko SK, et al. Kinetic modeling of amyloid binding in humans using PET imaging and Pittsburgh compound-B. J Cereb Blood Flow Metab 2005; 25:1528–47.
[53] Nelissen N, Van Laere K, 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. J Nucl Med 2009;50:1251–9.
[54] Yaqub M, Tolboom N, Boellaard R, van Berckel BNM, van Tilburg EW, Luurtsema G, et al. Simplified parametric methods for [11C]PIB studies. NeuroImage 2008;42:76–86.
[55] McNamee RL, Yee SH, Price JC, Klunk WE, Rosario B, Weissfeld L, et al. Consideration of optimal time window for Pittsburgh compound B PET summed uptake measurements. J Nucl Med 2009;50:348–55.
[56] Villemagne VL, Pike KE, Chételat G, Ellis KA, Mulligan RS, Bourgeat P, et al. Longitudinal assessment of Abeta and cognition in aging and Alzheimer disease. Ann Neurol 2011;69:181–92.
[57] Carson RE, Channing MA, Blasberg RG, Dunn BB, Cohen RM, Rice KC, et al. Comparison of bolus and infusion methods for receptor quantitation: application to [18F]cyclofoxy and positron emission tomography. J Cereb Blood Flow Metab 1993;13:24–42.
[58] van Berckel BNM, Ossenkoppele R, Tolboom N, Yaqub M, Foster-Dingley JC, Windhorst AD, et al. Longitudinal amyloid imaging using 11C-PiB: methodologic considerations. J Nucl Med 2013; 54:1570–6.
[59] Zhou Y, Sojkova J, Resnick SM, Wong DF. Relative equilibrium plot improves graphical analysis and allows bias correction of standardized uptake value ratio in quantitative 11C-PiB PET studies. J Nucl Med 2012;53:622–8.
[60] Villemagne VL, Ong K, Mulligan RS, Holl G, Pejoska S, Jones G, et al. Amyloid imaging with 18F-florbetaben in Alzheimer disease and other dementias. J Nucl Med 2011;52:1210–7.
[61] Joshi AD, Pontecorvo MJ, Clark CM, Carpenter AP, Jennings DL, Sadowsky CH, et al. Performance characteristics of amyloid PET with florbetapir F 18 in patients with Alzheimer's disease and cognitively normal subjects. J Nucl Med 2012;53:378–84.
[62] Vandenberghe R, Van Laere K, Ivanoiu A, Salmon E, Bastin C, Triau E, et al. 18F-flutemetamol amyloid imaging in Alzheimer disease and mild cognitive impairment: a phase 2 trial. Ann Neurol 2010;68:319–29.
[63] Cselényi Z, Jönhagen ME, Forsberg A, Halldin C, Julin P, Schou M, et al. Clinical validation of 18F-AZD4694, an amyloid-β–specific PET radioligand. J Nucl Med 2012;53:415–24.
[64] Rowe CC, Pejoska S, Mulligan RS, Jones G, Chan JG, Svensson S, et al. Head-to-head comparison of 11C-PiB and 18F-AZD4694 (NAV4694) for β-amyloid imaging in aging and dementia. J Nucl Med 2013;54:880–6.
[65] Rinne J, Wong D, Wolk D, Leinonen V, Arnold S, Buckley C, et al. [18F]Flutemetamol PET imaging and cortical biopsy histopathology for fibrillar amyloid β detection in living subjects with normal

pressure hydrocephalus: pooled analysis of four studies. Acta Neuropathologica 2012;124:833–45.
[66] Stankoff B, Freeman L, Aigrot MS, Chardain A, Dolle F, Williams A, et al. Imaging central nervous system myelin by positron emission tomography in multiple sclerosis using [methyl-(1)(1)C]-2-(4'-methylaminophenyl)-6-hydroxybenzothiazole. Ann Neurol 2011; 69:673–80.
[67] Landau SM, Breault C, Joshi AD, Pontecorvo M, Mathis CA, Jagust WJ, et al. Amyloid-β imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. J Nucl Med 2013;54:70–7.
[68] Landau SM, Thomas BA, Thurfjell L, Schmidt M, Margolin R, Mintun M, et al. Amyloid PET imaging in Alzheimer's disease: a comparison of three radiotracers. Eur J Nucl Med Mol Imaging 2014;41:1398–407.
[69] U.S. Department of Health and Human Services Food and Drug Administration, C.f.D.E.a.R.C. PET drugs—current good manufacturing practice (CGMP). 2009.
[70] USP Chapter <823> "Radiopharmaceuticals for Positron Emission Tomography— Compounding," (USP 32/NF 27) 2009.
[71] Elsinga P, Todde S, Penuelas I, Meyer G, Farstad B, Faivre-Chauvet A, et al. Guidance on current good radiopharmacy practice (cGRPP) for the small-scale preparation of radiopharmaceuticals. Eur J Nucl Med Mol Imaging 2010;37:1049–62.
[72] Verbruggen A, Coenen H, Deverre J-R, Guilloteau D, Langstrom B, Salvadori P, et al. Guideline to regulations for radiopharmaceuticals in early phase clinical trials in the EU. Eur J Nucl Med Mol Imaging 2008;35:2144–51.
[73] Andersson JL, Vagnhammar BE, Schneider H. Accurate attenuation correction despite movement during PET imaging. J Nucl Med 1995; 36:670–8.
[74] Ikari Y, Nishio T, Makishi Y, Miya Y, Ito K, Koeppe RA, et al. Head motion evaluation and correction for PET scans with 18F-FDG in the Japanese Alzheimer's Disease Neuroimaging Initiative (J-ADNI) multi-center study. Ann Nucl Med 2012;26:535–44.
[75] Boellaard R, Oyen WJG, Hoekstra CJ, Hoekstra OS, Visser EP, Willemsen AT, et al. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. Eur J Nucl Med Mol Imaging 2008;35:2320–33.
[76] Kinahan PE, Fletcher JW. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. Semin Ultrasound CT MR 2010; 31:496–505.
[77] Klein G, Scott D, Sharoyon V, Oh J, Koeppe RA, Suhy J. Scanner resolution effects on quantitative amyloid measurements, in Human amyloid imaging. Presented at the 7th Human Amyloid Imaging. January 16-18, 2013: Miami Beach, FL.
[78] Joshi A, Koeppe RA, Fessler JA. Reducing between scanner differences in multi-center PET studies. NeuroImage 2009;46:154–9.
[79] Koeppe, R.A. ADNI PET pre-processing. 2013. Available at: http://adni.loni.ucla.edu/methods/pet-analysis/pre-processing/. Accessed May 31, 2013.
[80] Erlandsson K, Buvat I, Pretorius PH, Thomas BA, Hutton BF. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. Phys Med Biol 2012;57:R119–59.
[81] Meltzer CC, Kinahan PE, Greer PJ, Nichols TE, Comtat C, Cantwell MN, et al. Comparative evaluation of MR-based partial-volume correction schemes for PET. J Nucl Med 1999; 40:2053–65.
[82] Thomas BA, Erlandsson K, Modat M, Thurfjell L, Vandenberghe R, Ourselin S, et al. The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. Eur J Nucl Med Mol Imaging 2011;38:1104–19.
[83] Raniga P, Bourgeat P, Fripp J, Acosta O, Ourselin S, Rowe C, et al. Alzheimer's disease detection using 11C-PiB with improved partial volume effect correction. *Med Imaging, Proc SPIE* 2009 [cited 7262].
[84] Quarantelli M, Berkouk K, Prinster A, Landeau B, Svarer C, Balkay L, et al. Integrated software for the analysis of brain PET/SPECT studies with partial-volume-effect correction. J Nucl Med 2004;45:192–201.
[85] Lowe VJ, Kemp BJ, Jack CR, Senjem M, Weigand S, Shiung M, et al. Comparison of 18F-FDG and PiB PET in cognitive impairment. J Nucl Med 2009;50:878–86.
[86] Nakamoto Y, Osman M, Cohade C, Marshall LT, Links JM, Kohlmyer S, et al. PET/CT: comparison of quantitative tracer uptake between germanium and CT transmission attenuation-corrected images. J Nucl Med 2002;43:1137–43.
[87] Fahey FH, Palmer MR, Strauss KJ, Zimmerman RE, Badawi RD, Treves ST. Dosimetry and adequacy of CT-based attenuation correction for pediatric PET: phantom study 1. Radiology 2007; 243:96–104.
[88] Ting X, Adam MA, Bruno De M, Ravindra M, Evren A, Paul EK. Ultra-low dose CT attenuation correction for PET/CT. Phys Med Biol 2012;57:309.
[89] Chang LT. A method for attenuation correction in radionuclide computed tomography. IEEE Trans Nucl Sci 1978;25:638–43.
[90] Varrone A, Asenbaum S, Vander Borght T, Booij J, Nobili F, Nagren K, et al. EANM procedure guidelines for PET brain imaging using [18F]FDG, version 2. Eur J Nucl Med Mol Imaging 2009; 36:2103–10.
[91] Malone IB, Ansorge RE, Williams GB, Nestor PJ, Carpenter TA, Fryer TD. Attenuation correction methods suitable for brain imaging with a PET/MRI scanner: a comparison of tissue atlas and template attenuation map approaches. J Nucl Med 2011;52:1142–9.
[92] Koeppe, R.A. Siemens HRRT PET scan warnings. 2007; Available from: http://adni.loni.ucla.edu/siemens-hrrt-pet-scan-warnings/.
[93] Keller S, Svarer C, Sibomana M. Attenuation correction for the HRRT PET-scanner using transmission scatter correction and total variation regularization. IEEE Trans Med Imaging 2013;32:1611–21.
[94] Doot RK, Scheuermann JS, Christian PE, Karp JS, Kinahan PE. Instrumentation factors affecting variance and bias of quantifying tracer uptake with PET/CT. Med Phys 2010;37:6035–46.
[95] Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. J Nucl Med 2004;45:1519–27.
[96] Andersson JLR. A rapid and accurate method to realign PET scans utilizing image edge information. J Nucl Med 1995;36:657–69.
[97] Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 2009; 46:786–802.
[98] Holland D, Dale AM, Alzheimer's Disease Neuroimaging Initiative. Nonlinear registration of longitudinal images and measurement of change in regions of interest. Med Image Anal 2011;15:489–97.
[99] Lundqvist R, Lilja J, Thomas BA, Lotjonen J, Villemagne VL, Rowe CC, et al. Implementation and validation of an adaptive template registration method for 18F-flutemetamol imaging data. J Nucl Med 2013;54:1472–8.
[100] Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh compound-B. Ann Neurol 2004;55:306–19.
[101] Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathologica 1991;82:239–59.
[102] Lopresti BJ, Klunk WE, Mathis CA, Hoge JA, Ziolko SK, Lu X, et al. Simplified quantification of Pittsburgh compound B amyloid imaging PET studies: a comparative analysis. J Nucl Med 2005;46:1959–72.
[103] Joachim CL, Morris JH, Selkoe DJ. Diffuse senile plaques occur commonly in the cerebellum in Alzheimer's disease. Am J Pathol 1989;135:309–19.
[104] Knight WD, Okello AA, Ryan NS, Turkheimer FE, Rodriguez Martinez de Llano S, Edison P, et al. Carbon-11-Pittsburgh compound B positron emission tomography imaging of amyloid deposition in presenilin 1 mutation carriers. Brain 2011;134(Pt 1):293–300.

[105] Edison, P, Hinz R, Ramlackhansingh A, Thomas J, Turkheimer FE, Brooks DJ. Can we use pons as a reference region for the analysis of [11C]PIB PET? Presented at the Human Amyloid Imaging Conference. April 9, 2010; Toronto, CA.
[106] Ziolko SK, Weissfeld LA, Klunk WE, Mathis CA, Hoge JA, Lopresti BJ, et al. Evaluation of voxel-based methods for the statistical analysis of PIB PET amyloid imaging studies in Alzheimer's disease. NeuroImage 2006;33:94–102.
[107] Engler H, Forsberg A, Almkvist O, Blomquist G, Larsson E, Savitcheva I, et al. Two-year follow-up of amyloid deposition in patients with Alzheimer's disease. Brain 2006;129:2856–66.
[108] Rowe CC, Ng S, Ackermann U, Gong SJ, Pike K, Savage G, et al. Imaging {beta}-amyloid burden in aging and dementia. Neurology 2007;68:1718–25.
[109] Rodrigue KM, Kennedy KM, Devous MD Sr, Rieck JR, Hebrank AC, Diaz-Arrastia R, et al. Beta-amyloid burden in healthy aging: regional distribution and cognitive consequences. Neurology 2012; 78:387–95.
[110] Yotter RA, Doshi J, Clark V, Sojkova J, Zhou Y, Wong DF, et al. Memory decline shows stronger associations with estimated spatial patterns of amyloid deposition progression than total amyloid burden. Neurobiol Aging 2013;34:2835–42.
[111] Li Y, Rinne JO, Mosconi L, Pirraglia E, Rusinek H, DeSanti S, et al. Regional analysis of FDG and PIB-PET images in normal aging, mild cognitive impairment, and Alzheimer's disease. Eur J Nucl Med Mol Imaging 2008;35:2169–81.
[112] Jack CR Jr, Lowe VJ, Senjem ML, Weigand SD, Kemp BJ, Shiung MM, et al. 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. Brain 2008;131:665–80.
[113] Mormino EC, Kluth JT, Madison CM, Rabinovici GD, Baker SL, Miller BL, et al. Episodic memory loss is related to hippocampal-mediated beta-amyloid deposition in elderly subjects. Brain 2009; 132(Pt 5):1310–23.
[114] Rowe, C. The Centiloid Scale: Standardization of amyloid imaging measures, in Alzheimer's Imaging Consortium - Alzheimer's Association International Conference. 2013: Boston, USA.
[115] Mattsson N, Zetterberg H, Blennow K. Lessons from multicenter studies on CSF biomarkers for Alzheimer's disease. Int J Alzheimers Dis 2010;2010.
[116] Mattsson N, Blennow K, Zetterberg H. Inter-laboratory variation in cerebrospinal fluid biomarkers for Alzheimer's disease: united we stand, divided we fall. Clin Chem Lab Med 2010;48:603–7.
[117] Jack CR Jr, Barkhof F, Bernstein MA, Cantillon M, Cole PE, Decarli C, et al. Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. Alzheimers Dement 2011; 7:474–4854.
[118] Braak H, Alafuzoff I, Arzberger T, Kretzschmar H, Del Tredici K. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. Acta Neuropathol 2006;112:389–404.