Economic Evaluation of Interventions for Screening of Dementia

Saha, Sanjib; Gerdtham, Ulf-Göran; Toresson, Håkan; Minthon, Lennart; Jarl, Johan

2018

Document Version:
Other version

Working Paper 2018:20

Department of Economics
School of Economics and Management

Economic Evaluation of Interventions for Screening of Dementia

Sanjib Saha
Ulf-G. Gerdtham
Håkan Toresson
Lennart Minthon
Johan Jarl

August 2018

LUND UNIVERSITY

Economic Evaluation of Interventions for Screening of Dementia

Sanjib Saha[a*], Ulf-G Gerdtham[a,b,c], Håkan Toresson[d], Lennart Minthon[d], Johan Jarl[a]

[a] Health Economics Unit, Department of Clinical Science (Malmö), Lund University, Sweden
[b] Centre for Economic Demography, Lund University, Lund, Sweden
[c] Department of Economics, Lund University, Sweden
[d] Clinical Memory Research Unit, Department of Clinical Science (Malmö), Lund University, Sweden

Abstract

Objective: The objective is to systematically review the literature on economic evaluations of screening interventions for early diagnosis of dementia disorders.

Methods: A systematic search of published economic evaluation studies in English was conducted using specified key words in relevant databased and websites. Data extracted included methods and empirical evidence (costs, effects, incremental cost-effectiveness ratio) and we assessed if the conclusions made in terms of cost-effectiveness were supported by the reported evidence. The included studies were also assessed for reporting quality using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist.

Results: Fourteen studies were identified and broadly fell into two groups: screening without biomarkers and screening using biomarkers. There was a considerable heterogeneity in methodological approaches, target populations, study time frames, and perspectives as well as types of biomarkers used. The sensitivity and specificity of screening instruments are one of the important aspects in estimating the cost-effectiveness of the interventions. Cost-effectivness of non-biomarker based interventions can not be judged due to lack of information. The biomarkers based screening have the potential to be cost-effective but their effectiveness has to be established first.

Conclusion: More economic evaluations studies as well as good quality effectiveness studies are required in screening strategies before these can be implemented in the clinical practice.

Key words: Dementia, Screening, early diagnostic, economic evaluation

JEL Classification: H43; I10; I18

Acknowledgement
The study was funded by Eli Lilly and Company, Copenhagen, Denmark. The Health Economics Unit at Lund University also receives core funding from Government Grant for Clinical Research [ALF F 2014/354], and Region Skåne (Gerdtham).

Background

Dementia is a syndrome with progressive deterioration in several cognitive domains that interfere with activities of daily living. Alzheimer's disease (AD) is the most common dementia disorder and accounts for 60 – 70% of dementia cases [1]. Current estimates demonstrate that there are over 40 million people suffering from AD with the number expected to rise to over 100 million by the year 2050 [2].

Dementia affects many levels of society. Firstly, the individual suffers from impairments in cognition and functioning as well as impaired quality of life and shortened life expectancy [3]. Secondly, the relatives suffer from gradually losing a family member and in return receive a high care burden for the affected person. Indeed, the need for informal care increases when dementia progresses with deteriorating cognition and functioning [4]. Thirdly, dementia has a strong economic impact on the society. Care for persons with dementia is very costly and resource-demanding [5]. The worldwide societal costs for dementia were estimated to be 604 billion US dollars in 2010, of which 252 billion dollars in costs for informal care (42%) [5]. These costs are expected to increase in the future due to an aging population.

There are currently no specific treatments to prevent the progression of cognitive decline in dementia patients. Still, early and specific diagnosis is considered important as it can help guiding therapy and allowing patients and families to properly prepare for the consequences of the disease. Early diagnosis is considered to enable an improved overall course of the disease and delay the conversion from early/mild to more severe disease stages [6]. The most commonly used diagnostic tool at primary health care for dementia diagnosis is Mini-Mental State Examination (MMSE) which has some limitations such as being too long to be used in primary care, requiring cautious interpretation of the scores as well as being affected by cultural and educational factors [7]. However, MMSE can not distinguish between different dementia subtypes[8]. Therefore, there are different diagnostic criteria for different types of

dementia [9]. For example, the current standard diagnostic procedure of AD given by the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINDS-ADRDA guidelines) [10] include identifying a person with AD based on risk factors (e.g. age, sex, family history of AD), medical history (eg. age associated memory changes, depression, delirium), cognitive tests (eg. MMSE score, Mini-Cog, MoCA), level and degree of independence (eg. IADL tool), presence and degree of behavioural symptoms (eg. NPI-Q tool), and caregivers' assessments [11]. These criteria are based on the detection of the dementia syndrome and the classic features of AD, being more oriented to the exclusion of other non-degenerative causes of dementia. Although these criteria have proven to be very sensitive to AD when compared to other dementia disorders, its specificity reaches only 50 to 60% [12].

Significant amount of research is showing AD as a clinical-biological entity, in which biomarkers, especially pathophysiological markers, are present in the cerebrospinal fluid (CSF) before the symptoms are visible [13]. In 2011, the National Institute of Aging Alzheimer Association (NIA-AA) workgroup has published biomarker supported diagnostic criteria to cover all disease stages [14]. These new criteria, which allow different level of diagnostics based on biomarker results, permit detecting the pre-clinical stage, the pre-dementia stage, *i.e.* mild cognitive impairment (MCI), and the dementia stage [14]. Recently, "A/T/N" system has been proposed in which 7 major AD biomarkers are divided into 3 binary categories based on the nature of the pathophysiology of each measures [15]. "A" refers to the value of a β-amyloid biomarker (amyloid PET or CSF $A\beta_{42}$); "T," the value of a tau biomarker (CSF phospho tau, or tau PET); and "N," biomarkers of neurodegeneration or neuronal injury ([^{18}F]-fluorodeoxyglucose–PET, structural MRI, or CSF total tau). Each biomarker category is rated as positive or negative. An individual score might appear as

A+/T+/N−, or A+/T−/N−, etc. A clinically normal individual with A+/T+/N+ would mean that s/he is at the pre-clinical stage [15].

Thus, there are currently two broad approaches to identifiy dementia early in the disease process, patients without symptoms can be screened either by using existing guidelines or by searching for certain biomarkers, although the latter is limited to AD. Screening strategies using biomarkers are yet to be routinely implemented in the clinical practice [16]. However, which approach to use in clinical practice depends not only on the effectiveness but also the cost-effectiveness of the screening strategies. Indeed screening, especially universial or broad population approaches, is costly and it is important to acertain that the costs are reasonable in relation to the benefits of the screening initiative. Economic evaluation (EE) is an analytic technique which identifies, measures, values and compares the cost and outcomes of two or more alternative programs or interventions. Economic evaluations can ensure that the limited available resources are allocated as efficiently as possible, helping decision makers to make informed decision [17].

Conducting systematic reviews is a good way to identify the common characteristics of existing studies, to evaluate the studies, and to find the areas where more research is required. There is only one systematic review of EE of early diagnosis of AD which included studies before April 2011 [18]. Recently there has been a surge of EE of biomarkers to identify AD early which require exploration. Therefore, the objective is to study whether the screening interventions for early diagnosis of dementia disorders are cost-effective.

Methodology

We performed a systematic literature review to answer the research question in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [19]. Moreover, the guidelines for incorporating economic evidence from the

Campbell and Cochrane Economics Methods Group [20] has been followed including search criteria, data extraction, synthesis and critical analysis.

Search strategy

A systematic search was performed to identify relevant articles published in both health economics and biomedical databases from 01.01.2000 till 31.12.2015. The databases were Medline (Pubmed), Embase and ECONbase, EconLit, Cumulative Index to Nursing and Allied Health (CINAHL), The National Bureau of Economic Research, Latin American and Caribbean Literature on Health Sciences Database (LILACS) and Popline. In addition, we searched specific economic evaluation databases: the Centre for Reviews and Dissemination database maintained by NHS (http://www.crd.york.ac.uk/CRDWeb/) and the Cost-effectiveness analysis registry (http://healtheconomics.tuftsmedicalcenter.org/cear4/Home.aspx). We also searched additional articles from the reference lists of included studies. The search was performed with search/key words and the details of the search strategy, key words, and initial hits are provided in Annex 1 for the reproducibility and transparency of the work.

Inclusion and exclusion criteria

The literature search covers EEs of all types of interventions targeting patients with dementia disorders, their caregivers, and the patient-caregiver dyad. We defined screening interventions where the aim was to achieve and early diagnose of mild cognitive impairment (MCI) and/or dementia disorders using any approach.This means that EEs of interventions focusing on (1) non-pharmacological interventions and (2) overall management of the dementia patients in home/community/residential care are not included, as well as EEs of (3) pharmaceuticals. These three categories will be presented in the subsequent papers.

Studies were included if they satisfied the following criteria: (1) Studies that estimated the cost-effectiveness of screening strategies with or without subsequent treatment regimes as

part of the intervention; (2) EE such as cost minimization analysis (CMA), cost- effectiveness analysis (CEA), cost utility analysis (CUA) and cost benefit analysis (CBA); and (3) Reported in English in the scholarly literature. Studies were excluded if they were (1) cost studies such as cost-of-illness analysis; (2) reviews, notes, commentaries, editorials in scientific journals; and (3) study protocol or study design of interventions.

Selection and data extraction

After each search in the above-mentioned databases the initial hits were exported into EndNote and duplicates were removed. All articles were screened based on the inclusion and exclusion criteria, first based on titles and abstracts and second based on the full text. The selection of the articles was done by one co-author while a second co-author reviewed all studies where assessment according to inclusion or exclusion criteria was challenging.

We extracted data from the selected articles along two main dimensions; the result of the study (empirical evidence) and how the results have been derived (methodology). In terms of result, we extracted the Incremental Cost Effectiveness Ratio (ICER) from the selected articles, as well as its components (costs and outcomes). Furthermore, we scrutinized whether the intervention was reported as cost-effective by the authors and whether the reported information support the conclusions, based on different scenarios presented in Table 1 and cost-effectiveness acceptability curves (CEAC), if presented. We also identified whether the health outcomes were measured as utility index or as other outcomes, e.g. person diagnosed accurately. We used the NICE threshold (£30,000 per QALY gain) to term an intervention cost-effective [21, 22]. We also consider an intervention (weakly) cost-effective if the probability of the intervention being cost-effective is 90% (80%) at the NICE threshold in the CEAC.

Furthermore, we extracted the information on the the sensivity (*i.e.* to correctly identify the patients who have the disease) and specificity (*i.e.* to correctly identify the patients who do

not have the disease) of the screening appraoches as well as the comparative strategy/ies from the included studies when available.

Studies were appraised for quality of reporting using the CHEERS statement [23]. This checklist was produced with the aim of harmonizing the presentation of information, raising the quality standard of EEs. The CHEERS guidelines has 24 items in six categories (title and abstract, introduction, methods, results, discussion and other). The items were scored as 'Yes' (reported in full), 'No' (not reported), and 'Not Applicable'. In order to assign a score of reporting, we assigned a score of 1 if the requirement of reporting was completely fulfilled for that item and 0 otherwise. Therefore, the maximum score was 24.

Results

The systematic literature search identified fourteen studies that subsequently are included in this study. A flow chart of the study selection procedure is presented in Figure 1, and the detailed characteristics of the studies are presented in Table 2. We found six studies evaluating screening without any biomarkers and eight studies where biomarkers have been used. In Table 3, we presented the sensitivity and specificity of the biomarkers and their comparators.

Screening without biomarkers

Saito, et al. [24] used a Markov model to estimate the cost reduction of community based dementia screening compared to no screening for adults. The Markov model had 6 health states: nondemented, MCI, mild dementia, moderate dementia, severe dementia, and death. Over a series of ten 1-year cycles, patients were allowed to move from less severe states to more severe states and vice versa except for death state. The screening was done at year one and persons having MCI or dementia were identified. The model assumed that treatment started immediately after diagnosis which prolonged the mild stage of dementia and reduced

the length of the moderate and severe stages. The treatment at the mild stage is less costly than moderate and severe stages of dementia. Over the ten-year period, the model predicted a cost reduction of 9.8% compared to the no screening strategy.

Banerjee et al. [25] used a mathematical equation for calculating reduced costs and increased Quality Adjusted Life Years (QALYs) due to early diagnosis of patients by a multi disciplinary and interagency team followed by proper treatment and management in the UK compared to usual care. They estimated that the program could save £950 million in ten years with a gain of 6,250 QALYs. The saving would come mostly from delaying the admission to institutional care. On the other hand, Dixon et al. [26] showed that a one-off screening for dementia for the 75 years old UK and Wales population would be cost neutral ranging from £3.6 million net cost to £4.7 million net savings. Although they have mentioned that the screening is associatied with several health benefit such as improved cognition, communication and quality of life for the dementia patients, this has not been included in the analysis. Another EE of a nationwide opportunistic screening for people aged 60 years and older compared to no screening was conducted in South Korea [27]. The screening was done at the primary health care centre by Mini Mental State Examination (MMSE) and if MCI was found, they were further screened by clinical and neuropsychological assessments, laboratory tests and brain imaging by psychiatrist or neurologists. A Markov model with a 10-year timeframe was used and resulted in an ICER ranging from $24,150 to $35,661 per QALY depending on age.

Only one of the included CEA was performed alongside a randomised controlled trial (RCT) [28]. In that RCT, conducted in the Netherlands, dementia patients were diagnosed with an integrated multidisciplinary component which included home visits, hospital visits, computed tomographic scan and a variety of blood tests. The comparator was the usual diagnostic procedure performed by General physicians (GPs) or at the regional facility centre. The

follow-up was performed for one year. The ICER was estimated to be €1,267 per QALY. The increased cost of screening was partly offset by reduced cost due to less time spent in the institutional care for the intervention group participants. The informal care cost was also lower for the intervention group.

One CBA showed that early identification followed by treatment is economically efficient for both state and federal governments in the USA compared to the current practice of diagnosis and treatment for the AD patients [29]. Using a monte carlo simulation model, Weimer et al. [29] showed that the early identification would have net social benefits in monetary values as well as reduced time at nursing home. Here social benefits refers to the sum of monetized values of impacts of the interventions on all persons, e.g. patients, caregivers and tax payer. The benefits depends on person's age, sex and MMSE score. For example, for a 70 years old married woman starting with MMSE score of 26, the social benefit will be $94,000 (2006 price year) and she will spend 1.2 years less time in a nursing home.

Biomarkers based screening

McMahon et al. performed two CEAs of functional neuroimaging test comparing to the conventional diagnostic strategy for patients suspected to have AD in the USA [30, 31]. The neuroimaging included either single photon emission computed tomography (SPECT) or dynamic susceptibility contrast-enhanced magnetic resonance imaging (DSC MRI). The standard diagnostic test for AD included detailed history, assessment of cognition and functional status, laboratory testing and a brain imaging examination such as nonenhanced computed tomography or MRI. The SPECT was found to be more costly and less effective compared to standard diagnostic test. The ICER for MRI was $479,500 /QALY (1998 price year) [30] and in a subsequent study the ICER for DSC MRI was $598,800 /QALY (1999 price year) [31]. The high ICERs were mainly driven by a very small effect on QALYs (0.004 and 0.0017, respectively) in the two studies compare to the standard approach.

We found three studies where the biomarkers in CSF was examined in order to screen for AD [32-34]. Guo et al. [32] used a Discrete Event Simulation (DES) model to estimate the cost-effectiveness of florbetaben with PET biomarkers for screening of pre-dementia and dementia patients in the USA, compared to standard diagnosis. Time to confirmation of diagnosis was estimated to be reduced by 2.15 months for pre-dementia patients and 2.42 months for dementia patients. The institutional care was estimated to be delayed by 0.24 years for pre-dementia patients and 0.12 years for dementia patients as patients stay longer in the mild stages of AD. QALY was estimated to be 0.15 higher for the pre-dementia group and 0.03 higher for the dementia group per patient over lifetime compard to standard diagnostics. The direct medical care costs was also estimated to be lower in the intervention group. The positive findings mainly derived not from early diagnosis but rather early confirmation of diagnosis. The early confirmed diagnosis allowed early treatment and thus yielded better clinical and economic outcomes. Younger patients had greater cost savings compared to older, due to a longer life expectancy.

Valcarcel-Nazco et al. found that the use of biomarkers in CSF as an option for early identification of AD in MCI and dementia patients followed by treatment had lower costs and better outcomes compared to current practice (clinical standard diagnostic) in Spain [33]. A decision tree from a lifetime perspective was used targeting patients 60 years and older with MCI (scenario 1) or symptoms of dementia (scenario 2). For MCI patients, the cost saving was €1,833 which comes mainly from less use of donepezil for misdiagnosed patients. For scenario 2, the cost saving was €1,134. The outcome was the probability of accurate diagnosis in both scenarios which were significantly higher compared to standard diagnostic.

Handels et al. [34] assumed a perfect biomarker in CSF for MCI subjects [34] in the Netherlands followed by disease modifying treatment. At first, the current practice was used to identify whether patients had AD or not and then CSF biomarker was used with the

assumption of 100% sensitivity and 100% specificity to identify AD patients in a hypothetical scenario. The researchers used a decision tree and DES model over the lifetime of the simulated (MCI) subjects. The perfect biomarker would result in a gain of 0.39 QALY, save €33,622 and delay the onset of dementia by 1.3 years compared to current practice.

Silverman et al. [35] and Moulin-Romsee et al. [36] used the same model in the USA and Belgium to estimate the cost per correct AD diagnosis comparing to current practice. The intervention included FDG PET (Fluorodeoxyglucose PET) in the screening process. The proposed algorithm including FDG PET can generate more accurate diagnoses compared to current practice, thus resulting in more accurate treatment with fewer side-effects of medication and better functional condition of patients.

Djalalov et al. [37] performed a CEA of genetic screening followed by pharmaceautical treatment for patients with amnestic mild cognitive imparment (AMCI) in Canada. They used genetic testing for the presence of APOE ε4 alleles arguing that AMCI patients who carry two APOE ε4 alleles have 90% chance of developing AD by the age of 80. They used a Markov model where the comparator was standard treatment. The ICER was estimated to Can$ 38,016 per QALY. However, the accuracy of APOE ε4 genotyping to predict the progression from MCI to AD is not yet established.

Discussion

The economic evaluation of a screening programme differs in a number of aspects comparing to other healthcare interventions. The total costs of screening programmes are relatively high. These include the costs of the screening procedure itself for a large number of people, the costs of follow-up procedures for people with a positive screening result, as well as the costs of implementing the programme. Screening most often includes confirmatory tests and treatments for those with a positive result [38]. In order to establish the value of identification

of a disease case the EE needs to incorporate potential treatments, as screening in the absence of effective treatments most likely will not be cost-effective.

The identified studies differ in many aspects such as type of screening strategies, biomarkers used, sensitivity and specificity of the biomarkers, length of study period, target groups, perspective, included costs and outcomes, and instruments to measure the outcomes. This makes general comparison across all studies difficult to achieve as there are also differences in the setting of the different studies, e.g. different healthcare systems, community or nursing home care, clinical practices, population values, availability and accessibility of drugs and technologies. However, we will discuss the main differences between studies in relation to the results.

Cost-effectiveness is at its heart a subjective concept as it refers to if an intervention is worth its costs, *i.e.* the decision-maker willingness-to-pay for the outcome under study. This will differ between settings but also between individuals, and it is therefore essential that the authors of EEs are clear about the valuation of the outcomes when determining an intervention's cost-effectiveness. Preferably a societal valuation should be used when reporting cost-effectiveness although this value is generally unknown. An exception is the value of a QALY where NICE in the UK uses a cost-effectiveness threshold range of £20,000 to £30,000 per QALY gained [21, 22]. There are no official guidelines for the USA and Australia although 50,000 US$/QALY is frequently employed as a threshold in the USA [39] and 50,000 AUS$/Disability Adjusted Life Year (DALY) in Australia [40]. For all other outcome measure, each study needs to establish the societal valuation of the used outcome in order to, potentially, claim cost-effectiveness. A few exceptions exist however; if the intervention is both better (worse) and less (more) costly than the comparator (scenarios 2 and 4 in the Table 1), it is (not) preferred irrespective of the valuation of the outcome measure.

For simplicity, we will term all clearly preferred alternatives as cost-effective in the continued discussion.

In table 4, we presented the cost-effectiveness of the included articles as reported by the authors as well as our own assessment based on the reported information. It is foremost important to establish that there is a difference in costs and/or outcomes between the intervention and the comparator before calculating the ICER, for example by reporting confidence intervals (CIs). If CIs were not available, we take a conservative approach and assess the cost-effectiveness as "unknown due to lack of information" as it cannot be established if the intervention is different from the comparator. Our assessments are in line with the reported conclusions in three studies [32-34] as the interventions were significantly better and cheaper. However, many studies do not report the 95% CIs (or corresponding test) for either costs or outcomes (Table 4) [25, 27, 28, 37]. We acknowledge that, in EE, costs and effects are very disperse and it may be difficult to find significant differences between two comparators. However, 95% CIs (or corresponding tests) of the differences in costs and effects should always be included.

Many studies handle the uncertainties around costs and effects by presenting Cost-Effectiveness Acceptability Curve (CEAC) [27, 28, 33, 37] (Table 4) which is a good practice that should be included in all EE. CEAC was developed as an alternative to producing CIs around the ICER. However, there is no agreement on when to claim an intervention cost-effective based on the findings from the CEAC. For ease of comparsion, we termed an intervention "cost-effective" if the intervention had 90% probability to be cost-effective at NICE threshold as it is clearly preferable to the alternative. It should be noted that decision makers are advised not to implement an intervention based on the findings from the CEAC [41-43] and that the correct interpretation of the CEAC is with regard to the uncertainty around the estimated cost-effectiveness ratio.

There is a monetary valuation (threshold) of QALY which researchers as well as policy makers can rely upon when comparing different interventions in terms of cost-effectiveness, despite the disagreement about the precise value. However, there is no agreed upon valuation for other effectiveness measurement such as MMSE score, number of person diagnosed, accurate diagnosis etc. (Table 2). In these cases, it falls upon the authors to try to establish the societal valuation of the used outcome, for example by comparing to the value that the society has been willing to pay in the past. However, none of the included studies makes a convincing case for the valuation which means that only in those cases where the intervention is either better but not more expensive, cheaper but not worse, or better and cheaper than the alternative (or the other way around) is the cost-effectiveness argument clear. For the majority of included studies, the intervention is better but also more costly, and it is impossible to establish if the intervention is cost-effective. In Wolfs et al [28] for example, improvement in MMSE score was considered cost-effective at an ICER of €333/ additional point on the MMSE. However, it is not established in the study that the societal valuation of one additional point in MMSE is above €333. It is therefore not possible, with any certainty, to make any conclusions regarding the cost-effectiveness of the intervention and labelling these interventions as "cost-effective" is inappropriate. We consider the cost-effectiveness of these intervention "unknown due to no agreed cost-effective threshold value" (Table 4). Future research is needed to reach an agreement among researchers and policymakers regarding the valuation of commonly used outcome measures, such as MMSE.

Effectiveness of the diagnostic tests

It is suggested that a screening program should take into account the sensitivity and specificity of the screening technology, the number of positive and negative results and the implications of false positive and false negative results [44]. The sensitivity and specificity are key factors that determine the effectiveness and cost-effective of a screening strategy [45].

We found that the sensitivity and specificity varies from study to study and also from test to test, both for intervention and comparator (Table 3). The sensitivity varies from as high as 100% for perfect CSF biomarkers [34] and genetic screening [37] to as low as 50% for visual SPECT [30]. The specificity also varies from 100% [30, 34] to 67% [36, 37]. It is interesting to note that four studies used FDG PET with almost similar sensitivity and specifity [31, 32, 35, 36] but the comparators' sensitivity and specificity varies widely. The FDG PET were cost-effective options [32, 35, 36] except for one study [31] where the specificity rates for comparator, i.e. standard examination, was higher than in other studies. This could be one explanation for the findings.

Another important aspect that could have an effect on the cost-effectiveness is the effect of false positive on patients and/or caregivers. It is possible that such patients and caregiver have increased levels of anxiety which could affect their quality of life. This would have an overall negative effect on the cost-effectiveness of screening strategies with low sensitivity.

We found four studies where the sensitivity and specifity of the screening or diagnostic strategies were not mentioned [24, 25, 28, 29], which is not a good practice. Sometimes the sensitivity and specificity values are based on assumption or expert opinion. Although it is a common practice in EE to use hypothetical estimate and/or expert opinion in the absence of real data, researchers need to verify the uncertainties in the sensitivity analyses. The readers or decision makers also need to consider this issues while making decision or implementing policies.

Early intervention

The purpose of early diagnosis is to introduce early treatment to improve cognitive dysfunction, delay conversion from MCI to dementia, decrease development of behaviour problems, allowing the patients and family members to plan ahead and make needed adjustments to keep patients living independently or with the family as long as possible and

delay institutional care. Given that institutional care is expensive, any delay to institutional care would have meaningful impact on the cost-effectiveness of any particular intervention. Most of the simulation models have introduced medication (e.g. acetylcholinesterase inhibitors and memantine). However, most of the studies have not incorporated non-pharmaceutical interventions as a treatment besides medication. It has been shown that physical exercise and psychological interventions such as cognitive rehabilitations, cognitive stimulation therapies, reminiscence therapy are effective [46] and some of the interventions are cost-effective as well for the patients and caregivers [47]. Only a few studies have included non-pharmaceuticals interventions for the patients [24, 26, 32] and interventions for caregivers was included only in one study [32]. Non-pharmacological interventions can improve the overall quality of life of the patients and caregiver which, in turn, will have the possibility to make screening interventions favourable in terms of cost-effectiveness.

Issues with simulation models

We found that researchers have used Markov model [24, 27, 37], unspecified simulation model [25, 26] and DES model with decision tree [32, 34] for their analyses (Table 1). When the entire AD lifecycle is modelled, starting from identification to various stages of AD and finally death, it is suggested that the outcomes should capture as many health-related variables as possible including cognition, behaviours, and functioning instead of a single outcome such as overall health-related quality of life since dementia is a multifaceted disease [48]. The DES models [32, 34] have captured more outcomes than the Markov models [24, 27, 37]. Moreover, Markov models have a limitation on managing individual level characteristics which can be overcome by DES model. However, these processes make the model complex and reduces transparency. The Markov models vary in included health states. For example, Saito et al. assumed that the patients can move forward as well as backward between disease severity states [24] while Yu et al. [27] only allowed patients to move forward towards more

severe health states. We found that study using Markov model [31] report that FDG PET is not cost-effective whereas study using DES models [32] report that FDG PET is a cost-effective option comparing to standard care.

Lack of transparency regarding the assumptions in the DES model in AD research has been raised as a cause for concern [48]. One imporant aspect is how patients have been stimulated and how the characteristics were assigned in the studies with DES models [32, 34]. The classical debate that a model developer faces is the balance between model's transparency and accuracy. Transparency means the ability to understand the logical arguments of a model to be able to reproduce it. Accuracy means to capture the real life situation [49]. The equilibrium between a model's accuracy and transparency is difficult to obtain because when the model is made more accurate, the complexity of the model increases concurrently, which in turn decreases the decision makers' ability to understand it. Therefore, some researchers emphasize the transparency of the model [50] whereas others argue that accuracy should be the main priority [51]. We agree with Oremus et al. [48, 52] that a model needs to emphasize on the transparency of model structure and inputs and also on the validity of assumptions regarding healthcare costs, health utilities and mortality across the dementia diseases' stages. It is recommended that the modellers should present a technical note of the model so that others can understand all the assumptions behind the model [53]. Since the space of a journal is less likely to be a problem nowadays, attaching a technical note as supplementary materials would be helpful.

Another important issue in modelling is the validity of the model itself which shows the robustness of the findings. Authors should consider possible variations using internal, external and structural validation of the models. Although most of the authors have performed sensitivity analyses (Table 2) to test the robustness of their results, the structural assumptions were untested. In case of probabilistic sensitivity analyses, the distribution of

the parameters and sources were missing in all studies. Therefore, findings from these need to be interpreated with cautious [27, 32-34, 37]. To minimize uncertainty in parameter estimates, the use of data from meta-analyses or systematic reviews has been suggested [48, 54]. However, data in reviews or meta-analysis are often not country specific and may introduce a bias, especially for costs.

Timeframe of analysis

The duration of the studies ranges from 1-year to lifetime (Table 2). Generally, short duration runs the risk of missing much of the benefits of the intervention while including most of the costs and thereby resulting in an upward biased cost-effectiveness ratio. FDG PET was cost-effective when the analysis time frame was lifetime [32] and not cost-effective when the duration of analysis was 18 months [31]. It is important for the CEA of dementia to have a lifetime perspective. Most of the costs of a screening intervention are incurred within a relatively short time period, but the benefits may not be accured for many years. In the western countries the costs for institutionalised care constitute a significant cost burden. A lifetime perspective has the potential to capture whether early identification delays and/or reduces days in institutionalised care. However, discounting has particular implications when the time horizon for analysis is long-term. A decision to discount costs or outcomes, or both, and the choice of the discount rate (s) may have a significant impact on the cost-effectiveness of the intervention and needs to be carefully considered. Most often the discount rates are country specific, for example, in the Dutch study different discount rates have been used for costs and outcomes which is according to Dutch national guideline [34].

Age of screening for dementia

We also find the starting age of the population groups differs from as early as 18 years [24] to as late as 75 years [26]. Guo et al. showed that younger patients would have greater gain in net cost saving and QALYs due to longer life expectancny compared to older [32]. The same

was observed in Saito et al. [24] where it was more cost saving to identify the disease at the younger age than at the later stage of life.

Issues with caregivers

Early diagnosis of dementia is considered helpful to provide caregivers enough time to cope with the problem [55, 56]. Dementia is expected to affect people close to the patient directly and indirectly through the burden of informal care. Therefore, the benefits of early intervention from the point of caregivers need to be included in the EE, especially as informal care is a major cost component [5]. However, there are some challenges with estimating the cost of informal care and thereby including it in an EE [57]. First, it is debatable what types of activities should be considered as caregiving. The World Alzheimer Report considers both time related to helping patients with activities of daily living and support with instrumental activities of daily living as caregiving. Second, it is difficult to measure caregiver activities over a long period of time which might lead to recall and interpolation bias. Third, there is much controversy regarding the valuation of time for informal care [57].

We found that some studies have included the cost for caregivers [24, 26, 28, 29, 31, 32, 36, 37] in their analysis whereas some studies did not [25, 27, 30, 33-35]. The cost for informal care varies between \$7.63-30.06/hour[1], for example, €10/hour (\$14.48) [36], €8.54/hour (\$12.32) [28], \$21.13/ hour (\$30.06) [31], \$14.69/hour (\$17.28) [29], and \$7.25/hour (\$7.63) [32]. In terms of the health benefit of the caregivers, only two studies included those [29, 32]. Future studies need to focus on incorporating both the costs and health outcomes of the caregivers in their analysis. Otherwise, the findings may be underestimated or overestimated.

Quality assessment

[1]Costs in parentheses are converted to 2015 USD using consumer price index and purchasing power parity

We assessed the articles based on the CHEERS statement and observed that the quality of reporting was insufficient for several articles. It can be argued that CHEERS statement is very recent and many of the articles were published before the CHEERS statement. However, other guidelines were available earlier, for example [58-61], and following any of these guidelines would have improved reporting. We found some studies that did not mention the funding source which is a major drawback [24, 25, 30-32, 36] and these studies performed poor on the CHEERS score overall. Several items were only partially or not reported at all in most articles such as description of costing methods and sources of costs items, which impede proper comparison between the studies. We also found that most studies did not have heterogeneity analyses. We hope that the availability of the CHEERS statement will lead to improvements in reporting. However, it should be kept in mind that these guidelines are to ensure the quality of the reporting and not the quality of the study, although high correlation is expected.

Role of funding source

A majority of the studies were funded by governmental or non-governmental organizations (Table 2) although many articles did not mention funding source. This is different compared to EE of drugs for dementia and AD patients which generally are funded by pharmaceutical industries [52, 62]. No differences in cost-effectiveness could be discerned based on funding source.

Strength and Limitation

The current literature review poses particular strengths. In line with recommendations, we searched key electronic bibliographic databases and other sources. Manual searching of reference lists of the reviewed articles was carried out to identify relevant studies. Identified studies were independently assessed for inclusion against a set of predetermined criteria. No

restrictions were applied on types of EE or country of origin including both trial- and model-based EE.

There may have been some potential limitations to our study. We assess the quality of reporting based on CHEERS statement but we have not assessed the methodological rigour of the identified studies. Our conclusions of cost-effectiveness are thus based on the presented information in the studies and do not account for potential weaknesses in methodology. In addition, we have not performed a systematic quantitative assessment to identify key drivers of the cost-effectiveness.

Conclusion

In conclusion, we find that within the different areas, there are different interventions some of which are cost-effective while others are not. The biomarkers have the potential to be cost-effective but their effectiveness has to be established first. More research is required to establish the cost-effectiveness of screening interventions overall. It is worth stating that the cost-effectiveness ratio is not the only aspect to consider in decision-making regarding implementation of interventions. Instead, a country and context-specific process for decision making should be considered, reflecting legislation and involving patients group, caregivers and civil society organizations [63, 64].

References

1. Mangialasche, F., et al., *Alzheimer's disease: clinical trials and drug development.* The Lancet Neurology, 2010. **9**(7): p. 702-716.
2. Alzheimer's, A., *2013 Alzheimer's disease facts and figures.* Alzheimer's & Dementia, 2013. **9**(2): p. 208-245.
3. Rizzuto, D., et al., *Dementia after age 75: survival in different severity stages and years of life lost.* Current Alzheimer Research, 2012. **9**(7): p. 795-800.
4. Nordberg, G., et al., *The amount of informal and formal care among non-demented and demented elderly persons—results from a Swedish population-based study.* International journal of geriatric psychiatry, 2005. **20**(9): p. 862-871.
5. Wimo, A., et al., *The worldwide economic impact of dementia 2010.* Alzheimer's & Dementia, 2013. **9**(1): p. 1-11. e3.
6. Ashford, J.W., et al., *Should older adults be screened for dementia? It is important to screen for evidence of dementia!* Alzheimer's & dementia : the journal of the Alzheimer's Association, 2007. **3**(2): p. 75-80.
7. Yokomizo, J.E., S.S. Simon, and C.M. Bottino, *Cognitive screening for dementia in primary care: a systematic review.* Int Psychogeriatr, 2014. **26**(11): p. 1783-804.
8. Creavin, S.T., et al., *Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations.* Cochrane Database Syst Rev, 2016(1): p. Cd011145.
9. Robinson, L., E. Tang, and J.-P. Taylor, *Dementia: timely diagnosis and early intervention.* BMJ : British Medical Journal, 2015. **350**.
10. McKhann, G., et al., *Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease.* Neurology, 1984. **34**(7): p. 939-44.
11. Galvin, J.E. and C.H. Sadowsky, *Practical Guidelines for the Recognition and Diagnosis of Dementia.* The Journal of the American Board of Family Medicine, 2012. **25**(3): p. 367-382.
12. Lim, A., et al., *Clinico-neuropathological correlation of Alzheimer's disease in a community-based case series.* J Am Geriatr Soc, 1999. **47**(5): p. 564-9.
13. Bloudek, L.M., et al., *Review and meta-analysis of biomarkers and diagnostic imaging in Alzheimer's disease.* J Alzheimers Dis, 2011. **26**(4): p. 627-45.
14. Sperling, R.A., et al., *Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.* Alzheimers Dement, 2011. **7**(3): p. 280-92.
15. Jack, C.R., et al., *A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers.* Neurology, 2016. **87**(5): p. 539-547.
16. Albert, M.S., et al., *The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.* Alzheimers Dement, 2011. **7**(3): p. 270-9.
17. Michael F. Drummond, et al., *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed. 2005: Oxford University Press.
18. Handels, R.L., et al., *Diagnosing Alzheimer's disease: a systematic review of economic evaluations.* Alzheimers Dement, 2014. **10**(2): p. 225-37.
19. Moher, D., et al., *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA StatementThe PRISMA Statement.* Annals of Internal Medicine, 2009. **151**(4): p. 264-269.

20. Higgins, J.P. and S. Green, *Cochrane handbook for systematic reviews of interventions*. Vol. 5. 2008: Wiley Online Library.
21. McCabe, C., K. Claxton, and A.J. Culyer, *The NICE Cost-Effectiveness Threshold: What it is and What that Means.* PharmacoEconomics, 2008. **26**(9): p. 733-744.
22. NICE, *NICE guide to the methods of health technology appraisal*. 2004, NICE.
23. Husereau, D., et al., *Consolidated health economic evaluation reporting standards (CHEERS) statement.* BMC medicine, 2013. **11**(1): p. 1.
24. Saito, E., et al., *Cost effective community based dementia screening: A markov model simulation.* International Journal of Alzheimer's Disease, 2014. **2014**.
25. Banarjee, S. and R. Wittenberg, *Clinical and cost effectiveness of services for early diagnosis and intervention in dementia.* International Journal of Geriatric Psychiatry, 2009. **24**(7): p. 748-754.
26. Dixon, J., et al., *Exploring the cost-effectiveness of a one-off screen for dementia (for people aged 75years in England and Wales).* International Journal of Geriatric Psychiatry, 2014. **30**(5): p. 446-452.
27. Yu, S.Y., et al., *Cost-effectiveness of nationwide opportunistic screening program for dementia in South Korea.* Journal of Alzheimer's Disease, 2014: p. epub.
28. Wolfs, C.A., et al., *Economic evaluation of an integrated diagnostic approach for psychogeriatric patients: results of a randomized controlled trial.* Arch Gen Psychiatry, 2009. **66**(3): p. 313-23.
29. Weimer, D.L. and M.A. Sager, *Early identification and treatment of Alzheimer's disease: social and fiscal outcomes.* Alzheimers Dement, 2009. **5**(3): p. 215-26.
30. McMahon, P.M., et al., *Cost-effectiveness of functional imaging tests in the diagnosis of Alzheimer disease.* Radiology, 2000. **217**(1): p. 58-68.
31. McMahon, P.M., et al., *Cost-effectiveness of PET in the diagnosis of Alzheimer disease.* Radiology, 2003. **228**(2): p. 515-22.
32. Guo, S., et al., *Florbetaben PET in the early diagnosis of alzheimer's disease: A discrete event simulation to explore its potential value and key data gaps.* International Journal of Alzheimer's Disease, 2012.
33. Valcarcel-Nazco, C., et al., *Cost-Effectiveness of the Use of Biomarkers in Cerebrospinal Fluid for Alzheimer's Disease.* Journal of Alzheimers Disease, 2014. **42**(3): p. 777-788.
34. Handels, R.L., et al., *Early cost-utility analysis of general and cerebrospinal fluid-specific Alzheimer's disease biomarkers for hypothetical disease-modifying treatment decision in mild cognitive impairment.* Alzheimers Dement, 2015. **11**(8): p. 896-905.
35. Silverman, D.H.S., et al., *Evaluating Early Dementia With and Without Assessment of Regional Cerebral Metabolism by PET: A Comparison of Predicted Costs and Benefits.* Journal of Nuclear Medicine, 2002. **43**(2): p. 253-266.
36. Moulin-Romsee, G., et al., *Cost-effectiveness of 18F-fluorodeoxyglucose positron emission tomography in the assessment of early dementia from a Belgian and European perspective.* Eur J Neurol, 2005. **12**(4): p. 254-63.
37. Djalalov, S., et al., *Genetic Testing in Combination with Preventive Donepezil Treatment for Patients with Amnestic Mild Cognitive Impairment.* Molecular Diagnosis & Therapy, 2012. **16**(6): p. 389-399.
38. Karnon, J., et al., *A review and critique of modelling in prioritising and designing screening programmes.* Health Technol Assess, 2007. **11**(52): p. iii-iv, ix-xi, 1-145.
39. Weinstein, M.C.P., - *How Much Are Americans Willing to Pay for a Quality-Adjusted Life Year? Editorial.* - Medical Care April 2008;46(4):343-345, 2008(- 0025-7079).

40. Moodie, M., et al., *Cost-effectiveness of a family-based GP-mediated intervention targeting overweight and moderately obese children.* Economics & Human Biology, 2008. **6**(3): p. 363-376.
41. Fenwick, E. and S. Byford, *A guide to cost-effectiveness acceptability curves.* Br J Psychiatry, 2005. **187**: p. 106-8.
42. Fenwick, E., B.J. O'Brien, and A. Briggs, *Cost-effectiveness acceptability curves--facts, fallacies and frequently asked questions.* Health Econ, 2004. **13**(5): p. 405-15.
43. Fenwick, E., K. Claxton, and M. Sculpher, *Representing uncertainty: the role of cost-effectiveness acceptability curves.* Health Econ, 2001. **10**(8): p. 779-87.
44. Mäklin, S., et al., *Costs and economic evaluation.* HTA Core Model for screening technologies, 2012: p. 79.
45. Steuten, L.M.G. and S.D. Ramsey, *Improving early cycle economic evaluation of diagnostic technologies.* Expert Review of Pharmacoeconomics & Outcomes Research, 2014. **14**(4): p. 491-498.
46. Freeman, M., *A systematic evidence review of non-pharmacological interventions for behavioral symptoms of dementia*. 2011.
47. Saha, S., et al., *Economic Evaluation of Nonpharmacological Interventions for Dementia Patients and their Caregivers-A Systematic Literature Review*. 2018.
48. Oremus, M. and J.-E. Tarride, *Modeling cost–effectiveness of pharmaceuticals in Alzheimer's disease.* Expert review of pharmacoeconomics & outcomes research, 2012. **12**(3): p. 275-277.
49. Eddy, D.M., *Accuracy versus transparency in pharmacoeconomic modelling.* Pharmacoeconomics, 2006. **24**(9): p. 837-844.
50. Weinstein, M.C., et al., *Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices-Modeling Studies.* Value in Health, 2003. **6**(1): p. 9-17.
51. Eddy, D.M., *Accuracy versus Transparency in Pharmacoeconomic Modelling: Finding the Right Balance.* PharmacoEconomics, 2006. **24**(9): p. 837-844.
52. Oremus, M., *Systematic review of economic evaluations of Alzheimer's disease medications.* Expert review of pharmacoeconomics & outcomes research, 2008. **8**(3): p. 273-289.
53. Gold, M., et al., *Cost-Effectiveness in Health and Medicine*. 1996, New York: Oxford University Press.
54. Drummond MF, et al., *Methods for the Economic Evaluation of Health Care Programmes*. Vol. 3rd. 2005, Oxford: Oxford University Press.
55. Dubois, B., et al., *Timely Diagnosis for Alzheimer's Disease: A Literature Review on Benefits and Challenges.* J Alzheimers Dis, 2016. **49**(3): p. 617-31.
56. de Vugt, M.E. and F.R. Verhey, *The impact of early dementia diagnosis and intervention on informal caregivers.* Prog Neurobiol, 2013. **110**: p. 54-62.
57. Koopmanschap, M.A., et al., *An overview of methods and applications to value informal care in economic evaluations of healthcare.* Pharmacoeconomics, 2008. **26**(4): p. 269-280.
58. Ramsey, S., et al., *Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report.* Value Health, 2005. **8**(5): p. 521-33.
59. Siegel, J.E., et al., *Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine.* Jama, 1996. **276**(16): p. 1339-41.
60. Drummond, M.F. and T.O. Jefferson, *Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party.* Bmj, 1996. **313**(7052): p. 275-83.

61. McGhan, W.F., et al., *The ISPOR Good Practices for Quality Improvement of Cost-Effectiveness Research Task Force Report.* Value Health, 2009. **12**(8): p. 1086-99.
62. Pouryamout, L., et al., *Economic Evaluation of Treatment Options in Patients with Alzheimer's Disease.* Drugs, 2012. **72**(6): p. 789-802.
63. Bertram, M.Y., et al., *Cost–effectiveness thresholds: pros and cons.*
64. Woods, B., et al., *Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research.* Value Health, 2016. **19**(8): p. 929-935.

Table 1: Decision rules for economic evaluations (new intervention vs. comparator)

Scenarios	Cost	Outcome	Interpretation
1	↑	↑	**Cost-effective** if the willingness-to-pay exceeds the ICER
2	↓	↑	**Cost-effective** (new intervention dominates the comparator)
3	≈	↑	**Cost-effective** (new intervention dominates the comparator)
4	↑	↓	**Not cost-effective** (comparator dominates the new intervention)
5	↓	↓	**Cost-effective** if the willingness-to-accept exceeds the ICER
6	≈	↓	**Not cost-effective** (comparator dominates the new intervention)
7	↑	≈	**Not cost-effective** (comparator dominates the new intervention)
8	↓	≈	**Cost-effective** (new intervention dominates the comparator *i.e.* cost-saving)
9	≈	≈	**Not cost-effective** (new intervention and comparator are equal)

Abbreviation: **↑:** statistically significantly higher; **↓:** statistically significantly lower; **≈:** no statistical significant differences

Table 2: Detailed characteristics of the selected studies

First author, year, country	Screening or diagnosis	Comparator	Instruments for screening	Target population; Sample size	Perspective time horizon	Currency; price year	Outcomes measures	Discount rate	Incremental cost-effectiveness ratio (ICER)	Sensitivity analysis	Model	CHEERS	Funding source
Screening without biomarkers													
Saito E, 2014, USA [24]	Community based dementia screening	No dementia screening	History of daily living and activity, physical and neurological examination, MMSE	Healthy individual aged > 18 years, 93	Health care; lifetime	US dollar; NM	Cost saved	NM	Cost saving ($209 per patient)	No	Markov	14	NM
Banerjee S, UK, 2009 [25]	Early diagnosis	Usual care	A multi-disciplinary and interagency team	Patients in early state of dementia; 600,000 (all dementia people in the UK)	Societal; 10 years	GBP; 2007/2008	QALY	3.5%	£950 million savings over 10 years and 6,250 to 12,500 QALY gain per year	No	Mathematical model	19	NM
Dixon J, 2014, UK [26]	One-off dementia screening	No screening	Screening by clinical nurses and GPs (3:1 ratio) by MMSE	All 75 years of age in the UK and Wales	Societal; lifetime	GBP; 2012	Person diagnosed	3.5%	3,514 people could be diagnosed. Costs range are 3,649,794 (net cost) to 4,685,768 (net savings)	Deterministic	Mathematical model	21	Industry
Yu Su-Yeon 2014, South Korea [27]	Opportunistic screening	No screening	MMSE at public health centers	All ≥60 years of age in South Korea	Societal; 10 years	US dollar; 2010	QALY	5%	$24,150 to $35,661 per QALY based on age-group	Deterministic and probabilistic	Markov model	23	Govt.
Wolfs CA, 2009, The Netherlands [28]	Multi-component screening follwed by treatment	Usual care	Home visit and hospital visit and computed tomographic scan plus various blood test	Patients ≥55 years and suspicion of dementia or cognitive disorder; Intervention (n=137) and usual care (n=93)	Societal; 12 months	Euro; 2005	QALY, MMSE score, NPI score	-	€1,267/QALY, -€333/additional point decreese in MMSE, dominant by NPI score	Deterministic	No model	22	Govt.

Weimer DL, 2009 USA [29]	Early diagnosis and treatment	Late diagnosis and treatment	Early screening involve MMSE, neuropsychologist, psychologist, geriatrician and geriatric psychiatrist	Patients suffering from cognitive decline	Societal; lifetime	US dollar; 2006	Net benefit	3.5%	Cost beneficial (For example, net societal benefit is $94,000 for a 70 years old married women starting with MMSE score of 26)	Deterministic	Monte carlo simulation	22	Govt.
Biomarkers based screening													
McMahon PM, 2000,USA [30]	Functional neuroimaging plus standard diagnostic	Standard diagnosis	Visual SPECT, computed SPECT and MRI + DSC MRI	Hypothetical patients referred to AD center	Societal; 18 months	US dollar; 1998	QALY	3%	Both visual and computed SPECT was dominated by standard diagnostic procedure. For MR imaging + DSC MR imaging $479,500/QALY	Deterministic	Decision tree and Markov model	20	NM
McMahon PM, 2003, USA [31]	Functional neuroimaging plus standard diagnostic	Standard diagnosis	DSC MRI, FDG PET, Computed SPECT	Hypothetical patients referred to AD center	Societal; 18 months	US dollar; 1999	QALY	3%	Both FDG PET and computed SPECT was dominated by standard diagnostic procedure. For DSC MRI $598,800/QALY	Deterministic	Decision tree and Markov model	20	NM
Guo S, 2012, USA [32]	Biomarkers	Standard diagnosis	Florbetaben with PET	Pre-dementia (n=320) and dementia (n=680)	Societal; lifetime	US dollar; 2011	Survival, time of diagnosis, time in different stage of AD, caregivers time and QALY	3%	Dominant Pre-dementia (cost saving $13,018, QALY gain 0.27.) Dementia (cost saving $11,389, QALY gain 0.03)	Deterministic and probabilistic	DES	21	NM
Valcarcel-Nazco C et al, 2014, Spain [33]	Biomarkers	Standard diagnosis	Amyloid-β peptide, total tau and phosphorylated tau in CSF	≥60 years with MCI and ≥60 years with dementia; NM (hypothetical)	Healthcare; lifetime	Euro; 2013	Cases identified and treated correctly	3%	Dominant for MCI (cost saving €1833, accurate diagnosis 0.37). Cost-effective for dementia (cost saving €1134, accurate diagnosis 0.24)	Probabilistic	Decision Tree	23	Govt.
Handels RLH, 2015, The Netherlands [34]	Biomarkers	Current practice	CSF markers	MCI subject, 2000 (hypothetical)	Societal; lifetime	Euro; 2012	QALY	4% on health outcomes and 1.5% on cost	Dominant (cost saving €33,622, QALY gains 0.39)	Probabilistic	Decision tree and DES	23	Govt, non-Govt. and industry
Silverman et al 2002, USA [35]	Biomarkers	Current practice	FDG PET	Patients with symptoms of cognitive decline	Healthcare, 12 months	US dollar; NM	Cost per accurate diagnosis	-	Cost savings was $1,138 per accurate diagnose	Deterministic	Decision tree	20	Govt.

Moulin-Romsee G, 2005, Belgium [36]	Biomarkers	Standard diagnosis	FDG PET incorporated in current practice	NM	NM	Euro; NM	Cost per accurate diagnosis	No discount	Cost saving ranges from €623 to €6,110 per accurate diagnose	Deterministic	Decision tree	20	NM
Djalalov S, 2012, Canada [37]	Genetic screening	Current practice	Genetic screening for APO e4+	Amnestic mild cognitive impairment	Societal; 30 years	Can$, 2009	QALY	5%	Can$38,016/QALY	DSA and PSA	Markov	23	Government

Abbreviations: AD, Alzheimer's Disease; DES, Descrete Event Simulation; DSC, Dynamic Susceptibility Contrast–enhanced; FDG, Fluorodeoxyglucose; GBP, British Pound; GPs, General Physicians; MMSE, Mini Mental State Examination; MRI, Magnetic Resonance Imaging; NPI, Neuropsychiatric Inventory; PET, Position Emission Tomography; NM, Not Mentioned; QALY, Quality Adjusted Life Years; SPECT, Single Photon Emission Computed Tomography.

Table 3: Sensitivity and specificity of the screening strategies and their comparators in the selected studies

First author, country	Screening strategy			Comparator		
	Instrument	Sensitivity	Specificity	Instrument	Sensitivity	Specificity
Dixon J, UK [26]	MMSE	92/86%	99/92%	No screening	-	-
Yu SY, South Korea [27]	MMSE	81.8%	80.5%	No screening	-	-
	Diagnostic test	90%	90%			
McMahon PM, USA [30]#	MRI+DSCMRI	88%	96%	Standard examination	75%	87%
	Visual SPECT	50%	100%			
	Computed SPECT	90%	87%			
McMahon PM, USA [31]#	FDG PET	94%	72%	Standard examination	70%	73%
Guo S, 2012, USA [32]	FDG PET	90%	90%	Clinical Guidelines only	87%	59%
Valcarcel-Nazco C, Spain [33]##	Biomarker	81%	87%	Diagnostic criteria	87%	58%
Handels RLH, The Netherlands [34]	Perfect CSF biomarker	100%	100%	Current practice	77%	68%
Silverman DHS, USA [35]	FDG PET	91.5 ± 3.5%	70 ± 3%	Clinical evaluation	66 ± 17%	77 ± 23%
Moulin-Romsee G, Belgium [36]	FDG PET	90% to 96%	67% to 97%	Conventional method	84%	52.5%
Djalalov S, 2012, Canada [37]###	Genetic screening for APO e4+	100%	100%	Current practice	NM	NM

for mild Alzheimer diseases ## for MCI, ### for AMCI

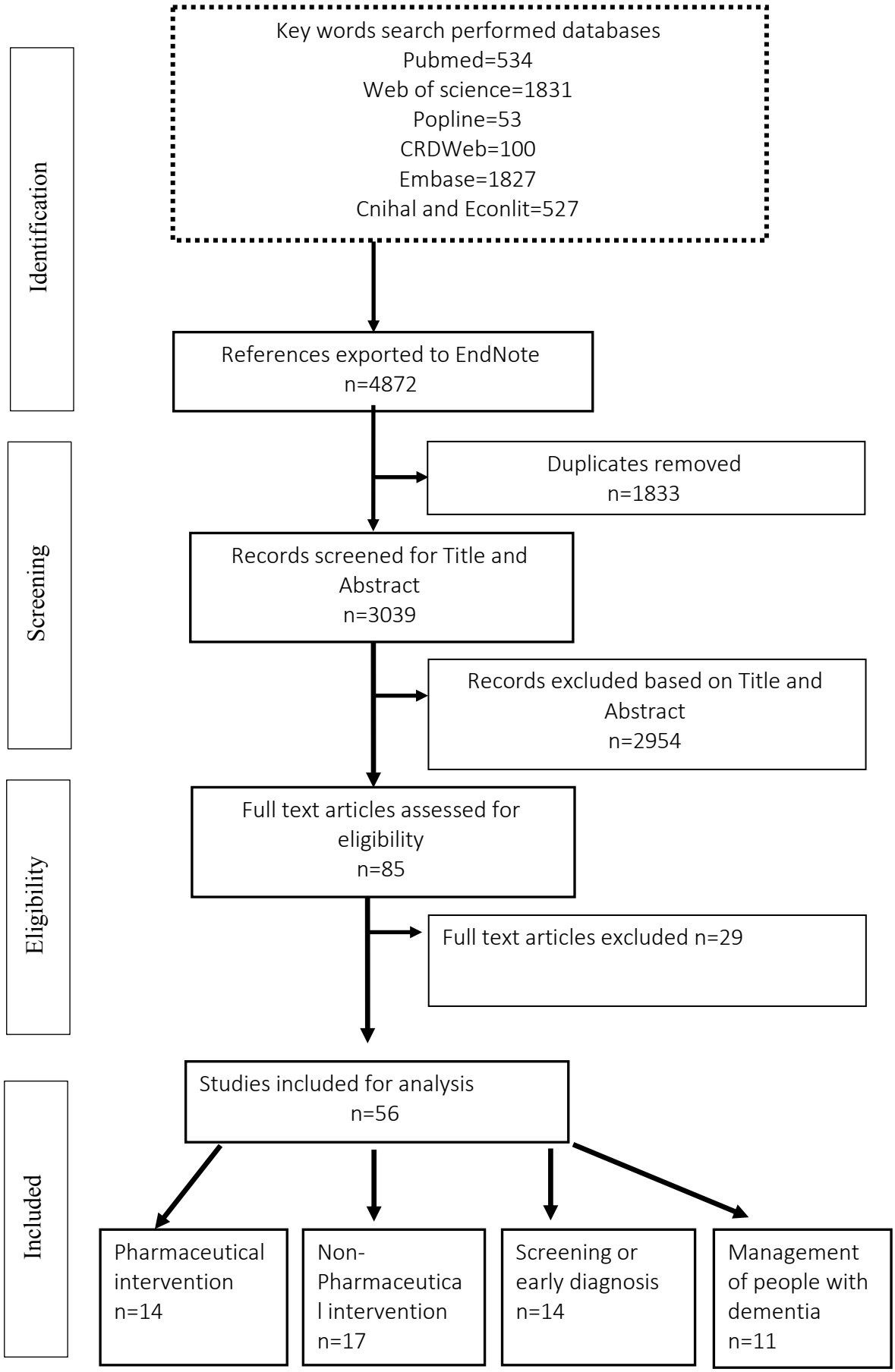
Abbreviations: AMCI, Amentic mild cognitive impairment; FDG, Fluorodeoxyglucose; MMSE, Mini Mental State Examination; MRI, Magnetic Resonance Imaging; PET, Position Emission Tomography; NM, Not Mentioned; QALY, Quality Adjusted Life Years; SPECT, Single Photon Emission Computed Tomography.

Table 4: Reported and evaluated cost-effectiveness of the identified screening interventions

First author, year, country	Effectiveness measure	Reported	Evaluation	Remarks on cost-effectiveness
Screening without biomarkers				
Saito E, 2014, USA [24]	Cost saved	Cost saving	95% CIs was not presented	Unknown due to lack of information
Banerjee S, UK, 2009 [25]	QALY (DEMQOL)	Dominant	95% CIs were not presented for either costs or QALYs.	Unknown due to lack of information
Dixon J, 2014, UK [26]	Person diagnosed	Cost neutral	No significant differences in costs were observed	Unknown due to no agreed cost-effective threshold value
Yu SY 2014, South Korea [27]	QALY	Potentially cost-effective	95% CIs were not presented for either costs or QALYs. For ≥65 years age group, CEAC showed 53% probability to be cost-effective at $20,536 WTP (GDP of south Korea in 2010) (65% by NICE threshold)	Not cost-effective
Wolfs CA, 2009, The Netherlands [28]	QALY, MMSE, NPI	Cost-effective for QALY, indecisive for MMSE, Dominant for NPI	95% CIs were not presented for costs, QALYs, MMSE or NPI. CEAC showed 72% probability to be cost-effective at €45,000 WTP for QALY (67% by NICE threshold)	Not cost-effective
Weimer DL, 2009 USA [29]	Cost	Cost beneficial	95% CIs was not presented	Unknown due to lack of information
Biomarkers based screening				
McMahon PM, 2000, USA [30]	QALY (HUI)	Not cost-effective	95% CIs were not presented for either costs or QALYs.	Unknown due to lack of information
McMahon PM, 2003, USA [31]	QALY (HUI)	Not cost-effective	95% CIs were not presented for either costs or QALYs.	Unknown due to lack of information
Guo S, 2012, USA [32]	QALY	Dominant for both pre-dementia and dementia patients.	No significant differences in costs or QALY gains were observed for pre-dementia patients. For dementia patients, cost differences was insignificant but QALY gain was significant.	Not cost-effective for pre-dementia patients. Cost-effective for dementia patients.
Valcarcel-Nazco C, 2014, Spain [33]	Accurate diagnosis	Dominant for both MCI and dementia	Significant differences in costs and effects were observed for MCI patients. For dementia patient, cost differences are not significant but effect differences are significant. CEAC showed 80% probability to be cost-effective at €1,000 WTP for dementia patient (98% by NICE threshold)	Cost-effective (Dominant) for MCI, cost-effective for dementia
Handels RLH, 2015, The Netherlands [34]	QALY	Dominant	Costs are significantly lower and QALY gains are significantly higher	Cost-effective (Dominant)
Silverman DHS, 2002, USA [35]	Cost per accurate diagnosis	Cost savings	95% CIs were not presented for costs difference	Unknown due to lack of information
Moulin-Romsee G, 2005, Belgium [36]	Cost per accurate diagnosis	Cost savings	95% CIs were not presented for costs difference	Unknown due to lack of information
Djalalov S, 2012, Canada [37]	QALY (HUI2)	Cost-effective	95% CIs were not presented for either costs or QALYs. CEAC showed 66% probability to be cost-effective at Can$100,000 WTP per QALY (50% by NICE threshold)	Not cost-effective

Abbreviations: CEAC, Cost-effectiveness acceptability curve; DEMQOL, Quality of life dementia; EQ-5D, HUI, Health Utility Index; Euroqol five dimensions; MMSE, Mini Mental State Examination; MCI, Mild Cognitive Impairment; NPI, Neuropsychiatric Inventory; QALY, Quality Adjusted Life Years; WTP, Willingness-to-pay

Figure 1: A flow chart for selection of articles



Annex 1: Detailed search history in databases with keywords

Pubmed

("economic evaluation"[All Fields] OR "cost-benefit analysis"[MeSH Terms] OR "cost-effectiveness"[All Fields] OR "cost-benefit analysis"[MeSH Terms] OR "cost-benefit analysis"[MeSH Terms] OR "cost benefit"[All Fields] OR "cost utility"[All Fields] OR "cost-utility"[All Fields]) AND ((("dementia"[MeSH Terms] OR "dementia"[All Fields]) OR "dementia"[MeSH Terms]) OR "mild cognitive impairment"[All Fields]) AND (("2000/01/01"[PDAT] : "2015/12/31"[PDAT]) AND English[lang])

= 534

CRDWeb

((dementia)) and ((Economic evaluation:ZDT and Bibliographic:ZPS) OR (Economic evaluation:ZDT and Abstract:ZPS)) IN NHSEED FROM 2000 TO 2016

=100

EMBASE

1. 'dementia'/exp OR dementia

2. 'cost effectiveness' OR 'cost utility' OR 'cost benefit analysis' OR 'economic evaluation'
3. 'mild cognitive impairment':ab
4. #1 OR #3
5. 'cost consequence analysis'
6. #2 OR #5
7. #4 AND #6
8. #7 AND (2000:py OR 2001:py OR 2002:py OR 2003:py OR 2004:py OR 2005:py OR 2006:py OR 2007:py OR 2008:py OR 2009:py OR 2010:py OR 2011:py OR 2012:py OR 2013:py OR 2014:py OR 2015:py)

=1827

Web of science

1. **TOPIC:** Dementia
2. **TOPIC:** Mild cognitive impairment
3. **TOPIC:** Alzheimer
4. **TOPIC:** Vascular dementia
5. **TOPIC:** Parkinson's disease
6. 1 OR 2 OR 3 OR 4 OR 5
7. **TOPIC:** (cost effectiveness) OR **TOPIC:** (cost-effectiveness analysis) OR **TOPIC:** (cost-effectiveness) OR **TOPIC:** (cost utility analysis) OR **TOPIC:** (cost-utility analysis) OR **TOPIC:** (cost benefit) OR **TOPIC:** (cost-benefit) OR **TOPIC:** (economic evaluation)
8. 6 AND 7 (Refined by: Publication Years (2000 to 2015))

9. 8 (Refined by: Language (English)

= 1831

1. Dementia
2. AB dementia
3. AB dementia OR mild cognitive impairment
4. Cost effectiveness
5. Cost benefit analysis
6. Cost utility analysis
7. Cost-utility analysis in healthcare
8. Economic evaluation
9. Cost consequences analysis in health economics
10. 4 OR 5 OR 6 OR7 OR 8 OR 9
11. 10 AND 3
12. 11 (limiters- 20000101-20151231)

= 527

Popline

1. ((((Title:dementia) OR (Title:alzheimer)))) AND ((Language:English) AND (Publication Year:[2000 TO 2015]) AND (Peer Reviewed:1) AND (Journal Article:1))

2. ((((Title:cost effectiveness analysis) OR (Title:cost utility analysis) OR (Title:economic evaluation) OR (Title:cost benefit analysis)))) AND ((Language:English) AND (Publication Year:[2000 TO 2015]) AND (Peer Reviewed:1) AND (Journal Article:1))

3. 1 OR 2

= 53