Clinical utility of FDG-PET for the clinical diagnosis in MCI
Javier Arbizu*, MD, PhD; Cristina Festari, MS; Daniele Altomare, MS; Zuzana Walker, MD; Femke Bouwman, MD; Jasmine Rivolta, MS; Stefania Orini, MD; Henryk Barthel, PhD; Federica Agosta, PhD; Alexander Drzezga, MD; Peter Nestor, MD; Marina Boccardi*, PhD; Giovanni Battista Frisoni, MD; Flavio Nobili, MD; for the EANM-EAN Task Force for the Prescription of FDG-PET for Dementing Neurodegenerative Disorders.

AFFILIATIONS
Cristina Festari: LANE – Laboratory of Alzheimer’s Neuroimaging & Epidemiology, IRCCS S. Giovanni di Dio, Fatebenefratelli, Brescia, Italy; Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy.
Daniele Altomare: LANE – Laboratory of Alzheimer’s Neuroimaging & Epidemiology, IRCCS S. Giovanni di Dio, Fatebenefratelli, Brescia, Italy; Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy.
Zuzana Walker: University College London, Division of Psychiatry & Essex Partnership University NHS Foundation Trust, UK.
Femke Bouwman: Department of Neurology & Alzheimer Center, Amsterdam Neuroscience, VU University Medical Center, Amsterdam, the Netherlands.
Jasmine Rivolta: LANE – Laboratory of Alzheimer’s Neuroimaging & Epidemiology, IRCCS S. Giovanni di Dio, Fatebenefratelli, Brescia, Italy.
Stefania Orini: Alzheimer Operative Unit, IRCCS S. Giovanni di Dio, Fatebenefratelli, Brescia, Italy.
Henryk Barthel: Department of Nuclear Medicine, Leipzig University Hospital, Leipzig, Germany
Federica Agosta: Neuroimaging Research Unit, Institute of Experimental Neurology, Division of Neuroscience, San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy.
Peter Nestor: German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany; Queensland Brain Institute, University of Queensland and at the Mater Hospital Brisbane.
Alexander Drzezga: Department of Nuclear Medicine, University Hospital of Cologne, University of Cologne and German Center for Neurodegenerative Diseases (DZNE), Germany.
Marina Boccardi: LANVIE (Laboratoire de Neuroimagerie du Vieillissement), Dept of Psychiatry, University of Geneva, Geneva, Switzerland; LANE – Laboratory of Alzheimer’s Neuroimaging & Epidemiology, IRCCS S. Giovanni di Dio Fatebenefratelli, Brescia, Italy.
Giovanni Battista Frisoni: LANVIE (Laboratoire de Neuroimagerie du Vieillissement), Dept of Psychiatry, University of Geneva, Geneva, Switzerland; Memory Clinic, University Hospitals, Geneva, Switzerland; LANE – Laboratory of Alzheimer’s Neuroimaging & Epidemiology, IRCCS S. Giovanni di Dio Fatebenefratelli, Brescia, Italy.

Flavio Nobili: Dept. of Neuroscience (DINOGMI) University of Genoa and Polyclinic San Martino Hospital, Genoa, Italy.

*Correspondents:
Javier Arbizu
Department of Nuclear Medicine. Clinica Universidad de Navarra,
University of Navarra,
Pamplona, Spain –
jarbizu@unav.es.

Marina Boccardi
LANVIE (Laboratoire de Neuroimagerie du Vieillissement), Dept of Psychiatry
University of Geneva, Geneva, Switzerland
e-mail: marina.boccardi@unige.ch
Tel.: 0041.(0)22.3055764
Fax.: 0041.(0)22.3054719
ABSTRACT

Purpose
We aim to report the quality of accuracy studies investigating the utility of [18F]fluorodeoxyglucose (FDG)-PET in supporting the diagnosis of prodromal Alzheimer’s Disease (AD), frontotemporal lobar degeneration (FTLD) and prodromal dementia with Lewy bodies (DLB) in mild cognitive impairment (MCI) subjects, and the corresponding recommendations made by a panel of experts.

Methods
Seven panellists, four from the European Association of Nuclear Medicine, and three from the European Academy of Neurology, produced recommendations taking into consideration the incremental value of FDG-PET, as added on clinical-neuropsychological examination, to ascertain the aetiology of MCI (AD, FTLD or DLB). A literature search using harmonized population, intervention, comparison, and outcome (PICO) strings was performed, and an evidence assessment consistent with the European Federation of Neurological Societies guidance was provided. The consensual recommendation was achieved based on Delphi rounds.

Results
Fifty-four papers reported the comparison of interest. The selected papers allowed the identification of FDG patterns that characterized MCI due to AD, FTLD and DLB. While clinical outcome studies supporting the diagnosis of MCI due to AD showed varying accuracies (ranging from 58 to 100%) and varying areas under the receiver-operator characteristic curves (0.66 to 0.97), no respective data were identified for MCI due to FTLD or for MCI due to DLB. However, the high negative predictive value of FDG-PET and the existence of different disease-specific patterns of hypometabolism support the consensus recommendations for the clinical use of this imaging technique in MCI subjects.

Conclusions
FDG-PET has clinical utility on a fair level of evidence in detecting MCI due to AD. Although promising also in detecting MCI due to FTLD and MCI due to DLB, more research is needed to ultimately judge on the clinical utility of FDG-PET in these entities.

Keywords
FDG-PET
MCI
Differential diagnosis
Alzheimer’s disease
Dementia with Lewy bodies
Frontotemporal lobar degeneration
1. BACKGROUND

The absence of international guidelines for the clinical use of brain $^{18}$F-fluorodeoxyglucose (FDG)-PET in the diagnosis of dementia motivated the European Association of Nuclear Medicine (EANM) and the European Academy of Neurology (EAN) to launch a joint initiative providing guidance to clinicians on the indications for the exam. The initiative included a set of 21 clinical questions to be addressed based on literature evidence and expert consensus [1].

In this paper, we report the assessment of accuracy studies which investigated the utility of brain PET by means of the glucose consumption tracer FDG in supporting the diagnosis of prodromal Alzheimer’s Disease (AD), prodromal frontotemporal lobar degeneration (FTLD) and prodromal dementia with Lewy bodies (DLB) in mild cognitive impairment (MCI) subjects. Moreover, we report the reasons for the decisions of the panel in supporting the use of the exam in the diagnosis of subjects with MCI.

Despite the current lack of effective treatments that reverse or stop the course of cognitive impairment due to AD, available treatments provide greater benefits in mild compared to moderate AD [2]. In addition, the early knowledge of an accurate diagnosis has an additional value to positively modify risk factors aggravating the effects of neurodegenerative conditions (e.g. cardiovascular), and to perform rest of life planning. Consequently, the search for useful diagnostic criteria moves towards the earliest manifestations of the disease.

Subjects with MCI represent an important clinical group as they are at increased risk of developing dementia. Nevertheless, MCI is a transitional syndrome that has heterogeneous presentations and aetiologies. Most subjects with amnestic MCI have AD pathology (prodromal AD), while one quarter or less have hippocampal sclerosis, FLTD or DLB pathology, and the rest have normal age-related changes, cerebrovascular disease or depression[3–5]. Therefore, not all individuals with MCI will develop AD dementia and, in fact, there are patients whose cognitive problems convert to other types of dementia (e.g. FTLD, DLB, vascular dementia), remain stable, and may even revert to normality.

In the last few years, specific cognitive impairment and neuropsychiatric profiles defined as prodromal DLB and FTLD have also been described in MCI subjects. However, the early clinical expression of these conditions is not as well understood as that of prodromal AD. The term mild behavioural impairment or prodromal behavioural variant frontotemporal dementia (bvFTD) has been adopted for patients with cognitive and/or behavioural impairment not fulfilling bvFTD criteria, lacking deficits in activities of daily living in spite of frontotemporal grey matter atrophy[6]. Besides, the clinical diagnosis of very early primary progressive aphasia (PPA) or MCI stage of PPA is usually based on the presence of mild but persistent isolated difficulty on tests of language (frequently
dissociated from one patient to another), with relative preservation of other cognitive domains and activities of daily living [7]. On the other hand, prodromal DLB is a complex disease to diagnose because of different clinical onsets and variable clinical trajectories. DLB-MCI subjects can exhibit low visuospatial function and letter fluency tests with increase mean UPDRS score, REM sleep behaviour disorder and fluctuating cognition.

In this framework, neuroimaging and specifically FDG-PET (currently included in the proposed clinical diagnostic criteria of all these conditions) might help clinicians to achieve an early diagnosis, mandatory for prognosis, treatment and patient management, and for improving subject selection for clinical trials based on prodromal stages.

FDG is the most available PET radiotracer. Further, National Health services, social security or private insurances provide reimbursement of brain FDG-PET for the evaluation of progressive neurodegenerative disorders associated with dementia in many European countries (e.g. France, Italy, Netherlands, Spain, Switzerland, UK)[5]. However, specific indications for clinical prescription of FDG-PET in these conditions vary from one country to another: FDG-PET is accepted for unclear cases (e.g. Sweden, Switzerland), in unexplained dementia or atypical presentations (e.g. Netherlands, France), or for differential diagnosis of AD. Interestingly, early diagnosis of AD is expressly included as a clinical prescription only in France[5].

Based on this background, literature searches were performed to assess the quality of the evidence on the efficiency of FDG-PET to support the diagnosis of prodromal AD, FTD, and DLB in subjects with progressive MCI of uncertain origin. The output of these assessments allowed panellists to decide on the applicability of FDG-PET via a Delphi procedure.

2. METHODS

Seven panellists, four from EANM and three from EAN, were appointed to develop recommendations on the clinical utility of FDG-PET, as added to clinical-neuropsychological examination, to ascertain the aetiology of MCI. By that, consensus recommendations were produced through a Delphi procedure based on the expertise of panellists. In this process, the panellists were also informed about the availability and quality of evidence as assessed by an independent methodology group[8] (in this issue).

Briefly, we performed literature searches using harmonized population, intervention, comparison, and outcome (PICO) question keywords edited by the experts, screened the studies for eligibility, extracted the data to assess their methodological quality, and provided an evidence
assessment consistent with the European Federation of Neurological Societies guidance [9] which was adapted to FDG-PET studies[8].

2.1 PICO questions
For this review, the PICO questions were whether FDG-PET should be performed, as adding diagnostic value (in terms of increased accuracy, and versus pathology, biomarker-based diagnosis or diagnosis at follow-up) as compared to standard clinical/neuropsychological assessment alone, to detect prodromal AD (PICO 1), prodromal FTLD (PICO 2), and prodromal DLB (PICO 3) in subjects with persistent MCI of uncertain origin.

2.2 Eligibility criteria
Only original full papers published in English on international impacted journals were considered, excluding reviews, management guidelines, abstracts and grey literature. Any sample size was allowed if pathology was the gold standard for diagnosis. Otherwise, in cases in which clinical follow-up was utilized as gold standard, only publications with a minimal sample size of 15 MCI subjects were considered.

2.3 Literature search
An electronic search strategy, developed and tested by the panellists, was performed through predefined keywords strings, grounding on the specific PICO question and including a selection of terms taken from a largely inclusive literature selection, in order to pick all variants for the same keyword [8].

Literature searches were performed using Medline and Embase databases, until December 2015. In reporting the findings of this review, we adhered to the standards of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [10]. An initial independent screening of all included studies was performed either by a neurologist or by a nuclear medicine physician with expertise in neurodegenerative dementing disorders, who could include additional papers based on personal knowledge or tracking from references of papers. The full texts of these potentially eligible studies were then independently assessed for eligibility by a methodological team.

2.4 Data extraction and quality assessment
Data extractors for this review were CF for PICO 1 and 2; JR for PICO 3. The quality of evidence was assessed consensually within the methodology group based on study design, gold/reference standard, FDG-PET image assessment (visual or semi-quantitative methods), risk of bias, index test
imprecision, applicability in memory clinics, effect size, and effect inconsistency. Critical outcomes were validated measures of FDG-PET performance (accuracy, sensitivity, specificity, AUC, positive and negative predictive values and likelihood ratios) against the above-discussed gold standard measures. A final assessment of relative availability of evidence was formulated, taking into account evidence availability from all PICO answers. This ranking was summarized as very poor/lacking, poor, fair or good[8].

In our terminology, we considered clinical syndromes separate from pathophysiologies, consistent with current NIA-AA and IWG criteria, [11–13]. Unless differently referenced, [14], we adopted the inclusive 1998 definition of frontotemporal lobar degeneration[15], treating separately only the linguistic variant when specified.

3. RESULTS

For the 3 PICOs included in this review, 184 papers were identified by panellists, and 54 reported the comparison of interest and were examined by the methodology group. Of these, 13 did contain the critical outcomes, properly quantifying FDG-PET diagnostic utility and all of these related to PICO 1 (diagnosis of AD in MCI; Figure 1). The evidence assessment denoted that these studies provided fair evidence for the utility of FDG-PET to support the diagnosis of AD in MCI subjects. Based on the relevance of the negative predictive value of FDG-PET, and on the different patterns of hypometabolism, (Figure 2) panellists supported the clinical use of the exam in the examined clinical scenarios, notwithstanding the lack of quantitative evidence (Table 1, Nobili et al., n.d.)

3.1 PICO 01: Use of FDG-PET to support diagnosis of AD in MCI subjects

Among the 74 papers identified and screened by the referent panellist (JA), 35 were sent to the methodological group for data extraction and assessment (see Figure 1 - PICO 1). Fourteen papers were excluded for the following reasons: 3 were cross-sectional studies, without information about subsequent conversion (or lack thereof) of MCI subjects[16–18]; 3 papers did not include the population of interest[19–21]; 3 studies correlated cognitive profiles with metabolic pattern [22–24]; Pagani et al. [25] compared MCI subjects and healthy controls; and Prestia et al. [26] assessed the accuracy of combinations of biomarkers in predicting decline. Lastly, 3 papers were meta-analyses or reviews [27–29]. As to the remaining 21 appropriate papers, 13 reported the critical outcomes for this PICO and 8 described only associative results [30–37] (see Table PICO 1; data extraction table available at (https://drive.google.com/file/d/0B0_JB3wzTvbpRXNkaXBjMUlLYU0/view?usp=sharing). On the
whole, these papers[38–48] obtained a large range of sensitivity (38-98%), specificity (41-97%) and accuracy (58-100%) values. This variability involved also the other types of critical outcomes: AUC ranged from 0.66 to 0.97[38,41–44,48,49] PPV: 41-85.2%; NPV: 77-95% [38,41,43–45,50]. The only paper reporting LR values obtained an LR+ = 8.14 and LR- = 0.12[45]. The pattern of hypometabolism that characterized MCI converting to AD mainly included posterior cingulate and posterior temporo-parietal areas (Figure 2). Besides limitations in applicability, due to the non-systematic use of semiquantitative assessment in clinical centres, we did not detect substantial risks of bias in the examined studies. Nonetheless, the variability in the values of critical outcomes does not allow to draw formal conclusions on the clinical utility of the exam for this PICO. However, taking into account the availability of formal evidence for all of the PICOs within the entire project, the relative availability of evidence supporting the clinical utility of FDG-PET in detecting AD in MCI subjects was ultimately been considered as fair. The consensual recommendation for PICO 1 was defined on Delphi round II, with 6 of 7 panelists voting for a clinical use of FDG-PET mainly based on its high negative predictive value, along with its typical posterior hypometabolic pattern.

3.2 PICO 02: Use of FDG-PET to support the diagnosis of FTLD in MCI subjects

Thirty-six papers were identified and screened by the referent panellist (FB), only 10 qualifying for further analysis and being sent to the methodology team (see Figure 1 - PICO 2). Of these, six papers were excluded for the following reasons. Mosconi et al. [16], Poljansky et al.[51] and Banzo et al.[52] were cross-sectional studies comparing FDG-PET in MCI of uncertain origin versus FTLD patients, whose clinical stage was dementia; Bergeron et al., [53] did not include the target population; Morbelli et al. [54] did not report diagnostic accuracy. Finally, Döbert et al.,[55] did not reach the minimum sample size, nor reported validated measures of test performance. The data extraction table for this PICO is available at (https://drive.google.com/file/d/0B0_JB3wTvbpWnFKNmZMcZFpaVU/view?usp=sharing).

For this PICO, the critical outcomes specific to the diagnosis of FTLD in MCI subjects were not available in any of the examined papers. This means a lack of any objective evidence supporting an incremental diagnostic value of FDG-PET for the diagnosis of FTLD in MCI. However, some papers were nonetheless of potential use to inform the panellists decisions. In Perani et al.[56], semi-quantitative assessment of FDG-PET images correctly identified at baseline the only 2 MCI subjects that later converted to FTLD. Grimmer et al. [57] found a visual rating accuracy in detecting conversion to dementia of different causes ranging from 64 to 68%. An SPM analysis reached a high accuracy in the early identification of converter MCI[58] and had a high predictive value in
identifying hypometabolic patterns in different neurodegenerative dementias and in MCI [59]. However, these data were not specific to the FTLD diagnosis.

Based on these data, the availability of formal evidence on a clinical utility of FDG-PET to support the diagnosis of FTLD in MCI subjects was considered to be lacking. However, the consensual recommendation was defined on Delphi Round II with 6 of 7 panellists deciding to recommend clinical use, based on the high negative predictive value in the above-cited study and the known typical metabolic pattern in FTLD already present at MCI stage (Figure 2), mainly including hypometabolism in at least one of the several aspects of the frontal lobes, the anterior temporal lobe, anterior cingulate gyrus, insula, amygdala and caudate nuclei.

3.3 PICO 03: FDG-PET to PET to support diagnosis of DLB in MCI

Among the 74 papers identified by the referent panellist (ZW), 9 were eligible for further assessment and sent to the methodology team (see Figure 1 - PICO 3). Among these, only two papers contained the comparison of interest. Cerami C. et al. [58], Fujishiro H. et al. [60], Perani D. et al. [61] Banzo I. et al.[52], Fujishiro H. et al.[62], Fujishiro H. et al. [63] and Pardo J.V. et al. [64] did not include the target sample or the requested sample size. The data extraction table is available at (https://drive.google.com/file/d/0B0_JB3wzTvbpMC1uM0JIblRGYUk/view?usp=sharing).

The two studies investigating the comparison of interest included 37 MCI subjects. They reported hypometabolic patterns for six subjects at the MCI stage, who progressed to a clinical diagnosis of DLB, confirmed at autopsy in one case. Clerici et al.[65] described 30 subjects suffering from amnestic or non-amnestic MCI at baseline, and found hypometabolism in the posterior and anterior cingulate gyrus and in the parietal lobe in the six subjects who converted to DLB. The pattern of hypometabolism described in Mosconi et al. [66] was similar to that of AD. At pathology, only 2 out of the 7 patients with clinical diagnosis of DLB had Lewy bodies (Figure 2).

Relative to the evidence available for PICO 1, the availability of formal evidence supporting diagnostic utility of FDG-PET for prodromal DLB was lacking. Nevertheless, the consensual recommendation was defined on Delphi round IV, with 5 of 7 panellists deciding to recommend its clinical use, mainly based on the potential clinical utility of the typical finding of hypometabolism in occipital cortex.

4. DISCUSSION

In this paper, we assessed the existing evidence on the clinical utility of FDG-PET in supporting the diagnoses of prodromal AD, FTLD and DLB in MCI subjects. We found fair availability of evidence
on a utility to support the diagnosis of AD in MCI subjects, while evidence for FTLD and DLB was lacking.

Some similar work has been published in the last few years. Perani D et al.,[67] performed a meta-analysis and GRADE analysis revealing positive critical outcomes for PET imaging (FDG- and amyloid-PET) that might be useful in considering recommendations to improve the early diagnosis of AD and differential diagnosis of dementia. They highlighted the additional value of parametric or voxel-wise analyses by comparing the subject's scan with a normative data set, which significantly increases the diagnostic accuracy of visual analysis largely used in most published series. On the other hand, Smailagic N,[68] et al. focused in the review publication on the early diagnosis of AD dementia and other dementias in MCI subjects using FDG-PET. They assessed the methodological quality of each study using the QUADAS-2 tool, as recommended by The Cochrane Collaboration. The set of papers identified for this analysis mostly overlapped with the papers we selected as eligible for PICO 1 in our study and, similar to our results, they also found large variability in diagnostic performance when using FDG-PET to support the diagnosis of AD in MCI subjects. However, our study differed from the study by Smailagic et al regarding conclusions. The Cochrane review concluded that a positive value of FDG-PET in people with MCI for prediction of developing AD dementia cannot be stated. They based this conclusion on results’ inconsistencies, on the non-specific nature of FDG and the allegedly inherent drawback of low spatial resolution of PET. Morbelli S, et al. on behalf of the European Association of Nuclear Medicine and Molecular Imaging (EANMMI), argued against the categorical negative conclusion of the Cochrane review by claiming that variability is not exclusively attributable to the method itself. More recently, Garibotto et al evaluated the validity of FDG-PET in diagnosing AD, in the prodromal stage by following the oncology framework adapted to the field of neurodegenerative disorders.[69] Although good analytical validity is demonstrated for FDG-PET to identify the AD in its earliest stage (prodromal AD) (demonstrated in phase 1–2 studies), they found that the integration of FDG-PET as a useful diagnostic marker in clinical practice, although potentially useful, still requires demonstration of clinical validity and utility.

The reasons proposed during the Delphi panel discussion to support the use of FDG-PET in the diagnosis of prodromal AD, FTLD and DLB in MCI subjects mainly relied on the patterns of hypometabolism reported in the available FDG-PET studies, and the importance of negative findings that exclude the presence of neurodegeneration. In particular, the pattern of hypometabolism in the posterior cingulate and posterior temporo-parietal areas that characterize MCI converting to AD was considered helpful in the diagnosis of AD in MCI subjects (PICO 1). An MCI constellation is challenging if diagnosed solely on clinical ground with regard to outcome prediction, because
declining memory is also a feature of normal aging, and some MCI cases may never progress to the dementia stage, or even reverse to normality. Therefore, one of the main strengths of FDG-PET over other biomarkers (i.e. amyloid imaging or CSF) lies on its high predictive value for short-term conversion to AD in MCI subjects, in turn offering clinically relevant prognostic information. Although a formal comparison with such biomarkers was not performed and is necessary to this regard, we anyway underline the informative value of hypometabolism in denoting the progression of possibly silent pathophysiology, consistent with current diagnostic criteria[70,71]. In addition, the different metabolic patterns that can be detected in MCI subjects may be useful in identifying non-Alzheimer types of neurodegeneration early in the course of the disease.

The diagnosis of FTLD in MCI subjects (PICO 2) may be challenging, especially early in the disease course. In the early phases, in fact, behavioural changes may be subtle, while cognitive impairment may be absent (and indeed mild behavioural -rather than cognitive- impairment may be a better denomination for the prodromal bvFTLD). Functional neuroimaging with FDG-PET has been included in the most recent diagnostic criteria for probable bv-FTD, in patients with dementia [14]. However, due to its sensitivity to synaptic failure even in early phases of neurodegeneration, the detection of frontotemporal hypometabolism may be useful well before a full-blown dementia syndrome. Indeed, in many expert centres, FDG-PET already entered the routine clinical workup to support the diagnosis of FTLD in MCI. It is also relevant to this discussion to consider ‘FTD phenocopies’, i.e., clinical syndromes evocative of FTD, but due to a miscellany of, typically non-degenerative aetiologies[72]. As agreed by the expert panellist in this project, the diagnosis of these cases can be improved by adding FDG-PET, because the image pattern in those patients is typically characterized by normal metabolism[73]. However, in the aphasic MCI (language variants of FTLD), subjects report feeling that finding words became more effortful, including infrequent word-finding pauses, occasional spelling errors, a slowing of reading, sporadic reversals of word order, and rare misinterpretations of word meaning. Unlike AD diagnosis in MCI subjects, these language abnormalities are not usually part of normal aging, so they might be attributed to a neurodegenerative process [7]. Therefore, the panelists agree that FDG-PET may assist the diagnosis of possible FTLD in early stages where symptoms are mild and dissociated from one patient to another.

In the diagnosis of MCI possibly due to DLB, again the primary role of FDG-PET may be the identification of non-neurodegenerative conditions, based in the negative predictive value of a normal scan. The classic metabolic profile in DLB is posterior hypometabolism that includes the occipital cortex but with relative preservation of the posterior cingulate (the “posterior cingulate island sign”) [74]. Although these features of FDG-PET are supportive of the DLB diagnosis when added to traditional clinical investigation, the consensus opinion of the panelists was that dopamine transporter
brain SPECT, F18 FDOPA PET or I-123 MIBG cardiac scintigraphy may represent a more informative exam.

There are a number of factors that can account for heterogeneity in the available literature. Firstly, MCI is not a homogeneous diagnostic syndrome and some clarifications have been made throughout the last 15 years (amnestic; amnestic and non-amnestic; single domain and multidomain; minimal involvement of the most complex activities of daily life). In the Cochrane review, authors included longitudinal cohort studies published between 1999 and 2013 in which MCI subjects had undergone an FDG PET scan at baseline. Therefore, subjects included in the Cochrane review corresponded to different definitions of MCI since this term was initially described in 1999. Although the range of search is slightly more recent in our review (2003 to 2015), some effect of heterogeneity in the definition of MCI (converter/non converter, amnestic/multi-domain, MCI due AD, undefined MCI) may also be present. In addition, the time those MCI subjects were followed-up clinically after the baseline FDG PET scan to define conversion to dementia differed considerably. A duration of follow-up of only 1 year, (as used in the Cochrane review), can be too short to assess whether MCI is due to AD, because conversion can occur early (1 to 2 years) but also later (4 to 10 years) in the disease course. On the other hand, MCI subjects who have pre-FTLD converted to probable bvFTD within 4 years from baseline (on average 2.1 years) [6]. Finally, differences in methodological approaches between studies for the acquisition (PET scanners) and analysis of FDG-PET scans were present in the evaluated studies. Visual read is the most frequent method for brain FDG-PET evaluation. However, recent publications advocate for a combination of visual qualitative and semiquantitative analyses with well-defined thresholds and scaling procedures. This approach seems to be of particular usefulness in the early disease stages when abnormalities are not as extensive or prominent as in the dementia stages, a recommendation that was also achieved within the present initiative[1,75].

As emphasized by Morbelli et al., the variability of critical outcome measures observed in the literature for FDG-PET supporting the diagnosis of AD in MCI subjects is rather a reflection of the extensive work that is on-going to compare and validate analytical tools for guiding interpretation of brain FDG-PET data within the heterogeneous MCI population[45]. Nevertheless, some assumptions can be extracted on how current work can be improved to help and encourage progress in this field. From the clinical perspective, homogeneous populations of MCI subjects (amnestic, multidomain, converter or non-converter), time to conversion window, and longitudinal follow up should be clearly defined before specifying whether the results of FDG-PET should be considered valid or not. Larger populations should also be recruited in order to establish different groups of MCI subjects (AD, bvFTLD, PPA, DLB). In this respect, it is important to note the contribution that is being made by
multicentre groups like the Alzheimer’s Disease Neuroimaging initiative (ADNI), the Network for Standardisation of Dementia Diagnosis (NEST-DD), and the European Alzheimer’s Disease Consortium (EADC) PET project, among others. Although data provided by these multicentre databases are a good testing set for new analytical approaches, prospective diagnostic studies should be conducted. Finally, the accuracy of FDG-PET (and another biomarkers) depends on the method of interpretation (i.e. visual, semiquantitative), resulting in improved diagnostic and prognostic accuracy with software-aided reading. Studies are increasingly using parametric or voxel-wise analyses that compare the subject’s scan with a normative data set. However, this readout of cortical hypometabolism should be harmonized and a cut-off for positivity needs to be established [69].

The demonstration of diagnostic gain by the implementation of FDG-PET into the diagnostic toolbox in MCI is still today limited by the lack of (i) pathology confirmation, of (ii) head-to-head comparison between FDG-PET and clinical assessment versus an independent gold standard, of (iii) prospective studies about the clinical usefulness of a routine and systematic use in MCI subjects, and of (iv) impact on health outcomes and costs studies. The lack of such studies may be due to difficulties to engage sponsors, as FDG is already widely available and as such commercially less attractive.

In summary, in this framework, we have provided consensus guidance on the clinical usefulness of FDG-PET to diagnose MCI due to AD, FTLD, and DLB. In view of the gaps in formal evidence mainly in case of MCI due to FTLD and DLB, this guidance should be seen as interim, deserving more research.

Acknowledgements
The procedure for assessing scientific evidence and defining consensual recommendations was funded by the EANM and by the EAN. We thank the Guidelines Working Group of the EAN, particularly Simona Arcuti and Maurizio Leone, for their methodological advice.
Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart of selected papers for population, intervention, comparison, and outcome (PICO) questions 1-3, addressing whether FDG-PET should be performed, as adding diagnostic value (in terms of increased accuracy, and versus pathology, biomarker-based diagnosis or diagnosis at follow-up) as compared to standard clinical/neuropsychological assessment alone, to support the diagnosis of prodromal Alzheimer’s disease (PICO 1), frontotemporal lobar degeneration (PICO 2), and prodromal dementia with Lewy bodies (PICO 3) in subjects with mild cognitive impairment of uncertain origin.
Figure 2.

Typical metabolic patterns in mild cognitive impairment (MCI). (A) **MCI due to Alzheimer’s disease (AD)**. Left parietal, precuneus and posterior cingulate hypometabolism in an 68 year-old woman with episodic memory deficit and mild visuospatial impairment. Mini-mental state examination (MMSE) score= 25/30. Preserved instrumental activities of daily living (IADL). The diagnosis was supported through [18F]-Florbetapir amyloid PET showing diffuse uptake increase in the grey matter. (B) **MCI due to fronto-temporal lobar degeneration (FTLD)**. Bilateral frontal hypometabolism (more marked in the left inferior frontal as showed in the 3D-Statistical Surface Projections (SSP) images) and striatum, in a 66 year-old woman with preserved IADL and episodic memory deficit, anomia, dysarthria and executive dysfunction that emerged in the last year. The diagnosis was further supported through a negative [18F]-Florbetapir amyloid PET. MMSE score= 27/30. (C) **MCI due to DLB**. Bilateral parieto-occipital hypometabolism, mainly left hemisphere with relative preservation in posterior cingulate cortex (cingulate island sign) in a 78 year-old man.
with episodic memory deficit, executive dysfunction, and parkinsonism. MMSE score 28/30. Preserved IAD. The [123I]-Ioflupane SPECT showed uptake reduction in right putamen.
Table 1. Availability of evidence and panellists’ decisions supporting the use of brain FDG-PET in the diagnostic work-up of the main forms of mild cognitive impairment.

<table>
<thead>
<tr>
<th>PICO</th>
<th>RELATIVE AVAILABILITY OF EVIDENCE</th>
<th>PANELISTS’ RECOMMENDATIONS</th>
<th>MAIN REASONS FOR FINAL DECISION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – MCI due to AD</td>
<td>Fair</td>
<td>YES</td>
<td>A normal FDG-PET scan excludes neurodegeneration due to AD</td>
</tr>
<tr>
<td>2 – MCI due to FTLD</td>
<td>Lacking</td>
<td>YES</td>
<td>Typical hypometabolism pattern.</td>
</tr>
<tr>
<td>3 – MCI due to DLB</td>
<td>Lacking</td>
<td>YES</td>
<td>Typical hypometabolism pattern.</td>
</tr>
</tbody>
</table>

PICO= population, intervention, comparison, and outcome, MCI=Mild cognitive impairment. AD=Alzheimer disease, FTLD=fronto-temporal lobe degeneration, DLB=dementia with Lewy bodies.
Table 2. Quality of evidence for each critical outcome measure for PICO question 1.

<table>
<thead>
<tr>
<th>Critical outcomes</th>
<th>N. of papers</th>
<th>Sample size</th>
<th>Gold/reference standard</th>
<th>FDG-PET assessment</th>
<th>Risk of bias</th>
<th>Index test method</th>
<th>Applicability</th>
<th>Effect (CI)</th>
<th>Effect assessment</th>
<th>Effect inconsistency</th>
<th>Outcome quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>10</td>
<td>388 MCI converter to AD 545 Stable MCI or converter to non-AD</td>
<td>Conversion at follow-up</td>
<td>2 Visual 10 Semi-quantitative</td>
<td>Not serious</td>
<td>Serious</td>
<td>Serious</td>
<td>38% (CI: 8-75%) – 98% (CI: 87-100%)</td>
<td>MODERATE</td>
<td>Very serious</td>
<td>LOW</td>
</tr>
<tr>
<td>Specificity</td>
<td>10</td>
<td>388 MCI converter to AD 545 Stable MCI or converter to non-AD</td>
<td>Conversion at follow-up</td>
<td>2 Visual 10 Semi-quantitative</td>
<td>Not serious</td>
<td>Serious</td>
<td>Serious</td>
<td>41% (CI: 26-56%) – 97% (CI: 82-100%)</td>
<td>MODERATE</td>
<td>Very serious</td>
<td>LOW</td>
</tr>
<tr>
<td>Accuracy</td>
<td>11</td>
<td>395 MCI converter to AD 555 Stable MCI or converter to non-AD</td>
<td>Conversion at follow-up</td>
<td>2 Visual 11 Semi-quantitative</td>
<td>Not serious</td>
<td>Serious</td>
<td>Serious</td>
<td>58.4% (CI: 43-74%) – 100% (CI: 80-100%)</td>
<td>MODERATE</td>
<td>Very serious</td>
<td>LOW</td>
</tr>
<tr>
<td>PPV</td>
<td>6</td>
<td>224 MCI converter to AD 317 Stable MCI or converter to non-AD</td>
<td>Conversion at follow-up</td>
<td>2 Visual 6 Semi-quantitative</td>
<td>Not serious</td>
<td>Serious</td>
<td>Serious</td>
<td>41% (CI: 22-59%) – 85.2% (CI: 75-92)</td>
<td>LOW</td>
<td>Very serious</td>
<td>LOW</td>
</tr>
<tr>
<td>NPV</td>
<td>6</td>
<td>224 MCI converter to AD 317 Stable MCI or converter to non-AD</td>
<td>Conversion at follow-up</td>
<td>2 Visual 6 Semi-quantitative</td>
<td>Not serious</td>
<td>Serious</td>
<td>Serious</td>
<td>77% (CI: 64-87%) – 95% (CI: 75-100%)</td>
<td>HIGH</td>
<td>Not serious</td>
<td>MODERATE</td>
</tr>
<tr>
<td>AUC</td>
<td>7</td>
<td>245 MCI converter to AD 421 Stable MCI or converter to non-AD</td>
<td>Conversion at follow-up</td>
<td>1 Visual 7 Semi-quantitative</td>
<td>Not serious</td>
<td>Serious</td>
<td>Serious</td>
<td>66% (CI NA) – 96.8% (CI: 91-100%)</td>
<td>HIGH</td>
<td>Serious</td>
<td>LOW</td>
</tr>
<tr>
<td>LR+</td>
<td>1</td>
<td>77 MCI converter to AD 50 Stable MCI or converter to non-AD</td>
<td>Conversion at follow-up</td>
<td>1 Visual 1 Semi-quantitative</td>
<td>Not serious</td>
<td>Serious</td>
<td>Not serious</td>
<td>8.14 (CI: 4.75–13.96)</td>
<td>HIGH</td>
<td>NA</td>
<td>MODERATE</td>
</tr>
<tr>
<td>LR-</td>
<td>1</td>
<td>77 MCI converter to AD 50 Stable MCI or converter to non-AD</td>
<td>Conversion at follow-up</td>
<td>1 Visual 1 Semi-quantitative</td>
<td>Not serious</td>
<td>Serious</td>
<td>Not serious</td>
<td>0.12 (CI: 0.06–0.23)</td>
<td>HIGH</td>
<td>NA</td>
<td>MODERATE</td>
</tr>
</tbody>
</table>

RELATIVE AVAILABILITY OF EVIDENCE: FAIR

Risk of bias: assessment of the study design and other methodological features (e.g., subject selection, clinical diagnostic criteria used).

Index test methods: assessment of index test methodology (e.g., technical details, image analysis methods and statistical analysis).

Applicability: representativeness of the studied population and index test reproducibility in clinical practice (semi-quantitative methods correspond to ‘serious’ indirectness, visual + semi-quantitative methods correspond to ‘not serious’ indirectness, due to partial implementation of quantitation in clinical practice).
Effect: lowest and highest values for each critical outcome; when more values were obtained for the same outcome, the highest was reported.

Effect assessment: 51-70% low, 71-80% moderate, 81-100% high.

Effect inconsistency: ‘Not serious’ if lowest and highest values difference was 0-20, ‘serious’ 21-40, ‘very serious’ >40.

Outcome quality: summary of evidence as from all columns. PICO= population, intervention, comparison, and outcome, AD=Alzheimer disease, CI=confidence index….
References PICO 1, 2 e 3


22


74. Nestor, Peter; Altomare, Daniele ; Festari, Cristina ; Drzezga, Alexander; Rivolta, Jasmine; Walker, Zuzana; Bouwman, Femke; Orini, Stefania; Law, Ian; Nobili, Flavio ; Frisoni GB. Diagnostic utility of FDG-PET for the differential diagnosis among the main forms ofDementia. Eur J Nucl Med Mol Imaging. 2018 in this issue;

75. Nobili, Flavio; Festari, Cristina; Altomare, Daniele; Agosta, Federica; Orini, Stefania; Van Laere, Koen; Arbizu, Javier; Bouwman, Femke; Drezga, Alexander; Nestor, Peter; Walker, Zuzana; Boccardi M. Automated assessment of FDG-PET for the differential diagnosis in patients with neurodegenerative disorders. Eur J Nucl Med Molec Imaging in this issue.