Review Article

Amyloid PET in clinical practice: Its place in the multidimensional space of Alzheimer's disease☆

Rik Vandenberghe a,b,c,⁎, Katarzyna Adamczuk a,c, Patrick Dupont a,c, Koen Van Laere d, Gaël Chételat e,f,g,h

a Laboratory for Cognitive Neurology, Department of Neurosciences, University of Leuven, Belgium
b Neurology Department, University Hospitals Leuven, Belgium
c Alzheimer Research Centre KU Leuven, Leuven Institute of Neuroscience and Disease, University of Leuven, Belgium
d Nuclear Medicine Department, University Hospitals Leuven, Belgium
e INSERM, U1077, Caen, France
f Université de Caen Basse-Normandie, UMR-S1077, Caen, France
g Ecole Pratique des Hautes Etudes, UMR-S1077, Caen, France
h CHU de Caen, U1077, Caen, France

Abstract

Amyloid imaging is currently introduced to the market for clinical use. We will review the evidence demonstrating that the different amyloid PET ligands that are currently available are valid biomarkers for Alzheimer-related β amyloidosis. Based on recent findings from cross-sectional and longitudinal imaging studies using different modalities, we will incorporate amyloid imaging into a multidimensional model of Alzheimer's disease. Aside from the critical role in improving clinical trial design for amyloid-lowering drugs, we will also propose a tentative algorithm for when it may be useful in a memory clinic environment. Gaps in our evidence-based knowledge of the added value of amyloid imaging in a clinical context will be identified and will need to be addressed by dedicated studies of clinical utility.

© 2013 The Authors. Published by Elsevier Inc. All rights reserved.

Contents

1. Introduction .............................................................................. 498
2. Some basic AD principles relevant to amyloid imaging ................. 498
3. Comparisons between amyloid ligands ........................................ 499
4. Comparisons between amyloid ligands and neuropathology ......... 499
   4.1. Large-scale prospective postmortem validation studies .......... 499
   4.2. Other clinicopathological studies ........................................ 500
      4.2.1. False-positive cases with CERAD and NIA-RI criteria as comparator .................. 500
      4.2.2. False-negative cases with CERAD and NIA-RI criteria as comparator .............. 500
   4.3. Affinity for white matter lesions .......................................... 500
   4.4. Summary ......................................................................... 501
5. The AD puzzle: putting the pieces together by means of multimodal imaging .................................................... 501
   5.1. Topography .................................................................. 501
   5.2. Relationship with disease severity ................................. 501
   5.3. Longitudinal change ..................................................... 501
   5.4. Comparison of diagnostic performance ............................. 502
   5.5. Task-related fMRI versus amyloid PET ......................... 502
   5.6. Summary ..................................................................... 503

☆ This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

⁎ Corresponding author at: Neurology Department, University Hospitals Leuven, Herestraat 49, 3000 Leuven, Belgium. Tel.: +32 16 344280; fax: +32 16 344285.
E-mail address: rik.vandenberghe@uz.kuleuven.ac.be (R. Vandenberghe).

2213-1582/5 – see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.nicl.2013.03.014
6. Amyloid imaging as a road sign in the clinical development path .................................................. 503
   6.1. Requirements ...................................................................................................................................... 503
   6.2. Patient selection .................................................................................................................................. 503
   6.3. Target engagement .............................................................................................................................. 504
7. Amyloid imaging in clinical practice ....................................................................................................... 504
   7.1. Amyloid imaging and the ideal biomarker .............................................................................................. 504
   7.2. Clinical utility of amyloid imaging ......................................................................................................... 505
   7.3. Predictive value in MCI ......................................................................................................................... 506
   7.4. Differential diagnosis ............................................................................................................................ 506
8. The multidimensional space of Alzheimer’s disease .................................................................................. 507
9. Conclusion .................................................................................................................................................. 507
Acknowledgments ......................................................................................................................................... 507
References ..................................................................................................................................................... 508

1. Introduction

Cognitive and behavioral decline due to neurodegenerative disease is highly prevalent across the world. Apart from the pervasive personal and familial impact, the public health burden may become unsurmountable for healthcare systems over the next decades (Brookmeyer et al., 1998; Hebert et al., 2001) except if more efficacious interventions to prevent, halt or slow down Alzheimer’s disease (AD) (Brookmeyer et al., 1998) are discovered. The ability to characterize directly in humans the underly- ing pathophysiological processes is fundamental to progress in AD research and therapy. After more than a decade of therapeutic stagnation, the time may have come to re-think the clinical concept of Alzheimer’s disease and to re-draw the clinical development path for testing novel candidate AD drugs. We will evaluate to which degree the availability of amyloid imaging, amid other techniques, could alter our clinical-diagnostic approach to AD and how this may enhance progress towards more efficacious therapy.

2. Some basic AD principles relevant to amyloid imaging

The initial abnormalities in AD probably occur at the functional level, involving synaptic and neuronal dysfunction, possibly initiated by abnormalities in soluble Aβ42. The exact temporal relationship between the diverse structural alterations that follow is still a matter of active neuropathological and in vivo research (see below). In the revised National Institute on Aging-Alzheimer’s Association (NIA-AA) neuropathological criteria (Hyman et al., 2012), three different classification schemes have been adopted in parallel, depending on the phase of the spread of amyloid plaques (Thal et al., 2002) and neurofibrillary tangles (Braak and Braak, 1991) and the amount of neuritic plaques as defined by the Consortium to Establish a Registry for Alzheimer’s Disease (CERAD) criteria.

The principal structural changes are:

- Loss of synaptic density (Davis et al., 1999; DeKosky and Scheff, 1990; Scheff et al., 2006; Terry et al., 1991), starting in the dentate gyrus and correlated to episodic memory scores (Scheff et al., 2006)
- Neuronal loss, starting in entorhinal cortex (Gómez-Isla et al., 1996), and correlated to cognitive scores (Giannakopoulos et al., 2003)
- Neurofibrillary tangles (NFT) and neuroil threads (Hyman et al., 2012). Initially, neurofibrillary tangles are mainly present in entorhinal and perirhinal cortex spreading to the hippocampus (stage I/II). They subsequently spread to inferior temporal and lateral temporal cortex (stage III/IV) and then become widely distributed over neocortical association zones (stages V/VI) (Braak and Braak, 1991; Hyman et al., 2012; Tirellabschi et al., 2004). These NFT stages correlate relatively well with neuronal loss and with the severity of clinical symptoms (Arriagada et al., 1992; Bancher et al., 1993; Giannakopoulos et al., 1998, 2003; Guillozet et al., 2003; Mesulam, 1999; Mesulam et al., 2004).
- Aβ amyloid aggregates, which may take different forms (Hyman et al., 2012). According to some authors, diffuse plaques should not be considered as pathological since they are not associated with synapse loss or neuronal loss, key features of Alzheimer’s disease (Masliah et al., 1990). Others have suggested that diffuse amyloid plaques are related to the presymptomatic stage of AD (Morris et al., 1996). In contrast to diffuse plaques, neuritic plaques stain with thioflavine S or Bielschowsky, indicative for the presence of tau pathology invading or surrounding the plaque. A time sequence analogous to that described for NFTs has been developed to describe the propagation of amyloid, starting in neocortical areas (phase 1) and then spreading to allocortical regions including, among other regions, entorhinal cortex, CA1, anterior and posterior cingulate (phase 2), basal forebrain nuclei, diencephalic nuclei and striatum (phase 3), brain stem nuclei (phase 4) and further into the molecular layer of the cerebellum (phase 5) (Thal et al., 2002). At a certain stage the increase in β amyloid aggregates may level off, also referred to as a ‘growth arrest’ of the amyloid plaques (Christie et al., 2001; Hyman et al., 1993).

The clinical phenotype and severity in AD is determined primarily by the topographic distribution and density of neuronal loss and NFTs, starting in the hippocampal formation and extending into inferior and lateral temporal cortex and beyond (Arriagada et al., 1992; Bancher et al., 1993; Giannakopoulos et al., 1998, 2003; Guillozet et al., 2003; Mesulam, 1999; Mesulam et al., 2004). In atypical variants of AD, such as posterior cortical atrophy (Hof et al., 1993, 1997; von Gunten et al., 2006) or the logopenic variant of primary progressive aphasia (Gefen et al., 2012), the unusual clinical phenotype at initial presentation is mirrored by an unusual distribution of neuronal loss and NFTs. Conversely, when other proteinopathies, such as TAR DNA Binding Protein (TDP-43) proteinopathies (Pao et al., 2011) or tauopathies (Hornberger et al., 2012) have an atypical distribution and predominantly affect medial temporal cortex, they may mimic the amnestic presentation of AD.

As age increases, the classical hallmark lesions of AD (NFT and neuritic plaques) allow one to discriminate less and less reliably between subjects with versus without dementia (Davis et al., 1999; Prohovnik et al., 2006). In a prospective population-based series of 456 autopsy cases between 70 and 100 years of age (Savva et al., 2009), a strong relationship between dementia and the presence of NFT and neuritic plaques was seen until 75 years of age (Savva et al., 2009). This relationship became weaker above the age of 80, especially for neuritic plaques. Diffuse plaques were only weakly associated with dementia, even at the age of 75 (Savva et al., 2009). In contrast, neuropathological measures of volume loss retained their strong relationship with the presence or absence of dementia over the entire age range (Savva et al., 2009). Prospective community-based neuropathological studies have highlighted the contribution of other lesions to cognitive decline, such as cerebrovascular lesions or Lewy bodies, independently (Schneider et al., 2012) or interactively (Snowdon et al., 1997) with Alzheimer pathology. Cognitive
decline therefore is the integral of a combination of different, at least partly independent factors, leading to the concept of ‘multifactorial AD’.

In line with this concept, NIA-AA guidelines for the neuropathologic assessment of AD put emphasis on the systematic evaluation of concomitant abnormalities, e.g. Lewy bodies, TDP43 inclusions and vascular changes (Hyman et al., 2012). In the same vein, the novel NIA-AA criteria for clinical AD separate out a category of possible AD with ‘etiologically mixed’ presentation (McKhan et al., 2011). This subgroup includes subjects who have concomitant neuropathological findings of cerebrovascular disease, TDP43-proteopathies or Lewy bodies. This subgroup also includes cases where medical comorbidity and use of drugs contribute to the cognitive dysfunction.

3. Comparisons between amyloid ligands

The amyloid ligands that are available for research use are the thioflavin T derived 11C-Pittsburgh Compound B (PIB) (Klunk et al., 2004), its 18F-labeled analog 18F-flutemetamol (Koole et al., 2009; Nelissen et al., 2009), the benzofuranes 11C-AZD-2184 (Nyberg et al., 2009)( Table 1). Overall, the discriminative power was higher for the latter compound (Table 1). Other studies compared 18F-AZD-4694 (Cselényi et al., 2012), the benzoxazole 11C-BF-227 (Furukawa et al., 2010), the stilbene compounds 11C-SB (Verhoeff et al., 2004), 18F-florbetaben (Barthel et al., 2011; Rowe et al., 2008) and 18F-florbetapir (Wong et al., 2010), and the naphthol 18F-FDDNP (Small et al., 2006). Of these, at the time of writing, only 18F-florbetapir is approved by the Food and Drugs Administration (FDA) and European Medicines Agency (EMA) for the clinical evaluation of cognitive deficits in patients. For review of the development and biochemical characteristics of these compounds as well as methodological nuclear imaging issues we refer to Herholz and Ebmeier (2011), and biochemical characteristics of these compounds as well as methodological nuclear imaging issues we refer to Herholz and Ebmeier (2011), and Mathis et al. (2012). Here we will focus on direct comparative studies between different amyloid ligands within the same subjects (Tables 1, 2).

Verhoeff et al. (2004) compared the stilbene 11C-SB and 11C-PIB (Verhoeff et al., 2004) (Table 1). Overall, the discriminative power was higher for the latter compound (Table 1). Other studies compared 18F-labeled amyloid ligands to 11C-PIB (Table 2). In 20 AD and 20 amnestic MCI patients (Vandenbergh et al., 2010), volume-of-interest based Standardized Uptake Value Ratios (SUVR) with cerebellar gray matter as reference region were compared between 11C-PIB and 18F-flutemetamol by means of linear regression (Table 2) (Vandenbergh et al., 2010). There was a close match of SUVR levels between the two compounds in neocortical volumes of interest (Table 2). The correlation was much weaker in subcortical white matter and pons where retention of the 18F amyloid ligand was higher than that of 11C-PIB (Vandenbergh et al., 2010).

Wolk et al. (2012) compared 18F-florbetapir and 11C-PIB in 14 AD and 15 healthy controls. The separation between controls and AD was wider for 11C-PIB, both in absolute terms and in terms of S.D. (healthy controls (HC) mean SUVR composite cortical VOI (SUVRcomp) = 1.07 (SD 0.23), AD mean SUVRcomp 1.94 (SD 0.36)) than for 18F-florbetapir (HC mean SUVRcomp 1.06 (SD 0.17), AD mean SUVRcomp 1.38 (SD 0.15)).

11C-PIB has also been directly compared with 18F-FDDNP in 14 patients with AD, 11 patients with amnestic MCI and 13 controls. The correlation of binding potentials between the two tracers was only 0.45 (Table 2). Values in hippocampus were higher for 18F-FDDNP than for 11C-PIB (Shin et al., 2010; Tolboom et al., 2009a). This was also true in inferior temporal gyrus and secondary visual association cortex (Shin et al., 2010), probably related to the predilection of NFTs for these regions (Giannakopoulos et al., 1998). In other neocortical association areas, binding potential values in the AD group were more than 9-fold lower for 18F-FDDNP than for 11C-PIB, reflecting lower affinity of 18F-FDDNP for neuritic plaques. The differences between groups (AD, MCI and controls) were more pronounced for 11C-PIB than for 18F-FDDNP (Tolboom et al., 2009a).

It is often assumed that 18F-amyloid ligands perform more or less equivalently for separating AD versus controls. Only direct comparisons between the 18F-labeled amyloid ligands within the same subjects can resolve this important question, and head-to-head comparisons are currently underway. As several competing ligands will enter the commercial arena, such direct comparisons will be of great clinical interest.

4. Comparisons between amyloid ligands and neuropathology

3H-PIB colocalizes with neuritic plaques and also with diffuse amyloid plaques and Aβ in the blood vessel wall (Lockhart et al., 2007), and much less so with tangles (Lockhart et al., 2007) or other β-sheets aggregates such as Lewy bodies (Lockhart et al., 2007). 6-CN-PIB, a highly fluorescent derivative of 11C-PIB, has affinity for plaques (Ikonomovic et al., 2008), more so for neuritic than for diffuse plaques (Ikonomovic et al., 2012), as well as affinity for β amyloid in the vessel walls (Backskoi et al., 2007) and for striatal plaques (Ikonomovic et al., 2008). There is no affinity for cerebellar plaques and only partial affinity for extracellular ‘ghost tangles’, probably due to adjacent neuritic plaques (Ikonomovic et al., 2008). 11C-PIB retention values correlate with Enzyme Linked ImmunoSorbent Assay (ELISA) measurements of Aβ42 more than α40 (Ikonomovic et al., 2012).

The level of exposure of the different types of aggregates to the tracer may drastically differ between the brain slices and the in vivo imaging situation (Lockhart et al., 2007). For that reason, quantitative correlational studies between in vivo measures of amyloid ligand retention and quantitative neuropathological measures of the presence of Aβ are an essential step for biomarker validation.

4.1. Large-scale prospective postmortem validation studies

In the clinicopathological 18F-florbetapir phase 3 study (Clark et al., 2011, 2012), 59 cases (mean age 79 years) received a 10-min florbetapir PET and a neuropathological assessment within 1–2 years following the PET (Clark et al., 2012). Twelve had no cognitive impairment, 29 had clinical AD, and 13 had a clinical diagnosis of some other form of dementia (Lewy body dementia (LDB), Parkinson’s Disease with Dementia (PDD), frontotemporal lobar degeneration (FTLD), unspecified, mixed). When binary scan reads were compared to binarized CERAD neuritic amyloid plaque scores (moderate/frequent vs sparse/none), specificity was 100% and sensitivity around 92–96%. 18F-florbetapir retention levels also correlated with β amyloid area as determined by 4G8 amyloid immunostaining (antibody against Aβ residues 17–24). The phase 3 clinicopathological studies with the other 18F- ligands have not been published at the time of writing.

A series of biopsy studies correlated β amyloid 4G8 immunostaining and 18F-flutemetamol retention in normal pressure hydrocephalus. These studies confirmed the correlation between ligand retention levels and β amyloid plaque surface area (Leinonen et al., 2013; Rinne et al., 2012; Wong et al., 2013).

| Table 1 | Cohen’s d effect size for discriminating between AD and controls (difference between the means of the two study populations divided by the pooled standard deviation). These values are for indicative purposes and, as indicated, may substantially vary between studies depending on the exact processing and analysis method used. For the sake of comparison, the effect size obtained for FDG-PET for discriminating AD from controls in a comparative study of FDG-PET and 11C-PIB-PET was 1.53 (Ng et al., 2007). |
|----|----|
| 11C-PIB | 2.23–3.8 | Verhoeff et al. (2004), Ng et al. (2007) and Villemagne et al. (2012) |
| 11C-SB | 1.22 | Verhoeff et al. (2004) |
| 18F-flutemetamol | 2.7 | Vandenbergh et al. (2010) |
| 18F-florbetapir | 2.3 | Wong et al. (2010) |
| 18F-florbetaben | 1.37–3 | Barthel et al. (2011) and Villemagne et al. (2011, 2012) |
4.2. Other clinicopathological studies

The large-scale clinicopathological studies have been pivotal in demonstrating the relationship between ligand retention and amyloid plaque surface area. Smaller case series from academic investigator-driven studies provide complimentary information to what can be obtained from larger-scale industry-sponsored trials. A number of cases exhibited a dissociation between amyloid positivity/negativity and the presence/absence of Alzheimer’s disease according to consensus CERAD or The National Institute on Aging and Reagan Institute (NIA-RI) neuropathological criteria. It is important to note that the terms ‘false-positive’ and ‘false-negative’ depend on the comparator used. If the amyloid plaque area is used as comparator and both neuritic and diffuse plaques are counted, false-positives have not been described until now and false-negatives only exceptionally (Cairns et al., 2009). If conventional neuropathological criteria are used as comparator, mismatches will inevitably occur as these criteria also integrate other AD-related features such as NFTs. Nevertheless, in a clinical review, when assessing clinicopathological correlations, it is useful to compare with such well-established nosological criteria. Conceivably, in the future biomarker research may lead to novel taxonomic classifications that are based on one specific lesion type (such as a diagnostic category of β amyloidosis), among the different abnormalities that currently define Alzheimer’s disease, but such novel classifications would first have to be validated and prove their utility in drug development and clinical practice.

4.2.1. False-positive cases with CERAD and NIA-RI criteria as comparator

In the Baltimore Longitudinal Study of Aging (BLSA), a community-recruited cohort is followed longitudinally with cognitive assessments until close to time of death and brain autopsy (Driscoll et al., 2012; Sojkova et al., 2011). In the 5 nondemented and 1 demented case reported, 11C-PIB retention values had an almost linear relationship with such well-established nosological criteria. Conceivably, in the future biomarker research may lead to novel taxonomic classifications that are based on one specific lesion type (such as a diagnostic category of β amyloidosis), among the different abnormalities that currently define Alzheimer’s disease, but such novel classifications would first have to be validated and prove their utility in drug development and clinical practice.

511

4.2.2. False-negative cases with CERAD and NIA-RI criteria as comparator

Another subject of the BLSA series (Driscoll et al., 2012; Sojkova et al., 2011) had a moderate amount of neuritic plaques according to CERAD classification while Distribution Volume Ratio (DVR) for 11C-PIB was only around the cut-off of 1.2. This would be considered a false-negative scan if conventional neuropathological criteria for AD diagnosis are used as comparator. Another case with clinically probable Lewy body disease met CERAD criteria for definite AD due to focal deposition of abundant neuritic plaques limited to a prefrontal region but was 11C-PIB-negative (Ikonomovic et al., 2012). In a biopsy study of cases with Normal-Pressure Hydrocephalus (NPH) (Leenonen et al., 2008), one 11C-PIB-negative case met neuropathological criteria for definite AD. Cairns et al. (2009) described a case above the age of 85 who had progressive cognitive decline, a negative 11C-PIB scan, abundant diffuse plaques and a neuropathological diagnosis of possible AD according to CERAD criteria and low-probability of AD according to NIA-RI criteria (Cairns et al., 2009). This case is not a false-negative when conventional criteria are used as comparator but could be regarded as such if the diffuse plaque count is used as comparator, as is the case in the Katchaturian criteria (Cairns et al., 2009).

Part of the mismatch between the amyloid scan and the neuropathological diagnosis can be attributed to the way neuropathologically definite AD is defined: The regions sampled in the CERAD protocol do not strictly match the regions of predilection of amyloid ligand retention. Furthermore, in the NIA-RI criteria NFTs play an important role which are obviously not detected by amyloid imaging. Some neuropathologically definite PIB-negative AD cases however have neuritic plaques that seem 11C-PIB refractory (Rosen et al., 2010). For yet unknown reasons, 11C-PIB may have a low affinity for specific and rare types of neuritic plaques (Leenonen et al., 2008). Other such examples are the plaques occurring in the Arctic mutation carriers (Schll et al., 2012).

4.3. Affinity for white matter lesions

White matter lesions have a highly heterogeneous neuropathological basis. White matter lesions due to cerebral amyloid angiopathy (CAA) would be expected to bind amyloid ligands given the affinity of...
$^{11}$Cl-PIB for CAA (Bacskai et al., 2007; Dholander et al., 2011; Dierksen et al., 2010; Johnson et al., 2007; Lockhart et al., 2007). PIB also binds to $\beta$ sheet structures contained in myelin, e.g. myelin basic protein. This can lead to a focal white matter signal decrease in demyelinating lesions in multiple sclerosis (Stankoff et al., 2011). In such inflammatory conditions it may be important to model changes in blood–brain-barrier permeability that are associated with the disease as this may affect tracer kinetics. The same could possibly be true for white matter lesions due to vascular disease.

4.4. Summary

Continuous measures of ‘amyloid plaque area’ highly correlate with amyloid ligand retention values (Clark et al., 2011, 2012; Driscoll et al., 2012; Sojkova et al., 2011). The amyloid ligands bind both neuritic and diffuse plaques, although with higher affinity for neuritic than for diffuse plaques (Ikonomovic et al., 2012). The three main explanations for mismatches between binary results from amyloid ligand scans and neuropathological CERAD or NIA-RI classification are:

1. Binding of the amyloid ligand to diffuse plaques (Lockhart et al., 2007) in cases where diffuse plaques are abundant and neuritic plaques sparse or absent.

2. PIB refractory cases’ where affinity of the amyloid ligand for neuritic plaques appears to be much lower for as yet poorly understood reasons.

3. A sampling bias that is inherent to the CERAD criteria and where the regions selected for neuropathological diagnosis do not match those that typically show high ligand retention in vivo (such as posterior and anterior cingulate and precuneus) (Driscoll et al., 2012).

Overall, the mismatches may be of high scientific interest but among the total group of amyloid PET scans performed, they probably represent only a minority of cases as evidenced by the prospective phase 3 neuropathological studies (Clark et al., 2011, 2012).

5. The AD puzzle: putting the pieces together by means of multimodal imaging

Let us turn to the relationship between amyloid imaging and other in vivo measures of AD pathology that are in clinical use (including cognitive assessment, structural and functional MRI or FDG PET), cross-sectionally and longitudinally. If correlations are calculated between ligand retention levels and cognitive scores across diagnostic groups (patients and cognitively intact controls), correlations will be almost certainly found but such correlations may be indirect: For example, as AD patients have lower cognitive scores than controls and also higher Aβ load, a correlation of cognitive scores with Aβ load across groups will be easily found. For a discussion of related issues in subjects with subjective cognitive impairment (SCI) (see Chételat et al., 2013).

5.1. Topography

Different brain regions have intrinsically different susceptibilities to different pathological expressions of AD (Jack et al., 2008; La Joie et al., 2012). A relatively consistent spatiotemporal pattern of volume loss has been described using structural MRI (Baron et al., 2001; Becker et al., 2011; Rombouts et al., 2000; Whitwell et al., 2008). According to this ‘AD signature’ (Dickerson et al., 2009), most severe cortical thinning occurs in rostral medial temporal and anterior temporal cortex extending posteriorly along the middle temporal gyrus, inferior parietal cortex, temporoparietal junction (TPJ), ventral premotor cortex and the prefrontal and posterior cingulate (Dickerson et al., 2009). The typical topography of hypometabolism in early-stage Alzheimer’s disease versus normal controls encompasses posterior cingulate and prefrontal, temporoparietal association cortex, prefrontal and ventromedial frontal cortex, with sparing of primary sensory and motor cortex, basal ganglia and thalamus (Herholz et al., 2002; Mosconi et al., 2008).

The regional distribution of $^{11}$Cl-PIB retention only partly overlaps with these well-known patterns of brain volume loss and hypometabolism. The overlap between $^{11}$Cl-PIB increase and volume loss mainly occurs in precuneus and lateral temporoparietal association cortex in controls, amnestic MCI and AD (Jack et al., 2008). In these regions, there is also concordance between Aβ deposition and hypometabolism (La Joie et al., 2012; Shin et al., 2010). In hippocampus, however, there is a discordance between volume loss and amyloid ligand retention: volume loss is pronouned with relatively little Aβ deposition and limited hypometabolism. A third pattern is seen in dorsolateral prefrontal cortex, with high ligand retention and limited volume loss or hypometabolism (Jack et al., 2008; La Joie et al., 2012).

Within the MCI or AD group, global $^{11}$Cl-PIB-retention levels do not correlate with structural volume loss (Bourgeat et al., 2010; Chételat et al., 2010). At the single-case level, several cases show dissociations between the presence of cortical $^{11}$Cl-PIB retention and hippocampal volume loss, in both directions (Jack et al., 2008; Vandenberghe et al., 2013).

5.2. Relationship with disease severity

In AD, a structural MRI-based linear model of medial temporal, inferior temporal and inferior frontal regions best predicts clinical severity, as measured by the Clinical Dementia Rating scale—Sum of Boxes (Dickerson et al., 2009). The pattern of hypometabolism also correlates with the type (Engler et al., 2006) and severity (Herholz et al., 2002; Jagust et al., 2009; Rabinovici et al., 2010) of clinical deficits in AD. In contrast, the correlation between PET measures of Aβ deposition and cognitive scores is weak or absent within the group of clinically probable AD subjects (Engler et al., 2006; Jagust et al., 2009; Rabinovici et al., 2010). Within the MCI group, episodic memory scores are lower in the amyloid-positive than the — negative MCI cases (Rowe et al., 2010; Wolk et al., 2009). If both hippocampal volume and $^{11}$Cl-PIB retention values are entered into a regression model in 39 $^{11}$Cl-PIB-positive MCI cases and 17 healthy controls, only hippocampal volume make an independent contribution to episodic memory scores (Mormino et al., 2009).

5.3. Longitudinal change

In one of the first multimodal imaging studies with longitudinal amyloid imaging in controls, MCI and AD (Jack et al., 2009), ventricular expansion rate measured with MRI differed between groups (controls 1.3 cm$^3$/yr, MCI 2.5 cm$^3$/yr, AD: 7.7 cm$^3$/yr). This correlated with clinical decline (Jack et al., 2009). In contrast, in controls, MCI and AD, amyloid ligand retention only showed a small increase over time in the absence of between-group differences (Jack et al., 2009) and without any relationship to cognitive decline (Fouquet et al., 2009; Furst et al., 2012; Landau et al., 2011). In MCI in particular, there was a wide variability in the amount of increase in amyloid ligand retention and in some cases the rate of change was higher than in any of the two other groups (Jack et al., 2009). Longitudinal studies based on FDG-PET have demonstrated widespread metabolic changes in MCI and AD, affecting, among other regions, posterior cingulate, precuneus, and medial temporal cortex (Chen et al., 2010; Fouquet et al., 2009). These changes correlate with longitudinal cognitive decline (Fouquet et al., 2009; Furst et al., 2012; Landau et al., 2011).

Other studies have confirmed that the longitudinal change in Aβ deposition in AD and MCI is relatively small. For instance, in both AD and controls, a 3–4% increase in $^{11}$C-PIB uptake over a two-year period was restricted to medial frontal cortex (Scheinin et al., 2009). In a study with 3–5 years follow-up, $^{11}$C-PIB uptake increased in MCI but remained...
stationary in the AD group, be it with considerable inter-individual heterogeneity (Kadir and Nordberg, 2010). One of the largest longitudinal studies reported so far included 103 controls, 49 MCI patients and 32 AD patients (Villain et al., 2012). There was a small but noticeable increase in amyloid ligand retention over time, even in the clinical stage. Higher rates of increase in Aβ deposition were seen with higher baseline levels, regardless of the group, leading to the novel concept of 11C-PiB accumulators and non-accumulators (Villain et al., 2012). This resonates with the basic neurobiological idea of self-propagation of Aβ deposition (Lee et al., 2010). The amount of increase in amyloid ligand retention was regionally dependent: At baseline, precuneus and temporoparietal junction showed the highest levels while the rate of change was most pronounced in posterior temporal cortex (see also Chételat et al., 2013). Between-subject variability, however, was high, which is also apparent from visual inspection of the data (Villain et al., 2012). Besides Aβ load at baseline, other baseline factors that have been associated with a higher increase in Aβ load over time in AD are APOE 4 status (Grimmer et al., 2010) and vascular white matter lesions (Grimmer et al., 2012). The latter effect has been attributed to reduced Aβ clearance (Grimmer et al., 2012).

Driscoll et al. (2011) evaluated whether longitudinal changes in structural MRI over the preceding 10 years differed between amyloid-positive versus amyloid-negative non-demented cases. No such correlation was seen at the global or the regional level, except for a trend in the precuneus (Driscoll et al., 2011). This paper therefore shows a within-subject dissociation between MRI structural volume loss estimated over a 10-year period and amyloid ligand retention levels. In MCI, Tosun et al. (2011) performed a parallel independent component analysis of baseline amyloid scans and brain atrophy rates over the subsequent year. The main components consisted of increased amyloid levels in medial parietal regions together with higher medial temporal atrophy rate (Tosun et al., 2011).

5.4. Comparison of diagnostic performance

From a clinical perspective, diagnostic and prognostic performance at the individual level is the critical parameter. For instance, while group comparisons robustly show more pronounced hippocampal volume loss in AD patients versus controls, at the individual level there is considerable overlap in hippocampal volumes between AD and healthy controls (Jack and Petersen, 2000).

When the diagnostic performance of 18F-flutemetamol PET and structural MRI is compared at the individual level in AD, MCI and controls using machine learning techniques, overall sensitivity was comparable between the two diagnostic techniques when the clinical diagnosis was used as comparator (85.2%) (Vandenberghe et al., 2013). Specificity however of the 18F-flutemetamol based classifier (92%) was higher than that of the gray matter based classifier (68%) (Vandenberghe et al., 2013). The gray matter based classifier classified more scans from healthy controls as pathological than the 18F-flutemetamol based classifier (Vandenberghe et al., 2013). This could be partly explained by the overlap in hippocampal volumes between the AD group and the controls (Duara et al., 2012; Thurjell et al., 2012). One of the regions that best discriminated between Alzheimer’s disease and normal controls for 18F-flutemetamol according to the classifier’s feature weights was the striatum, mainly in its anterior portion (Vandenberghe et al., 2013). Striatum is involved in AD-related β amyloidosis from Thal stage 3 onwards (see above) (Thal et al., 2002). Early neuropathological studies have revealed the presence of amyloid plaques in striatum (Brilliant et al., 1997; Gearing et al., 1993; Rudelli et al., 1984), in particular the ventral striatum (Suenaga et al., 1990). The plaques in dorsal and ventral striatum are immunochemically distinct from cortical plaques and also differ between dorsal and ventral striatum, with relatively more diffuse plaques in dorsal striatum and more dystrophic neurites in association with the plaques in ventral striatum (Suenaga et al., 1990). Possibly, the latter pattern may relate to the connections of ventral striatum with mesolimbic and allocortical regions and with orbitofrontal cortex (Aggleton et al., 1987; Cavada et al., 2000; Haber et al., 1990), structures that are affected early in the AD course by neuronal loss and tau pathology.

It is not common clinical practice in radiology to provide binary classifications of MRI scans in relation to AD. Binary reads, however, of FDG-PET scans are performed by nuclear medicine physicians on a daily basis and it is therefore of clinical relevance to compare the diagnostic accuracy of binary classification of FDG-PET with that of amyloid PET. When clinical diagnosis (AD versus healthy controls) is used as comparator, visual readings as well as Receiver Operating Characteristics (ROC) analysis show higher accuracy for 11C-PiB than for FDG-PET (90% vs 70% for the reads, 95% and 83% for the ROC) (Ng et al., 2007). Higher sensitivity and specificity of 11C-PiB than FDG-PET was also seen in a study by Devanand et al. (2010). Values were compared at a global level and also for the regions that showed highest diagnostic accuracy per technique (precuneus for 11C-PiB and parietal cortex for FDG-PET). Sensitivity and specificity of global 11C-PiB retention levels were 94% while sensitivity of mean regional cerebral metabolic glucose rate (CMRGlucose) was 81% and specificity 71%. In precuneus, sensitivity and specificity of 11C-PiB was 94%, while sensitivity and specificity of rCMRGlucose in parietal cortex was 87% and 88%, respectively (Devanand et al., 2010). In MCI, discrepancies between 11C-PiB PET and FDG-PET were more marked and 11C-PiB PET had higher diagnostic accuracy for the distinction between healthy controls and MCI than FDG-PET (Devanand et al., 2010).

When clinically probable AD patients (n = 62) were compared with FTLD patients (n = 45, including behavioral variant frontotemporal dementia and nonfluent and semantic variants of primary progressive aphasia (PPA)) and the clinical diagnosis was used as standard-of-truth, sensitivity of visual reads of 11C-PiB scans for AD (89%) was higher than that of FDG-PET (73%), with similar specificity (83–84%) (Rabinovici et al., 2007, 2008, 2011). Interrater agreement of the visual reads was higher for 11C-PiB (Fleiss κ 0.96) than for FDG-PET (Fleiss κ 0.72) (Rabinovici et al., 2011). Accordingly, when neuropathology is used as gold standard, unanimity in the overall interpretation of FDG-PET scans between readers for discriminating AD from FTLD was relatively low, in particular for the scans from FTLD cases (unanimity in 7 out of 14 cases, compared to 27 out of 31 AD cases) (Womack et al., 2011). With quantitative analysis, PiB had a higher sensitivity (89% vs 73%) but FDG a higher specificity (83% vs 98%) (Rabinovici et al., 2011). The higher specificity probably relates to the occurrence of positive amyloid scans even in cognitively intact older adults. High interrater agreement (Rabinovici et al., 2011) together with higher sensitivity (Rabinovici et al., 2011) may be important clinical advantages that amyloid PET would bring compared to FDG PET in the differential diagnosis between AD and FTLD, although further studies with neuropathological diagnosis as comparator will be required.

5.5. Task-related fMRI versus amyloid PET

Task-related fMRI can reveal adaptive changes in cognitive brain circuits during cognitive processing which may provide resilience against the functional impact of amyloid pathology. Brain resilience and functional reorganization may gain importance in the field as AD therapies may increasingly target multiple facets of AD. To the best of our knowledge, two studies combined amyloid PET and task-related fMRI within the same subjects in MCI or AD.

The first such study focused on changes in the language circuit in early-stage clinically probable AD and how this related to amyloid ligand retention (Nelissen et al., 2007). This study built on a preceding study in amnestic MCI: in amnestic MCI, the earliest changes in the language network occur in the posterior part of the left superior temporal sulcus and these changes correlate with subclinical changes in written word identification speed, in line with the critical role of

R. Vandenberghe et al. / NeuroImage: Clinical 2 (2013) 497–511
this region in lexical-semantic retrieval (Vandenbulcke et al., 2007). In
early-stage AD, hypooactivity is also seen in left posterior temporal cor-
tex during associative-semantic versus visuooperceptual processing.
These fMRI changes in the left posterior STS in AD correlate with offline
measures of confrontation naming: clinically probable AD subjects with
hypooactivity in this region are impaired on confrontation naming
(Nelissen et al., 2007). There is no correlation between 11C-PIB levels
in this region and the clinical deficit (Nelissen et al., 2007). fMRI activity
levels to the homotopical right side show increased activity in those pa-
tients who had preserved naming on the confrontation naming, sugges-
tive of a compensatory increase (Nelissen et al., 2007).

In the episodic memory domain, a combined 11C-PIB-fMRI study in
subjects with a Clinical Dementia Rating Scale of 0–0.5 revealed mainly
differential activity decreases as a function of Aβ deposition. During
encoding of associations between faces and names, the precuneus and
temporoparietal cortex, partially overlapping with the posterior
nodes of the default mode network, were less deactivated in amyloid-
positive subjects compared to amyloid-negative subjects (Sperling et al.,
2009). Deactivation in PCC is inversely related to hippocampal activation
(Celone et al., 2006).

5.6. Summary

Imaging techniques such as structural MRI and FDG PET reveal the
topography of structurally and functionally affected regions at an
anatomical level. This correlates with the type and degree of clinical
deficit, cross-sectionally and longitudinally. Uniquely, amyloid imag-
ning provides direct information about lesion type, yielding a superior
accurate prediction compared to structural imaging techniques. Thus, it is
essential to consider lesion type in the context of disease diagnosis.
Within the AD group, however, the extent and topographical distribution of
amyloid plaques due to the passive immunization.

The more fundamental question concerns the biological signifi-
cance of a change in amyloid ligand retention and how this translates into
clinical benefit given the fact that the relationship between amyloid
ligand retention and clinical expression is limited (see above). In our
view, the critical contribution of amyloid imaging to amyloid lowering
drug development lies in the selection of study patients and the demon-
stration of drug target engagement.

6. Amyloid Imaging as a Road Sign in the Clinical Development Path

6.1. Requirements

In clinical trial research, the issues related to evaluating longitudi-

1. The change over time in the normal population and the normal vari-
   ance needs to be quantified. This is necessary to determine if any
   change observed in an individual, either positive or negative, falls
   within the normal range or not.

2. Test–retest variability of the technique at a short interval must be
   known: Test–retest variability of 11C-PIB SUVR in AD is 8.0% (SD 7)
   (Tolboom et al., 2009b), but can be improved to 3% by using dynamic
   scans with reference tissue methods (Tolboom et al., 2009b). The
test–retest variability of the 18F-labeled amyloid ligand is provided
in Table 3 (Cselényi et al., 2012; Joshi et al., 2012; Vandenberghe et
al., 2010; Villemagne et al., 2011). To detect a drug-induced differ-
ence, test–retest variability within the same center needs to be
kept as low as possible: this will depend not only on the analysis
method (Tolboom et al., 2009b) but also on changes in structural
brain volume over time and partial volume effect, motion during
scanning etc.

3. Between-center variability should be measured: differences in scanner,
   reconstruction algorithm etc. may affect retention levels. This is
an active area of investigation which is highly relevant for multicenter
trials. A European multicenter study of 11C-PIB using different scan-
ers and reconstruction algorithm suggested that the range of values
in AD, MCI and controls are relatively comparable between centers
(Nordberg et al., 2013)

4. The study drug should not affect the metrics of the technique
directly. Both in Rinne et al. (2010) and Ostrowitzki et al. (2012),
competition assays did not reveal any alteration of the tracer affinity
of amyloid plaques due to the passive immunization.

6.2. Patient selection

In most mono- and multicentric diagnostic trials of early-stage clini-
cally probable AD from academic memory clinics, the proportion of
AD cases with a negative amyloid PET scan is approximately 10% (Devanand
et al., 2010; Nelissen et al., 2007, 2009; Vandenberghe et
al., 2010). In other multicenter diagnostic trials these percentages have
been as high as 32% (Doraiswamy et al., 2012). In the PET substudy of
the bapineuzumab phase 2 therapeutic trial, 8 of the 53 patients screened
for inclusion did not meet 11C-PIB criteria (Rinne et al., 2010). It is likely
that similar or higher proportions of amyloid-negative cases have en-
tered amyloid-lowering drug trials as these trials most often did not
verify amyloid status at inclusion, possibly leading to dismissal of poten-
tially useful drugs. The proportion of false-positive clinical AD diagnoses
may further increase as one moves towards the predementia stage,
reducing the chance of a positive study outcome. Regardless of the
exact mechanism of action of the drug tested, improved patient selection
may be one of the principal advantages that amyloid imaging at inclusion will bring. This may restrict the applicability of the drug once it is licensed. Such considerations however lose their validity in the face of the current crisis in AD therapeutic development.

6.3. Target engagement

For drugs that have as a purpose to lower amyloid levels, amyloid imaging also allows one to evaluate target engagement. If a drug engages its target but still does not exert a beneficial effect, this is highly informative from a scientific viewpoint, while a negative trial in the absence of any knowledge about target engagement is open to many more interpretations. Rinne et al. (2010) and Ostrowitzki et al. (2012) reported longitudinal changes during passive immunization with bapineuzumab (a humanized monoclonal antibody) and gantenerumab (a fully human monoclonal antibody), respectively (Table 5). In both studies, the relative small sample size led to imbalances in the $^{11}$C-PIB baseline values between groups. In the Ostrowitzki et al. (2012) study, the highest initial SUVR value in the placebo group was 2.3 while in the actively treated groups at least 5 subjects had values above 3, and in two of them between 3.5 and 4 (Ostrowitzki et al., 2012). This mismatch in baseline value makes interpretation of differences in change between groups difficult, in particular since initial $^{11}$C-PIB level affects the subsequent rate of increase in $^{11}$C-PIB retention (Villain et al., 2012). Furthermore, in the Ostrowitzki et al. (2012) study, the number of infusions received also differed between groups: two of the placebo received all 7 infusions, 6 of the 60 mg group and only 1 of the 200 mg group (1 case only two, 2 cases only 3 infusions, 2 four infusions) (Table 5). From the report it is unclear whether the number of infusions received affected the interval between first and end-of-treatment scan so that there could be a mismatch in duration between cases. The confidence intervals for change to baseline also included zero in all treatment groups (Ostrowitzki et al., 2012). For these reasons one has to be cautious in drawing strong conclusions from these initial studies using $^{11}$C-PIB to provide proof of target engagement by amyloid-lowering drugs. Based on these proof-of-concept studies, one could wonder whether prior stratification based on the biomarker value would help prevent initial mismatches at baseline.

### Table 4

| Reference | AD (n) | MCI (n) | Controls (n) | $\kappa$ |
|-----------|--------|---------|--------------|---------|
| $^{11}$C-PIB | 15     | 25      | 0.90         | Ng et al. (2007) |
| $^{18}$F-flutemetamol | 27     | 20      | 0.96         | Vandenbergh et al. (2010) |
| $^{18}$F-florbetapir | 10     | 30      | 0.76         | Joshi et al. (2012) |
| $^{18}$F-florbetaben | 81     | –       | 0.60         | Barthel et al. (2011) |

7. Amyloid imaging in clinical practice

While the validation of amyloid PET as a biomarker is relatively advanced, there have been very few studies that directly tested its utility (added value) and cost-effectiveness in diagnosis and patient management in a clinical context.

7.1. Amyloid imaging and the ideal biomarker

So far, clinical development of amyloid imaging markers has mainly focused on validating the imaging measure as a biomarker. Ideally, AD biomarkers have to fulfill a number of requirements (The Ronald and Nancy Reagan Research Institute of the Alzheimer Association, 1998). The biomarker should:

1. Detect a fundamental feature of pathophysiological processes active in AD. There is general consensus that $\beta$ amyloidosis is a fundamental feature to AD (Hardy and Selkoe, 2002). There however remains debate about the pathological nature of diffuse as opposed to neuritic plaques (Masliah et al., 1990; Morris et al., 1996). The selectivity of amyloid ligands for neuritic as opposed to diffuse plaques is discussed above.

2. Be validated in neuropathologically confirmed cases. As described above, the correlational studies with neuropathology have shown a sensitivity of 90% and specificity of 100% (Clark et al., 2011, 2012). It remains to be seen to which degree this can be extrapolated to the clinical populations in which the test would be applied. These patients would generally be in a much earlier disease stage, when, for instance, the relative proportion of diffuse versus neuritic plaques might be different. Conversely, in advanced cases the scans may be relatively more difficult to classify due to advance cortical atrophy.

3. Be precise, i.e. should detect AD early in its course and distinguish it from other dementias. For instance, in the phase 2 $^{18}$F-flutemetamol study, in which the prevalence of amyloid-positive healthy controls was relatively low, sensitivity and specificity to discriminate AD from controls with the clinical diagnosis as comparator was 93–96% (Vandenbergh et al., 2010). These values vary markedly between studies (Doraiswamy et al., 2012). This can be partly accounted for by differences in the prevalence of positive scans in the healthy control groups and differences in the prevalence of negative scans in the AD group. High sensitivity and specificity values have also been obtained in the phase 3 clinicopathological studies, which included a variety of non-AD dementias, each in low number (Clark et al., 2011, 2012). A growing number of papers have evaluated diagnostic performance in AD compared to other disease groups, such as Lewy body disease (Edison et al., 2008; Rowe et al., 2007; Villemagne et al., 2011), cerebral amyloid angiopathy (Dhollander et al., 2011; Dierksen et al., 2010; Johnson et al., 2007) and frontotemporal degeneration (Rabinovici et al., 2007, 2008, 2011; Rowe et al., 2007; Villemagne et al., 2011) and determined sensitivity and specificity with the clinical diagnosis as standard-of-truth (Villemagne et al., 2011). The added value compared to clinical diagnosis will only become clear once an independent standard-of-truth becomes available, e.g. through clinicopathological examination.

4. Should be reliable: Test–retest variability is reported in Table 3 and interrater reliability $\kappa$ in Table 4. For the sake of comparison, the interrater reliability of visual reads of FDG-PET in 20 healthy controls and 15 AD patients was 0.56 according to a comparative study of $^{11}$C-PIB and FDG-PET (Ng et al., 2007), but this may improve when statistical maps are used based on the contrast between an individual’s scan and a normal database rather than visual reads of native images (Foster et al., 2007; Herholz et al., 2002; Mosconi et al., 2008).

5. Simple to perform and inexpensive: For a health system to be sustainable given the demographic evolution, the cost of a diagnostic technique should be outweighed by the benefit the technique brings...
in terms of patient’s well-being. This benefit does not purely relate to therapeutic benefit but also in terms of communication of a correct diagnosis.

6. Recommended steps to establish a biomarker include confirmation by at least two independent studies from different groups of researchers.

7.2. Clinical utility of amyloid imaging

Both amyloid imaging and CSF biomarkers directly reflect one of the underlying neuropathological hallmark lesions of AD. The place that we will define for these biomarkers in our tentative diagnostic algorithm will be generally applicable to in vivo measures of AD-related amyloidosis, be it by means of PET or CSF (Fig. 1). The difference in utility between imaging and biochemical biomarkers will lie in the metrics of the tests, i.e. between-center and within-center replicability, the prevalence of intermediate values, and the between-reader variability in diagnostic interpretation of the values for amyloid imaging and CSF, respectively. It has been shown that binary assignment of cases based on 11C-PIB uptake into positive and negative cases closely corresponds to values of Aβ42 in cerebrospinal fluid but not with any of the other CSF variables such as total tau or Aβ40 (Fagan et al., 2006). From the viewpoint of implementation in a clinical environment, CSF biomarkers still have to go through the standardization steps that other diagnostic laboratory tests, e.g. for measuring protein levels in blood, have gone through (Bartlett et al., 2012; Mattsson et al., 2012). The values measured for a same sample vary between centers and the interpretation of a given CSF profile may also be center- and examiner-dependent. The factors that contribute to the between-center variability of CSF measurements are relatively poorly understood. In this sense, the strength of amyloid imaging compared to CSF at the time of writing seems to lie in its performance in terms of replicability between and within centers, between-reader replicability and standardization (see above).

The occurrence of AD neuropathology in cognitively intact individuals (see side-to-side review by Chételat et al. (2013)) necessarily puts an upper limit to the maximal specificity the technique can reach in comparison with healthy controls. The occurrence of a positive scan in a patient who presents at the memory clinic with cognitive complaints may reflect the underlying cause of cognitive dysfunction or, alternatively, may constitute a coincidental finding given the high prevalence of positive scans even in cognitively intact older adults in the higher age range (see Chételat et al. (2013)).

A second clinically relevant issue is the prevalence of ‘intermediate values’ (see also side-to-side review by Chételat et al. (2013)). This is a classical problem in diagnostic tests, which also exists for CSF and for FDG-PET. The number of cases within the gray zone will depend on the population tested and may differ substantially between modalities.

The added value of biomarkers for cerebral amyloidosis depends on:

1. The clinical context in which a diagnosis is made: primary care, neurology/psychiatry/geriatrics practice, academic memory clinic, ... Different clinical contexts are associated with differences in a priori probability of Alzheimer’s disease and non-AD dementias, age differences, differences in a priori probability of comorbidity or reversible cognitive causes, etc. These differences may affect the accuracy of a clinical AD diagnosis and the clinical impact of an additional biomarker measurement.

2. The healthcare system in which amyloid imaging is implemented. Factors that play a role and that are outside the scope of the current review are the degree of equality of access to healthcare that a society wishes to guarantee and budgetary restraints that have to be put in place to be able to provide this guarantee. Budgetary restraints will determine the cost-benefit ratio that is required for granting reimbursement of the biomarker.

3. The societal perception of Alzheimer’s disease. The latter determines the emotional impact of an AD diagnosis on the individual’s wellbeing in the different stages of the disease course.

Paradoxically, the added value of a biomarker for β amyloidosis may be higher in a clinical context where clinical expertise is relatively limited. Under such circumstances, the added value and cost-benefit of amyloid imaging would have to be compared with corrective measures directed at enhancing knowledge and clinical skills in this domain at different healthcare levels. Nevertheless, even in centers of excellence false-positive diagnoses of AD in the early disease stage may be more frequent than previously assumed. For instance, awareness is increasing that FTLD may present initially with an amnestic syndrome that is hardly distinguishable from what is seen in AD (Hornberger et al., 2012; Pao et al., 2011).

Possible indications for 18F amyloid PET in patients presenting with objective cognitive deficits in memory clinic

- Amnestic syndrome and at least one other cognitive domain affected, clinically suspect for a neurodegenerative cause
- Age < 60 yrs
- Age > 60 yrs
- Significant comorbidity that may explain cognitive syndrome
  - Psychiatric
  - Cerebrovascular
  - Cardiac
  - Metabolic
- Atypical associated features
  - Prominent aphasia
  - Comportmental changes
  - Early in continence
  - Early gait insufficiency
- Atypical presentations
  - Primary progressive aphasia
  - Comportmental disorder
- Rapid progression

Fig. 1. Tentative clinical algorithm for the utility of amyloid imaging in patients presenting at a memory clinic with a cognitive deficit that is confirmed by clinical assessment of cognitive functions or by neuropsychological evaluation. This algorithm is tentative and subject to further empirical evaluation.
The role of clinical context is also reflected by the NIA-AA criteria (Albert et al., 2011; McKhann et al., 2011). One set of clinical criteria is intended for use in all clinical settings. A separate set of criteria refers to ‘probable or possible AD dementia or mild cognitive impairment due to AD with low, high or intermediate level of evidence of the AD pathophysiological process’. A high level is defined as a concordance between markers for Aβ and markers of neuronal injury. The latter set is for use in three circumstances: investigational use, clinical trials (Section 5), and as optional clinical tools for use where available and when deemed appropriate by the clinician (McKhann et al., 2011). The topic of the remainder of this review is to delineate a number of circumstances where the use of biomarkers may be appropriate in clinical practice in our opinion.

The main clinical situations that we discern are predictive value in MCI and differential diagnosis of AD under specific clinical circumstances.

7.3. Predictive value in MCI

Approximately 50% of MCI patients have a positive amyloid scan (Okello et al., 2009; Pike et al., 2007; Vandenberghge et al., 2010; Wolk et al., 2009), with highest proportions (89%) in the multidomain amnestic MCI subtype (Wolk et al., 2009). It has been suggested that in MCI the amyloid ligand retention values have a bimodal distribution, but currently, to the best of our knowledge, there is no statistical proof of such a bimodal distribution (Mormino et al., 2009). A positive amyloid PET has a predictive value for future conversion to Alzheimer’s disease according to early studies with follow-up periods up to 3 years (Forsberg et al., 2008; Nordberg et al., 2013; Okello et al., 2009). In such studies, the clinician who decides about conversion to clinically probable AD ought to be blinded to the result of the initial PET scan. The interval over which MCI patients may convert to AD can be relatively long and a positive scan in a subject who has not converted after a 3 or 5 year interval is not necessarily a false-positive. It is also of clinical importance to note that a negative scan in a case with amnestic MCI does not preclude progression. Other diseases not associated with brain amyloidosis may cause an amnestic syndrome, such as frontotemporal degeneration, hippocampal sclerosis, argyrophilic grain disease or tangle-only AD. The risk for ‘conversion’ in such cases is probably as high as for MCI due to AD. A final point relates to the clinical utility of prediction in MCI patients in the absence of efficacious therapy. Ethically, it is important to avoid situations where the patient knows more than what he or she would consider desirable or beneficial. Such situations could for instance be patients with a diagnosis of amnestic MCI in whom acetylcholinesterase inhibitors are not of proven benefit and who would receive an earlier AD diagnosis due to the diagnostic advances.

In the phase 2 18F-florbetapir study (Doraiswamy et al., 2012), 51 MCI patients participated along with 69 healthy controls and 31 AD patients. 37% of MCI patients were amyloid positive (Doraiswamy et al., 2012). Forty-six of the MCI patients were followed for 18 months. This is one of the few studies that explicitly mentions blinding of the evaluator. Of the amyloid positive MCI subjects (n = 17) 5 progressed in an 18-months follow-up period. Of those who were amyloid negative (n = 29), 3 progressed. This study however suffers from a relatively low sample size, short follow-up time, and a relatively poor interrater reliability of the visual reads. Furthermore, the low prevalence of amyloid positivity in the AD group raises concerns about the clinical diagnostic accuracy prior to inclusion also in the MCI group.

7.4. Differential diagnosis

Accessibility of amyloid imaging in clinical practice will probably precede the evidence for clinical utility based on strict evidence-based medicine criteria. In our generic algorithm for use of biomarkers in an academic memory clinic, the indications for additional biomarkers are based on the a priori probability of Alzheimer’s disease in a given individual and the level of diagnostic certainty that one wishes to obtain.

We discern four situations where biomarker evidence may contribute to the differential diagnostic process in a way that is relevant for the patient and the caregivers concerned:

1. Early-onset dementia (onset before 60 years of age), in the absence of a known genetic mutation in the family: in early-onset dementia, the a priori probability of AD as a cause of cognitive decline is equal to that of FTLD. Furthermore, given the impact of the diagnosis that is communicated, a high level of diagnostic certainty is desirable (Hodges, 2001). The impact of a correct and early diagnosis in young-onset dementia is situated at the personal and familial emotional and relational level, and also at the level of third-party payers who have to make decisions on the strength of evidence for an underlying neurological cause.

2. Atypical presentations such as PPA or predominant compartmental disorder. Although detailed neurolinguistic phenotyping allows to estimate the probability of AD as an underlying cause of PPA (Gorno-Tempini et al., 2011), biomarker evidence may be very helpful. In an early stage of the disease, subtyping of individual cases may sometimes be difficult and may require a high level of neurolinguistic expertise that may not be widely available. Even with adequate clinical subtyping, the relationship between clinical phenotype and underlying pathophysiological substrate remains probabilistic, meaning that inferring the underlying lesion type based on the clinical phenotype will always be liable to a substantial degree of uncertainty (Mesulam, 2000). Furthermore, when the distribution of the AD pathology is atypical, imaging measures that reflect topography will provide less diagnostic information than evidence based on biomarkers measuring the underlying lesion type.

3. Cases where concomitant pathology could explain the cognitive decline but a diagnosis of possible AD is also entertained.

• A high vascular load on structural MRI does not exclude the concomitant presence of AD. In a study of 45 cases with subcortical vascular disease in the absence of large-vessel stroke or macrohemorrhage, approximately one third of cases were 11C-PIB positive (Lee et al., 2011). The amount of white matter hyperintensities correlates with 11C-PIB based measures of Aβ load in patients with clinically probable CAA but not in AD (Gurrol et al., 2012).

• Cases with a long-standing psychiatric history of depression may develop Alzheimer’s disease and it may be impossible to make a diagnosis of clinically probable AD under such conditions. A false-negative diagnosis will lead to poor caregiver comprehension of the drastic changes that occur over the more recent period of time and it is important to disentangle the psychiatric origin from any acquired neurodegenerative contribution. A false-positive diagnosis would also have serious consequences in this vulnerable population.

• Medical comorbidity such as chronic renal insufficiency, heart failure or obstructive pulmonary disease, to the degree that the presence of AD would impact on patient management in such cases (e.g. candidates for transplantation surgery).

• Normal-pressure hydrocephalus: Among 433 patients with a clinical diagnosis of NPH based on neurological and radiological criteria, half of whom received a ventriculoperitoneal shunt, 20% showed cognitive decline over the following years. The presence of Aβ alone or in combination with hyperphosphorylated tau at the time of shunting was a strong predictor for subsequent cognitive decline. In such cases amyloid imaging may increase preoperative diagnostic accuracy and contribute to the decision making with regards to a surgical intervention (Leinonen et al., 2010).

4. Rapidly progressive cognitive deterioration: There is no strict definition of ‘rapidly progressive AD’: an interval from clinical onset to death of less than two years would be a reasonable definition but is essentially retrospective. Besides rapidly progressive AD, the differential diagnosis of rapidly progressive dementia...
includes conditions such as Lewy body dementia and Creutzfeldt–Jakob but also potentially reversible causes such as paraneoplastic or non-paraneoplastic autoimmune encephalitis, central nervous system lymphoma or angiitis. To reach a high level of confidence that reversible causes of rapid progression are excluded, hospitalization and costly diagnostic investigations will often be required. Evidence for the presence or absence of cerebral amyloidosis may help direct the investigations. Of course, under such dramatic conditions, biomarkers for cerebral amyloidosis will not replace other tests such as diffusion-weighted MRI or cerebrospinal fluid assays. The added value of biomarkers for cerebral amyloidosis in rapidly progressive dementia is an important topic for further investigation which will probably require a multicenter approach given the relative rarity of this syndrome.

Under such conditions the clinical utility of amyloid imaging has to be weighed against the clinical utility of other tools such as FDG-PET and CSF biomarkers. Structural MRI will likely retain its place in the diagnostic work-up, for exclusion of neoplastic lesions (basal forebrain tumors, non-Hodgkin lymphoma, low-grade glioma of medial temporal cortex…) and for assessment of the vascular lesion load. A high vascular load may direct diagnostics and therapeutics into the vascular risk prevention strategies. Also, if disease-modifying drugs are introduced into the clinic, MRI Gradient Echo sequence may allow to detect cerebral amyloid angiopathy which is associated with a higher risk for Amyloid-Related Imaging Abnormalities (Sperling et al., 2011), as seen in most passive immunization trials.

Under the conditions outlined above, the benefit of amyloid imaging is a benefit in terms of diagnostic clarity. This should not be underestimated as under the aforementioned conditions, accurate diagnosis determines patient management. It is also one of the principal questions by patients and caregivers to the physician. It is hard to quantify this benefit by monetary values and the benefit of accurate diagnosis early in the disease course is a complex issue. Furthermore, with most other diagnostics (e.g. $^{123}$I β-CIT SPECT in Parkinson’s disease), there is a clear cost-effectiveness based on appropriate therapy choices, which is difficult for dementia when therapeutic options are still limited.

8. The multidimensional space of Alzheimer’s disease

According to one of the most influential current AD models (Jack et al., 2010, 2013), CSF alterations in Aβ42, amyloid aggregation, MRI volume loss and cognitive decline follow an orderly temporal sequence, initiated by changes in amyloid and culminating in the clinical expression of cognitive symptoms. This model has a high heuristic value and has engendered a number of testable hypotheses from which a clearer picture of the time-dependent trajectories of AD biomarkers relative to clinical disease stage and to each other will be derived (Jack et al., 2010). This model could be viewed as a translation of the amyloid cascade hypothesis (Hardy and Selkoe, 2002) into a sequence of changes that are measurable in humans in vivo. Inherent to the amyloid cascade hypothesis is a linear sequence of events that is initiated by an amyloid related change.

Alternatives exist to such linear sequential models. For instance, the different in vivo imaging modalities may reflect partly independent events. According to that scenario, AD would have to be depicted in a multidimensional space where different measures co-determine cognitive outcome rather than as a linear sequence from Aβ42 to cognitive symptoms. The term ‘multidimensional space’ emphasizes the partial independence of the different in vivo measures of Alzheimer’s disease. It refers to the multidimensional scaling of sets of parameters that is often used when one wants to discern separate clusters within complex datasets. The multidimensional space of Alzheimer’s disease relates to the concept of multifactorial AD but remains more at the level of description of the complex relationship between different features while the term multifactorial refers more to the factors that contribute to AD, such as neuropathological core lesions or concomitant Lewy bodies (see above).

A multidimensional view fits with recent genetic evidence in sporadic AD (Lambert et al., 2009; Sleegers et al., 2010): AD can be regarded as a network of multiple pathways with multiple hubs that may separately and interactively result in a given behavioral phenotype. The multidimensional model incorporates the partial independence between processes that we can measure in vivo by means of the different techniques, in particular between amyloid ligand retention levels and volume loss (Driscoll et al., 2012; La Joie et al., 2012). There is also neuropathological evidence for partial independence: Synapse loss in dentate gyrus does not correlate with the amount of NFT (Scheff et al., 2006) and volume loss does not correlate with amyloid load or NFT (Driscoll et al., 2012). Such multidimensional models resonate with the multifactorial model of AD (Savva et al., 2009) and are distinct from an amyloidocentric linear sequential model. The way in which different measures combine may give rise to clusters or patterns that characterize a specific subtype of patients. The multidimensional space therefore not only allows for partial independence between measures but also for inter-individual heterogeneity.

The distinction between a linear sequential model and a multidimensional model hinges on the relative independence of the different processes that can be imaged with current techniques. At the group level the evidence for a probabilistic relationship between brain volume loss, presence of amyloid aggregates, and zones of hypometabolism is strong. At the regional and at the individual level, the interdependence between these measures may be more variable. The distinction between a multidimensional and parallel model of AD versus a linear model bears resemblance to the distinction between serial box-and-arrow models of cognitive function and parallel distributed models. As is the case in such neuropsychological models, linear models have a strong heuristic advantage as they engender clearly defined hypotheses that can be empirically tested. Such models have contributed a lot to our understanding of how cognitive processes are related to each other. Multidimensional models however may allow one to capture the complexity of the brain’s operations in a more truthful manner. Empirical data obtained from multimodal longitudinal studies will have to determine which of the two types of models is correct.

9. Conclusion

Amyloid imaging provides a direct window on one of the components of AD. This component stands in a relatively complex relationship to a diversity of other component processes. Only some of these processes can be imaged and there is an ongoing search for techniques to detect and quantify in vivo some of the other pathogenetic mechanisms, such as glial cell involvement (Cagnin et al., 2001) or tau hyperphosphorylation (Fodero-Tavolletti et al., 2011; Okamura et al., 2005). Potential key early aspects of AD pathophysiology, such as Aβ oligomers, remain below the radar of current biomarker detection capabilities (Jack et al., 2013). Amyloid imaging provides hope for progress as it allows for direct measurement of one component contributing to AD-related cognitive decline. This can enhance the chance of success of trials by allowing to restrict inclusion to those patients who have the target of interest and by showing whether the drug engages the target or not. While amyloid imaging is already being introduced to the market, many gaps can be identified in our evidence-based medical knowledge of its role in clinical practice. These gaps will have to be filled over the years to come by studies of clinical utility and added value.

Acknowledgments

Supported by FWO grants G.0076.02 (R.V.), KU Leuven Research grants OT/08/056 and OT/12/097 (R.V.), Federaal Wetenschapsbeleid belslo Inter-University Attraction Pole P6/29 and P7/11, and Stichting Alzheimer Onderzoek grant 11020, R.V. and K.V.L. are senior clinical
investigators of the Fund for Scientific Research, Flanders (FWO), and K.A. is a doctoral research fellow of the Fund for Scientific Research, Flanders (FWO).

References
Aggleton, J.P., Friedman, D.P., Mishkin, M., 1987. A comparison between the connections of the amygdala and hippocampus with the basal forebrain in the macaque. Experimental Brain Research 67, 556–568.
Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gambelas, S., Holtzman, D.M., Jack, W.J., Petersen, R.C., et al., 2011. The enigma of amyloid detection in Alzheimer’s disease: recommendations from the national institute on aging-Alzheimer’s association workgroups on diagnostic guidelines for Alzheimer’s disease. Alzheimers & Dementia 7, 279–295.
Arráigada, P.V., Crowdon, J.H., Hedley-Whyte, E.T., Hyman, B.T., 1992. Neurofibillary tangles but not senile plaques parallel duration and severity of Alzheimer’s disease. Neurology 42, 631–639.
Backsai, B.J., Frosch, M.P., Freeman, S.H., Raymond, S.B., Augustinack, J.C., Johnson, K.A., Irizarry, M.C., Klunk, W.E., Mathis, C.A., DeKosky, S.T., Greenberg, S.M., Hyman, B.T., Crowdon, J.H., 2007. Molecular imaging with Pittsburgh Compound B confirmed at autopsy: a case report. Archives of Neurolgy 64, 431–433.
Baroni, C., Braak, H., Fischer, P., Jellinger, K.A., 1999. Neuropathological staging of Alzheimer lesions and intellectual status in Alzheimer’s and Parkinson’s disease patients. Neuroscience Letters 162, 179–182.
Baron, J.C., Chételat, G., Desgranges, B., Perchey, G., Landeau, B., de la Sayette, V., Eustache, F., 2001. In vivo mapping of gray matter loss with voxel-based morphometry in mild Alzheimer’s disease. NeuroImage 14, 298–309.
Barthel, H., Gertz, H.-J., Dresel, S., Peters, O., Hauenstein, V., Liu, X., van Heurten, R., Mollenhauer, B., 2009. Pittsburgh B florbetapir compared with neocortical amyloid imaging in Alzheimer’s disease: a prospective cohort study. Lancet Neurology 11, 669–678.
Cselényi, Z., Jönghagen, M.E., Forsberg, A., Hallidin, C., Julin, P., Schou, M., Johnström, P., Varnäs, K., Svensson, S., Farde, L., 2012. Amyloid-β-specific PET radioligand. Journal of Nuclear Medicine 415–424.
Davis, D.G., Schmitt, F.A., Weiskel, D.R., Markesbery, W.R., 1999. Alzheimer neurban changing in aged cognitively normal subjects. Journal of Neuropathology and Experimental Neurology 58, 376–388.
DeKosky, S.T., Scheff, S.W., 1990. Synapse loss in frontal cortex biopsies in Alzheimer’s disease: correlation with cognitive severity. Annals of Neurology 27, 457–464.
Devanand, D.P., Mikno, A., Petlon, G.H., Cuasay, K., Pradhaban, G., Kumar, J.S., Upadhyay, L., Rai, G., Hyman, B.T., Jack, W.J., Liu, X., van Heurten, R., Mollenhauer, B., 2010. Pittsburgh B Florbetapir PET and fluorodeoxyglucose (18F-FDG) PET in patients with Alzheimer disease, mild cognitive impairment, and healthy controls. Journal of Geriatric Psychiatry and Neurology 23, 185–193.
Dhollander, A., Neillen, N.V., De peppere, P., De Peutter, P., van Paesschen, W., Thijs, V., Vandenbunde, R., 2011. In vivo amyloid imaging in cortical superficial siderosis. Journal of Neuroimaging and Neurosurgery, Psychiatry 82, 469–471.
Duffy, D., Bakkour, A., Sadowsky, C.M., Pacheco, J.A., DePeters, S., Wright, C.L., Black, D., Rosand, W.H., Sperling, R.A., Atri, A., Crowdon, J.H., Hyman, B.T., Morris, J.C., Fischl, B., Buckner, R.L., 2009. The cortical signature of Alzheimer’s disease: regionally specific cortical thinning relates to symptom severity in very mild to mild ad dementia and is detectable in asymptomatic amyloid-positive individuals. Cerebral Cortex 19, 497–510.
Dierssen, G.A., Skene, M.H., Khan, A., Jeng, J., Nandigam, R.N.K., Becker, J.A., Kumar, A., Neal, K.L., Betensky, R.A., Frosch, M.P., Rosand, J., Johnson, K.A., Wiswanathan, A., Sahakian, B.J., Greenberg, S.M., Hyman, B.T., 2004. Microbleeds and amyloid deposits in amyloid angiopathy. Annals of Neurology 68, 545–548.
Dorsainvil, M., Sperling, R.A., Coleman, R., Johnson, K.A., Reiman, E.M., Graff-Radford, N.R., Morris, J.C., Grundman, M., Sabbagh, M.N., Sadosky, C.H., Filler, A.S., Carpenter, A., Clark, C.M., Jozefiak, T., Mintun, M.A., Skovronsky, D.M., Ponce, R.C., for the AV45-A07 group, 2012. Amyloid-β-assessed by florbetapir 18F PET and 18-month cognitive decline: a multicenter study. Neurology 79, 1636–1644.
Drsicoll, I., Zhou, Y., An, Y., Sojkova, J., Davatzikos, C., Kraut, M.A., Ye, W., Ferrucci, L., Mathis, C.A., Klunk, W.E., Wong, D.F., Reinsch, S.M., 2011. Lack of association between 11C-PiB and longitudinal brain atrophy in non-demented older individuals. Neurobiology of Aging 32, 2123–2130.
Drsicoll, I., Troncoso, J.C., Rudow, G., Sojkova, P., Chudnovsky, O., Zhou, Y., Cran, M.A., Furlow, D.L., Mathis, C.A., Knudtson, M.E., Reiman, E.M., Doraiswamy, P.M., O’Keefe, J.E., Ferrucci, L., Reinsch, S.M., 2012. Correspondence between in vivo 11C-PiB amyloid imaging and postmortem, region-matched assessment of plaques. Acta Neuropathologica 124, 831–831.
Duara, R., Loewenstein, D.A., Shon, Q., Barler, W., Potter, E., Don, V., Heurin, K., Rand, B., 2012. Amyloid positron emission tomography with 18F-flutemetamol and structural magnetic resonance imaging in the classification of mild cognitive impairment and Alzheimer’s disease. Alzheimers & Dementia. http://dx.doi.org/10.1016/j.jalz.2012.01.006.
Edison, P., Rowe, C.C., Rinne, J.O., Ng, S., Ahmed, I., Kemppainen, N., Villemagne, V.L., O’Keefe, J., Nagren, K., Chaudhury, K.R., Masters, C.L., Brooks, D.J., 2008. Amyloid load in Parkinson’s disease dementia and lewy body dementia measured with 11C-PiB PET emission tomography. Journal of Neurology, Neurosurgery, and Psychiatry 79, 1331–1338.
Engler, H., Forberg, A., Almkvist, O., Blomqvist, G., Larsson, E., Savicevich, I., Wall, A., Reiman, E.M., Langström, B., 2006. Amyloid detection in patients with Alzheimer’s disease: Brain 2389, 2865–2866.
Fagin, A.M., Mintun, M.A., Mach, R.H., Lee, S.Y., Dence, C.S., Shah, A.R., LaRoossa, G., Spinner, M.L., Klunk, W.E., Mathis, C.A., DeKosky, S.T., Morris, J.C., Holtzman, D.M., 2006. Inverse relation between in vivo amyloid imaging load and cerebral fluid beta42 in humans. Annals of Neurology 59, 512–519.
Fodero-Tavoletti, M.T., Okamura, N., Furumoto, S., Mulligan, R.S., Connor, A.R., McLean, C.A., Cao, D., Rigopoulos, A., Cartwright, C.A., O’Keefe, G., Gung, A., Salland, P., Barcelo, H., Foster, K.J., Rowe, C.C., Hong, J., Kudo, Y., Yonai, T., Enomoto, K., Vinters, H.V., 2011. 18F-Tik523: a novel in vivo tau imaging ligand for Alzheimer’s disease. Brain 134, 1089–1100.
Foster, N.L., Heidebrink, J.L., Clark, C.M., Jagust, W.J., Arnold, S.E., Babars, N.R., DeCarli, C.S., Turner, R.S., Koeppa, R.A., Higdon, R., Minoshima, S., 2007. FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer’s disease. Brain 130, 2616–2635.
