Where are we at with model-based economic evaluations of interventions for dementia? A systematic review and quality assessment

Kim-Huong Nguyen (PhD)

Centre for Health Services Research, The University of Queensland, PAH Campus, Brisbane, Australia

NHMRC's Partnership Centre on Dealing with Cognitive and Related Functional Decline in Older People, The University of Sydney, Sydney, Australia

Tracy Comans (Associate Professor)

Centre for Health Services Research, The University of Queensland, PAH Campus, Brisbane

NHMRC's Partnership Centre on Dealing with Cognitive and Related Functional Decline in Older People, The University of Sydney, Sydney, Australia

Colin Green (Professor)

Medical School, The University of Exeter, St Luke's Campus, Exeter, United Kingdom

Corresponding author: Kim-Huong Nguyen

Email: kim.h.nguyen@uq.edu.au

Centre for Health Services Research, Building 33, PAH, Woolloongabba QLD 4102

Abstract:

Objective: To identify, review and critically appraise model-based economic evaluations of all types of interventions for people with dementia and their carers.

Design: A systematic literature search was undertaken to identify model-based evaluations of dementia interventions. A critical appraisal of included studies was carried out using guidance on good practice methods for decision-analytic models in health technology assessment, with a focus on model structure, data, and model consistency.

Setting: Interventions for people with dementia and their carers, across prevention, diagnostic, treatment and disease management.

Results: We identified 67 studies, with 43 evaluating pharmacological products, 19 covering prevention or diagnostic strategies, and five studies reporting non-pharmacological interventions. The majority of studies use Markov models with a simple structure to represent dementia symptoms and disease progression. Half of all studies reported taking a societal perspective, with the other half adopting a third-party payer perspective. Most studies follow good practices in modelling, particularly related to the decision problem description, perspective, model structure and data inputs. Many studies perform poorly in areas related to reporting of pre-modelling analyses, justifying data inputs, evaluating data quality, considering alternative modelling options, validating models and assessing uncertainty.

Conclusions: There is a growing literature on model-based evaluations of interventions for dementia. The literature predominantly reports on pharmaceutical interventions for Alzheimer's disease, but there is a growing literature for dementia prevention and non-pharmacological interventions. Our findings demonstrate that decision-makers need to critically appraise and understand the model-based evaluations and their limitations to ensure they are used, interpreted and applied appropriately.

1 Introduction

Economic considerations play an important role in financing and service provision for the diagnosis of dementia, and the treatment and care available for people with dementia. Economic evidence is presented in many formats, including cost of illness studies, within-trial cost effectiveness analyses, and model-based economic evaluations. We focus here on the latter and the use of decision-analytic models to assess the cost effectiveness of interventions for dementia. The use of model-based evaluations is attractive in a health policy context as it provides an explicit framework to systematically synthesise evidence available from multiple sources, and to estimate and compare longer term outcomes (e.g. life-expectancy, quality adjusted life years) and inputs (costs) for alternative interventions and/or treatment pathways. Model-based economic evaluations are commonly used to support decision-makers in a health policy context, where assessment of effectiveness and cost effectiveness in a policy-relevant context is required.

The usefulness and validity of the evidence available from model-based evaluations is dependent on the methods and data used to estimate and report the expected costs, outcomes and cost effectiveness of interventions. Decision-makers must assess the quality of the model-based evaluations and be familiar with the pros and cons of the evaluations when interpreting findings and using them to inform health policy. There is a growing literature on the use of decision-analytic models in the dementia context, including a number of high-quality systematic reviews. While each of the available reviews has focused on a distinctive segment of the care continuum, i.e. diagnosis (Handels et al., 2014), pharmacological treatment (Cohen and Neumann, 2008; Green, 2007; Green et al., 2011; Gustavsson et al., 2009; Hernandez et al., 2016; Oremus, 2008), non-pharmacological treatment (Jones et al., 2012) and interventions for dementia carers (Sopina and Sørensen, 2018), they shared similar objectives and a number of common themes. A majority of reviews aimed to understand the analytical approaches used to evaluate cost effectiveness through providing a detailed description of study characteristics, methodologies, model elements and main findings. Common findings across reviews have been that (i) many studies failed to present sufficient information to fully assess methodologies used, thus increasing the uncertainty surrounding the interpretation and generalisation of the results, (ii) considerable variation exists in methodological assumptions applied and target populations among studies, preventing a meaningful comparison of cost effectiveness results, and (iii) evaluations should use better quality and up-to-date data, particularly outcome data that better reflect the breadth of dementia symptoms. Prior reviews have taken a varied approach to the critical appraisal and assessment of methods, with few providing a comprehensive quality assessment of modelling methods and reporting practices against published decision-analytic models in dementia. None of the prior reviews have reported across the full continuum of dementia care, from

prevention, screening, early detection and diagnosis, to treatment and care programs for people with dementia and their carers.

In the current review we present a comprehensive overview of contemporary model-based economic evaluations of dementia interventions. We systematically review the literature reporting decision-analytic models used in economic evaluation of interventions for people with dementia and their carers (informal and family carers) across prevention, early detection, diagnosis and treatment and care strategies, including all types of dementia. Additionally, we critically assess the modelling methods used and reporting quality against good practice guidance on methods for decision-analytic modelling in health technology assessment, across key domains of model structure, data, and model consistency.

2 Methods

2.1 The systematic search and review of the literature

We have undertaken the systematic review in line with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). The systematic review was registered in PROSPERO international prospective register of systematic reviews (ID: 42017059154) in April 2017. The search strategy and keywords are presented in Table A1.1 (Appendix 1, published as supplementary material online attached to the electronic version of this paper at https://www.cambridge.org/core/journals/international-psychogeriatrics). The search was applied to literature available up to February 2018, in the electronic databases of MEDLINE, CINAHL, Cochrane Library, Embase, PsycINFO, Econlit, HTA database, CEA registry and PubMed. Reference lists of identified publications were checked for additional articles meeting the inclusion criteria.

Inclusion criteria (Table A1.2, Appendix 1) were (i) evaluations of interventions for people with dementia and/or their carers, covering surveillance, screening, early diagnosis, treatment, management and care, where (ii) economic evaluation used a decision-analytic model-based framework. The identification of studies was undertaken by two independent reviewers, who initially screened the titles and abstracts for eligibility. A study was included if both reviewers agreed that it met all the inclusion criteria. Dissimilarities in reviewers' conclusion were resolved by discussion and/or examining the full article. Full texts of all included studies were reviewed independently by both reviewers. Resulting citations and reference lists in these reviewed studies, studies that published systematic reviews and meta-analyses were examined to ensure that no eligible studies were missed. Both reviewers systematically and independently documented the properties of the included studies.

2.2   Quality assessment of decision-analytic models

For quality assessment of included studies, we used the framework for assessment of good practice in decision-analytic models in a health technology assessment context, presented by Philips et al., (2006), (hereafter, referred to as the Philips checklist). The checklist is a broad framework which assesses three key dimensions of modelling: structure, data (including uncertainty) and consistency. It includes 58 items; with 9 distinct areas for model structure (23 items), 11 areas for data (30 items), and 2 areas for model consistency (5 items). See Table A1.3 (Appendix 1) for detail on the Philips checklist. This checklist was used because of its relevance to the HTA context (Philips et al., 2004), and as it is considered to be consistent with the ISPOR principles of good practice for decision-analytic modelling for health care evaluation (Caro et al., 2012; Weinstein et al., 2003).

The quality assessment using the Philips checklist was conducted by two independent reviewers. Differences between the two reviewers' responses for each item were resolved by discussion and/or consultation with a third reviewer. We followed the questions and examples presented in Philips et al., (2004, 2006) to score a response for each item (Y=Yes, N=No, U=Unsure, NA=Not Applicable). The scoring decisions for a study were also influenced by the available literature up to that point in time, particularly with respect to other published models, data sources, and best practice guidelines. The standards for assessment of uncertainty increased over time as the number of published studies using competing models increased. For instance, the lack of discussion on alternative models was not heavily penalised prior to 2000 (before the good practice guidelines released and widely adopted in the mid-2000s) while studies published from late 2000s required more comprehensive assessment of uncertainty.

Many model-based evaluations of pharmacological treatments for Alzheimer's disease (AD) have been subject to critical appraisal in the prior review by Green (2007) which reviewed literature reported up to 2005. For this reason, and for pragmatic reasons, we use this prior review as a point of departure and limit the quality assessment of studies in this category to those published from 2005 onwards. All studies identified on prevention, diagnosis, and non-pharmacological interventions are subject to quality assessment in the current review.

3   Results

3.1   Basic properties of studies included in the systematic review

We identified 67 studies that met the inclusion criteria (evaluating a dementia intervention using decision-analytic model/s). See Figure A2.1 (Appendix 2, published as supplementary material   online   attached   to   the   electronic   version   of   this   paper   at

https://www.cambridge.org/core/journals/international-psychogeriatrics) for detail on screening, identification and selection of studies and Tables A2.1 (Appendix 2) for basic characteristics of the studies identified by intervention type, target population, types of decision-analytic models and their major features.

Two thirds of identified studies (42/67) were published after 2005 (see Figures 1 and A2.2, Appendix 2). All but three studies (Fuh and Wang, 2008; Hu et al., 2015; Yang and Chen, 2016) were conducted in OECD countries. More than half of the studies had some degree of industry involvement (e.g. authors received direct funding from industry or were affiliated with industry-funded projects or groups); ten studies did not declare any information about industry involvement.

INSERT FIGURE 1

In terms of the perspective/s stated and employed in included studies, in many cases the perspective is dictated by the research context (and decision problem) being addressed. For example, specific reimbursement-setting studies commonly follow recommendations, or reference case methods, for regulatory agencies, for government subsidy or insurance approval. Approximately 40% of studies applied a societal perspective, in which both (formal) health and care service costs and the opportunity cost of informal care (provided by family carers) were accounted for, and another 40% used a narrower health system perspective (including Third Party Payer, Department of Health). A small number of studies (10/67) used more than one perspective, often the combination of societal and health system (Figure 1).

Evaluations of pharmacological treatments accounted for 43 of the 67 included studies. A majority of studies (40/43) assessed the three cholinesterase inhibitors (donepezil, galantamine and rivastigmine) and a receptor antagonist (memantine), which have demonstrated modest effectiveness for the symptomatic treatment of Alzheimer's disease (Birks, 2006). Three studies evaluated tacrine (Henke and Burchmore, 1997), olanpazine (Kirbach et al., 2008) and an immunotherapy (Aβ42) (Yang and Chen, 2016). In the majority of these studies the decision context was treatment for people with a diagnosis of Alzheimer's disease (42/43); two studies covered dementia with Lewy bodies (Gustavsson et al., 2009) and vascular dementia (Wong et al., 2009). The main comparator in these studies was standard care or no treatment (35/43). None compared a pharmacological agent with a non-pharmacological intervention.

Two studies (2/67) evaluated a hypothetical disease modification treatment for people with mild cognitive impairment and pre-dementia (Budd et al., 2011; Skoldunger et al., 2013), and three other studies evaluated non-pharmacological interventions: a family-based program for mild AD (Martikainen et al., 2004), deep brain stimulation for mild AD (Mirsaeedi-Farahani et al., 2015) and a connected health care program for mild dementia (Dowd et al., 2018).

We identified 19 studies evaluating strategies for prevention (Zhang et al., 2011), early assessment (Getsios et al., 2012; Yu et al., 2015), screening (Djalalov et al., 2012; Retchin and Hillner, 1994) or diagnosis of dementia (14/19). Diagnosis studies covered a large variety of strategies, from imaging (magnetic resonance imaging, positron emission tomography, etc.) to biomarker (Aβ41, cerebrospinal fluid) detection. In these studies, the main target populations were elderly people or those with mild cognitive impairment.

Almost all studies used a cost utility framework (48/67), in which the primary outcome was quality adjusted life years, or a cost effectiveness analysis (15/67). Four studies used either cost benefit analysis or evaluated either outcomes or cost only. In cost effectiveness studies, the final outcomes were measured by clinical endpoints such as life years, the number of accurate diagnoses, and the number of years before full-time care or before institutionalisation.

The most commonly used modelling framework was a Markov state transition model (52/67), followed by a decision tree approach (12/67). Four studies combined a decision-tree (for diagnosis) with either a Markov or a discrete event simulation model (for dementia progression). Six studies used a discrete event simulation (DES) framework, four of which were based on the model structure developed in (Getsios et al., 2010). More than two-thirds of all studies evaluating pharmacological treatments applied a previously published model-based framework (see Figure 2). For instance, 16 studies adopted the model published by Neumann et al., (1999), and 13 studies were based on the Assessment of Health Economics in Alzheimer's Disease (AHEAD) model (Caro et al., 2001). Approximately a half of studies (10/19) evaluating prevention, screening and diagnosis of dementia developed their own original models, while the rest translated previously published models into other settings.

INSERT FIGURE 2

Neumann et al., (1999) estimated the cost effectiveness of donepezil, compared to no treatment, using a Markov transition model with four health states: mild, moderate, severe and dead. Three AD severity states/stages were used based on the Clinical Dementia Rating (CDR) scale, and patients were either living in the community or in a nursing home. The underlying AD progression probabilities were estimated from Consortium to Establish a Registry for Alzheimer's Disease (CERAD) data, which is a longitudinal database of 1,145 dementia patients in the United States between 1986 and 1995 (Morris et al., 1989). Data on the impact of donepezil on progression and mortality were sourced from the 24-week randomized, double-blinded, placebo-controlled clinical trial that investigated the impact of donepezil on patients with mild or moderate AD (Rogers et al., 1998). Subsequent studies applying the Neumann model kept the four-state model structures and the underlying AD progression while updating treatment impacts, mortality risks, costs and outcomes using the

relevant settings. Model durations (e.g. six months, two years, 10 years to life-time) and cycle lengths (e.g. one month, six weeks, one year) were applied to suit the study objectives.

The AHEAD model (Caro et al., 2001) applied a different interpretation of disease progression, using "need for full-time care (FTC)" as the main end-point to classify three health states: not requiring FTC, requiring FTC and dead. FTC was defined as the consistent requirement of caregiving for a significantly large amount of time per day, which was indicated by the need for institutional care (in a health-related facility). The likelihood of FTC at any point in time was estimated using patient characteristics observed in the Predictors Study Cohort of 236 patients with AD observed for a period of up to seven years in the United States (Stern et al., 1997). Patient characteristics included the (modified) Mini Mental Sate Examination (m-MMSE) scores, extrapyramidal signs, psychotic symptoms, disease duration and whether or not the patient had a young onset AD. Subsequent studies using the AHEAD model continued with the three-state model and its progression probabilities while revising treatment impacts, mortality risks, costs and outcomes. In some occasions the definition of FTC was revised to "institutionalisation". The majority of these studies used a 5-year or a 10-year horizon. The cycle length of one month was explicitly reported in five studies (out of 13).

A small number of alternative (dementia progression) models to AHEAD and Neumann incorporate more disease characteristics than just cognitive impairment. Some combinations of physical function (e.g. using Activities of Daily Living, instrumental Activities of Daily Living or Severe Impairment Battery instruments), care requirement (e.g. full-time care) and/or residence (community versus nursing home), or behaviour symptoms (e.g. using Neuropsychiatric Inventory) were used in five Markov transition models to define health states (nine to 12 health states), and in six DES models to initiate events and assign attributes to patients (see Table A2.1, Appendix 2 for more details).

Similar to other areas of disease modelling literature, models in this dementia literature relied upon disease progression data estimated from observational studies, many of which date from the 1980s and 1990s (such as CERAD, Kungsholmen or Predictors Study Cohort). Mortality risks were often assumed either constant over time in studies published in the early half of the literature, or age- and gender-specific (based on country-specific all-cause mortality risk) in the recent half of the literature. In a small number of cases, a hazard ratio due to dementia was added to the base mortality (Budd et al., 2011; Gagnon et al., 2007; Kirbach et al., 2008; Lachaine et al., 2011; Lee et al., 2017; Thibault et al., 2015; Touchon et al., 2014). Impacts of treatment were extrapolated from short-duration clinical trials (e.g. six months) to the longer term (e.g. five or ten years) to fit the model duration. All studies grouped costs into the three main categories: intervention costs; direct medical expenditure (e.g. doctor visits, monitoring tests, hospitalisation); and paid (formal) care cost (e.g. community-based social and nursing

services, nursing home fee). In cost utility analysis models, most of the studies sourced quality of life weights from valuation studies that used one of two generic instruments: the Health Utility Index (developed with a US population using HUI2, in Neumann et al., 1998 and 1999 or a Japanese population using HUI3 in Ikeda et al., 2001) or the EuroQol, developed with a mostly European population (including the UK in Jonsson et al., 2006, Spain in Lopez-Bastida et al., 2006, and Sweden in Ekman et al., 2007). A dementia-specific quality of life weight (e.g. DEMQOL developed in Rowen et al., 2012) was used in one study (Dowd et al., 2018) and the quality of life instrument for people with Alzheimer's disease (QoL-AD) was used in a small number of studies, in which the quality of life scores were mapped into a utility scale (e.g., Gagnon et al., 2007; Rive et al., 2010). Out of 36 studies using a societal perspective, 32 incorporated the informal care cost, approximated by total number of hours of caregiving and hourly opportunity cost, but only seven studies accounted for carers' quality of life.

3.2   Quality assessment of decision analytic models using the Philips checklist

We conducted quality assessment using the Philips checklist for 55 (out of 67) studies: all 19 studies evaluating interventions/strategies for prevention, screening, early detection and diagnosis of dementia; five studies evaluating non-pharmacological interventions and programs for people with dementia and their carers; and 31 studies (published after 2005) evaluating pharmacological treatments.

Overall, "Structure" (23 items) was the area where a majority of studies scored more positively ("Yes", providing sufficient details) than negatively ("No", insufficient information or unclear evidence), while "Data" (30 items) and "Consistency" (5 items) showed greater limitations in reporting quality. It is clear that some aspects of modelling and reporting practices have improved overtime, evidenced by the rising percentage of items scored positively ("Yes") in the Data domain over time (see Figure 3). However, information/evidence provided in certain domains was still inadequate to meet the checklist criteria, including comprehensive assessments of uncertainty (e.g. covering all four principal types of uncertainty: methodological; structural; heterogeneity; and parameter), validation of internal consistency and calibration against independent data. Only two studies scored "Yes" in more than 85% of all items; both of them were comprehensive (long) health technology assessment (HTA) reports submitted to NICE (Bond et al., 2012; Loveman et al., 2006). They both explicitly described a systematic approach to the literature search and thorough consideration and selection of models, parameters, assumptions, and pre-modelling methods, which were commonly omitted in other publications. We present the full quality assessment findings in Table A2.2 (Appendix 2), and a summary of key observations on the changes of modelling and reporting quality over time in Table 1.

INSERT FIGURE 3

INSERT TABLE 1

Structure: Most studies clearly stated their objective and scope, the options under evaluation, and the analytical perspective/s and outcomes. The evidence of model types, disease states and pathways, evaluation horizon and cycle length were described clearly. However, those modelling decisions were often not accompanied by an explicit rationale and justification, such as why competing model structures were not considered, or why other comparator/s should be excluded. Studies that replicated existing models in another setting were more likely to omit rationale for comparators, duration and cycle length, compared to studies that developed original models. Unexpectedly, the latter failed to justify their (newly developed) model choices more often than the former (22/24 versus 22/31).

Data: Modelling quality has improved over time. Strong improvements were observed in the reporting of baseline data, quality of life weights and analysis of parameter uncertainty. While the adoption of good modelling and reporting practices is also reflected in critical areas such as selecting data sources and using appropriate techniques for pre-modelling analyses, a majority of studies simply reported the modelling decision/approach without providing a rationale. For instance, while more than 80% of all studies discussed, reported and referenced the model data properly (items D11, D2c1, D31, D33, and D35), fewer than 20% provided sufficient justification for its use (items D12, D2a4, D2b2, D2b3, D2c3, D32, and D42). Surprisingly, only a third of studies clearly reported their systematic approach to data identification and selection, and 25% provided any quality assessment of the data used in their models (including risk of bias). With regard to uncertainty analyses, while the assessment of parameter uncertainty improved significantly over time (from <50% in studies published before 2005 to >90% in those published after 2010), little improvement was observed in the assessment of methodological and heterogeneity uncertainties (approximately 20-25% of all studies).

Consistency: This domain consists of fewer assessment items than the other domains: one internal and four external consistency checks. Out of 55 studies, two explicitly stated that the mathematical logic of their models was examined (both of them were HTA reports, Bond et al., 2012 for the UK and Bermingham, 2014 for Canada) and another two reported that they calibrated their models using external data (Gagnon et al., 2007; Gustavsson et al., 2009). All studies drew consistent conclusions (of cost-effectiveness) from the model results and more than half (32/55) compared their findings with other published models. Only a few studies (4/55) reported "counter-intuitive results" and attempted some explanation. Overall, we did not see strong evidence of improvement in the quality of modelling and reporting of evidence in the "Consistency" domain across the 30-year period of this literature review.

4 Discussion

Decision-analytic models have been used widely in the economic evaluation of dementia interventions in the past three decades. While this approach has been attractive in a health policy context, the validity of (cost-effectiveness) findings depends entirely on the modelling quality. That is, how well the model addresses the decision problem through modelling of disease progression and treatment pathways, and capturing the relevant costs and outcomes, and whether or not the model can be validated against external data.

As expected, recent studies (e.g., those published after 2010) showed significant improvement in terms of modelling quality and reporting standard, which may be a result of the introduction of guidelines on decision-analytic modelling and good practice checklists (e.g. CHEERs, Philips checklist). Despite this positive trend, there remain areas requiring substantial improvement. Some would be relatively simple, such as clear documentation of model characteristics (model type, number of health states, model duration, cycle length, and half-cycle correction), data sources and assumptions underlying the model structure. Others require substantial improvements in modelling practice, such as a systematic approach to literature searches and selection of data sources and alternative models (including assessment of data quality), appropriate application and rationale for pre-data modelling methods, and comprehensive assessment of model validity and consistency using internal and external data sources. We acknowledge that the critical appraisal using the Philips checklist was a subjective process and may be more reflective of the quality of reporting than the quality of modelling. Studies with a higher percentage of items scored positively ("Yes") did not necessarily have higher quality modelling. For instance, many items were easier to report and to evaluate (i.e. whether or not there was sufficient information): evaluation objective, target population, interventions, model type, number of health states, duration and cycle length. Additionally, studies that translated existing modelling frameworks into new settings/evaluations (with minimal revisions) could often refer to the original models (for pre-modelling and data sources), which may have freed up space in the main text for content on rationale and discussion surrounding modelling methods. Our approach to critical appraisal, and use of the Philips checklist, does not allow a detailed assessment of the appropriateness of the model structure, assumptions and data sources. Nor was it possible to make meaningful direct comparisons of models against each other, which is only feasible through replicating the models used in those studies. Despite these limitations, the Philips checklist remains useful as a self-assessment tool for model developers to improve both modelling and reporting quality.

In this review, we found the majority of models use a crude representation of dementia and fail to account for distinguishing features of the condition, such as behavioural and

psychological symptoms and functional performance, all of which affect care needs and quality of life. This finding supports other interpretations of the literature in this field (Cohen and Neumann, 2008; Green et al., 2011). A few recent studies (Gagnon et al., 2007; Getsios et al., 2010; Hartz et al., 2012) incorporated, at least conceptually, a combination of domains to better reflect the impact of the underlying biological process on disease states and progression. Recent contributions to the modelling literature in this area, Green and Zhang (2016) have argued for, and provided empirical evidence to support, modelling of Alzheimer's disease progression using a multi-domain descriptive system, for example using cognition, behaviour and function as core symptom domains. Such recommendations respond to earlier findings in the literature on methods for model-based evaluations, and it is suggested that this practice becomes an integrated part of future modelling initiatives and developments for assessment of the progression of dementia in an HTA setting.

While evaluation of pharmacological agents accounted for two thirds of all studies, attention on non-pharmacological interventions, prevention and diagnosis has been consistent over time, although infrequent, and possibly on the rise over recent years. The dominance of drug evaluation for people with AD strongly mirrors both the target population of pharmacological agents and the distribution pattern of dementia prevalence by types of dementia, where Alzheimer's disease accounts for 50%-75% of dementia cases (Prince et al., 2014). It also reflects the environment in which many economic evaluations are conducted: prior to approval of public funding for wide-scale implementation of interventions, most of which were pharmacological. The question remains whether or not a pharmacological-based evaluation model can be reapplied directly to evaluating prevention, diagnosis or non-pharmacological interventions. This current literature indicates that it is a reasonable approach: a half of non-pharmacological studies sourced their models from pharmacological evaluations. For instance, a prevention strategy that reduced the dementia incidence rate (Zhang et al., 2011) or a more precise diagnosis method that led to a higher number of patients correctly diagnosed at early stage (Biasutti et al., 2012; Djalalov et al., 2012; Getsios et al., 2012; Yu et al., 2015) or a care program (for the dementia dyads) impacting on the disease progression (Dowd et al., 2018; Martikainen et al., 2004) can all share a disease progression model where subsequent costs and QALYs are accumulated and compared. This approach can potentially allow for the comparison of all interventions along the continuum of dementia care from prevention to active disease management. It will also facilitate a better understanding of the impacts and potential trade-offs between interventions, thus offering decision-makers an opportunity to consider low-value versus high-value strategies, and the possibility of combining complementary therapies to suit patient settings and preferences. This is particularly relevant in countries where health and social services are both funded centrally through tax and/or social insurance and where the primary goal of HTA is ensuring efficient

allocation of scarce resources. A caveat here is that the feasibility of this approach rests entirely on the quality of the core models that can successfully represent the disease progression patterns, capture relevant dementia symptoms and outcomes (of both person with dementia and their carers), and account for important disease modification factors.

Our last observation relates to the existing capacity to perform cross (external) validation and convincingly draw a cost-effectiveness conclusion. It may be that cost effectiveness is evidenced when many studies, even when using different modelling methods, arrive at a similar conclusion. For instance, the cost effectiveness of memantine was established in four different studies/models (Jones et al., 2004; Jonsson, 2005; Rive et al., 2010; Hu et al., 2015) and multiple reapplications of those models in other settings. A similar observation can be seen for donepezil with its economic value confirmed by at least six different studies/models, (Neumann et al., 1999; Jonsson et al., 1999; O'Brien et al., 1999; Teipel et al., 2007; Getsios et al., 2010; Peters et al., 2013). While this consistency could be interpreted as a source of strong evidence of external validity for those models/evaluations, it may be that the direction of the evidence correlates with the source of funding/sponsor. Many of these studies have received direct funding from industry, and in many instances may reflect the economic evaluation models that were likely to have been submitted by industry for subsidy consideration (reimbursement purposes). Studies conducted independently of industry, or for HTA authority, showed a less optimistic conclusion regarding the economic value of some interventions. This observation is supported by previous authors (Bond et al., 2012; Peters et al., 2013). In particular, Bond et al., (2012) stated that when using more conservative assumptions on elements with high uncertainty, the (estimated) incremental cost effectiveness ratio (using industry submission models) could be at the margins of what would normally be considered cost-effective. Some consideration of this issue may be useful in future research on the validity and consistency of results and conclusions in modelling methods.

5 Conclusion

We identify an expanding literature on model-based evaluations of interventions for dementia. The literature is predominantly reporting on pharmaceutical interventions for Alzheimer's disease, but there is a growing literature for dementia prevention and non-pharmacological interventions. We provide a comprehensive summary and a critical review of the literature, finding that reporting of decision-analytic models has improved with the recent introduction of checklists and guidelines. However, we find some important aspects of model reporting still require substantial improvement, including justification of key model components and data selection. Findings here clearly demonstrate that decision-makers using model-based evaluations of dementia care need to critically appraise and understand the methods and data

applied, and the limitations of any evaluation to ensure they are used, interpreted and applied appropriately.

References

1. Bermingham, S., 2014. The Appropriate Use of Neuroimaging in the Diagnostic Work-Up of Dementia. Ont. Health Technol. Assess. Ser. 14, 1–67.
2. Biasutti, M., Dufour, N., Ferroud, C., Dab, W., Temime, L., 2012. Cost-Effectiveness of Magnetic Resonance Imaging with a New Contrast Agent for the Early Diagnosis of Alzheimer's Disease. PLoS ONE 7, e35559. https://doi.org/10.1371/journal.pone.0035559
3. Birks, J.S., 2006. Cholinesterase inhibitors for Alzheimer's disease, in: The Cochrane Collaboration (Ed.), Cochrane Database of Systematic Reviews. John Wiley & Sons, Ltd, Chichester, UK. https://doi.org/10.1002/14651858.CD005593
4. Bond, M., Rogers, G., Peters, J., Anderson, R., Hoyle, M., Miners, A., Moxham, T., Davis, S., Thokala, P., Wailoo, A., Jeffreys, M., Hyde, C., 2012. The effectiveness and cost-effectiveness of donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (review of Technology Appraisal No. 111): a systematic review and economic model. Health Technol. Assess. 16. https://doi.org/10.3310/hta16210
5. Budd, D., Burns, Guo, L'Italien, Lapuerta, 2011. Impact of early intervention and disease modification in patients with predementia Alzheimer's disease: a Markov model simulation. Clin. Outcomes Res. 189. https://doi.org/10.2147/CEOR.S22265
6. Caro, J.J., Briggs, A.H., Siebert, U., Kuntz, K.M., 2012. Modeling Good Research Practices—Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1. Med. Decis. Making 32, 667–677. https://doi.org/10.1177/0272989X12454577
7. Caro, J.J., Getsios, D., Migliaccio–Walle, K., Raggio, G., Ward, A., Group, A.S., others, 2001. Assessment of health economics in Alzheimer's disease (AHEAD) based on need for full-time care. Neurology 57, 964–971.
8. Cohen, J.T., Neumann, P.J., 2008. Decision analytic models for Alzheimer's disease: State of the art and future directions. Alzheimers Dement. 4, 212–222. https://doi.org/10.1016/j.jalz.2008.02.003
9. Djalalov, S., Yong, J., Beca, J., Black, S., Saposnik, G., Musa, Z., Siminovitch, K., Moretti, M., Hoch, J.S., 2012. Genetic Testing in Combination with Preventive Donepezil Treatment for Patients with Amnestic Mild Cognitive Impairment: An Exploratory Economic Evaluation of Personalized Medicine. Mol. Diagn. Ther. 16, 389–399. https://doi.org/10.1007/s40291-012-0010-7
10. Dowd, W.N., Cowell, A.J., Regan, D., Moran, K., Slevin, P., Doyle, G., Bray, J.W., 2018. An Exploratory Cost-Effectiveness Analysis of the Connected Health Intervention to Improve Care for People with Dementia: A Simulation Analysis. Health Serv. Outcomes Res. Methodol. 18, 47–62. https://doi.org/10.1007/s10742-017-0175-y
11. Ekman, M., Berg, J., Wimo, A., Jönsson, L., McBurney, C., 2007. Health utilities in mild cognitive impairment and dementia: a population study in Sweden. Int. J. Geriatr. Psychiatry 22, 649–655. https://doi.org/10.1002/gps.1725
12. Fuh, J.-L., Wang, S.-J., 2008. Cost-effectiveness analysis of donepezil for mild to moderate Alzheimer's disease in Taiwan. Int. J. Geriatr. Psychiatry 23, 73–8. https://doi.org/10.1002/gps.1842
13. Gagnon, M., Rive, B., Hux, M., Guilhaume, C., 2007. Cost-effectiveness of memantine compared with standard care in moderate-to-severe Alzheimer disease in Canada. Can. J. Psychiatry Rev. Can. Psychiatr. 52, 519–26.
14. Getsios D., Blume S., Ishak K.J., MacLaine G., Hernández L., 2012. An economic evaluation of early assessment for Alzheimer's disease in the United Kingdom. Alzheimers Dement. 8, 22–30. https://doi.org/10.1016/j.jalz.2010.07.001
15. Getsios, D., Blume, S., Ishak, K.J., Maclaine, G.D.H., 2010. Cost effectiveness of donepezil in the treatment of mild to moderate Alzheimer's disease: a UK evaluation using discrete-event simulation. PharmacoEconomics 28, 411–27. https://doi.org/10.2165/11531870-000000000-00000
16. Green, C., 2007. Modelling disease progression in Alzheimer's disease. Pharmacoeconomics 25, 735–750.

17. Green, C., Shearer, J., Ritchie, C.W., Zajicek, J.P., 2011. Model-Based Economic Evaluation in Alzheimer's Disease: A Review of the Methods Available to Model Alzheimer's Disease Progression. Value Health 14, 621–630. https://doi.org/10.1016/j.jval.2010.12.008

18. Green, C., Zhang, S., 2016. Predicting the progression of Alzheimer's disease dementia: A multidomain health policy model. Alzheimers Dement. https://doi.org/10.1016/j.jalz.2016.01.011

19. Gustavsson, A., Van Der Putt, R., Jönsson, L., McShane, R., 2009. Economic evaluation of cholinesterase inhibitor therapy for dementia: comparison of Alzheimer's disease and Dementia with Lewy bodies. Int. J. Geriatr. Psychiatry 24, 1072–1078. https://doi.org/10.1002/gps.2223

20. Handels, R.L.H., Wolfs, C.A.G., Aalten, P., Joore, M.A., Verhey, F.R.J., Severens, J.L., 2014. Diagnosing Alzheimer's disease: A systematic review of economic evaluations. Alzheimers Dement. 10, 225–237. https://doi.org/10.1016/j.jalz.2013.02.005

21. Hartz, S., Getsios, D., Tao, S., Blume, S., Maclaine, G., 2012. Evaluating the cost effectiveness of donepezil in the treatment of Alzheimer's disease in Germany using discrete event simulation. BMC Neurol. 12, 2–2. https://doi.org/10.1186/1471-2377-12-2

22. Henke, C., Burchmore, M., 1997. The economic impact of Tacrine in the treatment of Alzheimer's disease. Clin. Ther. 9.

23. Hernandez, L., Ozen, A., DosSantos, R., Getsios, D., 2016. Systematic Review of Model-Based Economic Evaluations of Treatments for Alzheimer's Disease. PharmacoEconomics 34, 681–707. https://doi.org/10.1007/s40273-016-0392-1

24. Hu, S., Yu, X., Chen, S., Clay, E., Toumi, M., Milea, D., 2015. Memantine for treatment of moderate or severe Alzheimer's disease patients in urban China: clinical and economic outcomes from a health economic model. Expert Rev. Pharmacoecon. Outcomes Res. 15, 565–578. https://doi.org/10.1586/14737167.2015.1065734

25. Ikeda, S., Niwata, S., Igarashi, Y., 2001. Evaluating QoL of elderly people with dementia. Jpn. J. Pharmacoepidemiol 5, 99–105.

26. Jones, C., Edwards, R.T., Hounsome, B., 2012. Health economics research into supporting carers of people with dementia: a systematic review of outcome measures. Health Qual. Life Outcomes 10, 1.

27. Jones, R.W., McCrone, P., Guilhaume, C., 2004. Cost effectiveness of memantine in Alzheimer's disease: an analysis based on a probabilistic Markov model from a UK perspective. Drugs Aging 21, 607–20. https://doi.org/10.2165/00002512-200421090-00005

28. Jonsson, L., 2005. Cost-effectiveness of memantine for moderate to severe Alzheimer's disease in Sweden. Am. J. Geriatr. Pharmacother. 3, 77–86. https://doi.org/10.1016/j.amjopharm.2005.05.002

29. Jonsson, L., Andreasen, N., Kilander, L., Soininen, H., Waldemar, G., Nygaard, H., Winblad, B., Jönhagen, M.E., Hallikainen, M., Wimo, A., 2006. Patient-and proxy-reported utility in Alzheimer disease using the EuroQoL. Alzheimer Dis. Assoc. Disord. 20, 49–55.

30. Jonsson, L., Lindgren, P., Wimo, A., Jonsson, B., Winblad, B., 1999. The cost-effectiveness of donepezil therapy in Swedish patients with Alzheimer's disease: a Markov model. Clin. Ther. 21, 1230–40. https://doi.org/10.1016/S0149-2918(00)80025-9

31. Kirbach, S., Simpson, K., Nietert, P.J., Mintzer, J., 2008. A Markov Model of the Cost Effectiveness of Olanzapine Treatment for Agitation and Psychosis in Alzheimer's Disease. Clin. Drug Investig. 28, 291–303. https://doi.org/10.2165/00044011-200828050-00003

32. Lachaine, J., Beauchemin, C., Legault, M., Bineau, S., 2011. Economic evaluation of the impact of memanine on time to nursing home admission in the treatment of Alzhiemer disease. Can. J. Psychiatry 56.

33. Lee, S.A.W., Sposato, L.A., Hachinski, V., Cipriano, L.E., 2017. Cost-effectiveness of cerebrospinal biomarkers for the diagnosis of Alzheimer's disease. Alzheimers Res. Ther. 9, 18. https://doi.org/10.1186/s13195-017-0243-0

34. Lopez-Bastida, J., Serrano-Aguilar, P., Perestelo-Perez, L., Oliva-Moreno, J., 2006. Social-economic costs and quality of life of Alzheimer disease in the Canary Islands, Spain. Neurology 67, 2186–2191. https://doi.org/10.1212/01.wnl.0000249311.80411.93

35. Loveman, E., Green, C., Kirby, J., Takeda, A., Picot, J., Payne, E., Clegg, A., 2006. The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease. Health Technol. Assess. Winch. Engl. 10, iii–iv, ix–xi, 1–160.

36. Martikainen, J., Valtonen, H., Pirttila, T., 2004. Potential cost-effectiveness of a family-based program in mild Alzheimer's disease patients. Eur. J. Health Econ. 5, 136–42.

37. Mirsaeedi-Farahani, K., Halpern, C.H., Baltuch, G.H., Wolk, D.A., Stein, S.C., 2015. Deep brain stimulation for Alzheimer disease: a decision and cost-effectiveness analysis. J. Neurol. 262, 1191–7. https://doi.org/10.1007/s00415-015-7688-5

38. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann. Intern. Med. 151, 264–269, W64.

39. Morris, J.C., Heyman, A., Mohs, R.C., Hughes, J.P., van Belle, G., Fillenbaum, G., Mellits, E.D., Clark, C., 1989. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 39, 1159–1165.

40. Neumann, P., Hermann, R.C., Weistein, M., 1998. Measuring QALYs in dementia, in: Health Economics of Dementia. John Wiley & Sons, Inc., Sussex, UK, p. 359-370.

41. Neumann, P. J., Hermann, R.C., Kuntz, K.M., Araki, S.S., Duff, S.B., Leon, J., Berenbaum, P.A., Goldman, P.A., Williams, L.W., Weinstein, M.C., 1999. Cost-effectiveness of donepezil in the treatment of mild or moderate Alzheimer's disease. Neurology 52, 1138–45.

42. Neumann, Peter J., Kuntz, K.M., Leon, J., Araki, S.S., Hermann, R.C., Hsu, M.-A., Weinstein, M.C., 1999. Health Utilities in Alzheimer's Disease: A Cross-Sectional Study of Patients and Caregivers. Med. Care 37, 27–32.

43. O'Brien, B.J., Goeree, R., Hux, M., Iskedjian, M., Blackhouse, G., Gagnon, M., Gauthier, S., 1999. Economic evaluation of donepezil for the treatment of Alzheimer's disease in Canada. J. Am. Geriatr. Soc. 47, 570–578.

44. Oremus, M., 2008. Systematic review of economic evaluations of Alzheimer's disease medications. Expert Rev. Pharmacoecon. Outcomes Res. 8, 273–289. https://doi.org/10.1586/14737167.8.3.273

45. Peters, J.L., Anderson, R., Hoyle, M., Hyde, C., 2013. Evolution of a cost utility model of donepezil for Alzheimer's disease. Int. J. Technol. Assess. Health Care 29, 147–154. https://doi.org/10.1017/S026646231300007X

46. Philips, Z., Bojke, L., Sculpher, M., Claxton, K., Golder, S., 2006. Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment. PharmacoEconomics 24, 355–371. https://doi.org/10.2165/00019053-200624040-00006

47. Philips, Z., Ginnelly, L., Sculpher, M., Claxton, K., Golder, S., 2004. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technol. Assess. 8.

48. Prince, M., Albanese, E., Guerchet, M., Prina, M., 2014. Dementia and Risk Reduction: An analysis of protective and modifiable factors, World Alzheimer Report. Alzheimer's Disease International (ADI), London, UK.

49. Retchin, S.M., Hillner, B.E., 1994. The costs and benefits of a screening program to detect dementia in older drivers. Med. Decis. Making 14, 315–324.

50. Rive, B., Grishchenko, M., Guilhaume-Goulant, C., Katona, C., Livingston, G., Lamure, M., Toumi, M., Francois, C., 2010. Cost effectiveness of memantine in Alzheimer's disease in the UK. J. Med. Econ. 13, 371–80. https://doi.org/10.3111/13696998.2010.491347

51. Rogers, S., Farlow, M., Doody, R., Mohs, R., Friedhoff, T., Donepezil Study Group, 1998. A 24-week, double-blind, placebocontrolled trial of donepezil in patients with Alzheimer's disease. Neurology 50, 137.

52. Rowen, D., Mulhern, B., Banerjee, S., van Hout, B., Young, T.A., Knapp, M., Smith, S.C., Lamping, D.L., Brazier, J.E., 2012. Estimating Preference-Based Single Index Measures

for Dementia Using DEMQOL and DEMQOL-Proxy. Value Health 15, 346–356. https://doi.org/10.1016/j.jval.2011.10.016

53. Skoldunger, A., Johnell, K., Winblad, B., Wimo, A., 2013. Mortality and treatment costs have a great impact on the cost-effectiveness of disease modifying treatment in Alzheimer's disease--a simulation study. Curr. Alzheimer Res. 10, 207–16.

54. Sopina, E., Sørensen, J., 2018. Decision modelling of non-pharmacological interventions for individuals with dementia: a systematic review of methodologies. Health Econ. Rev. 8. https://doi.org/10.1186/s13561-018-0192-8

55. Stern, Y., Tang, M.-X., Albert, M.S., Brandt, J., Jacobs, D.M., Bell, K., Marder, K., Sano, M., Devanand, D., Albert, S.M., others, 1997. Predicting time to nursing home care and death in individuals with Alzheimer disease. Jama 277, 806–812.

56. Teipel, S.J., Ewers, M., Reisig, V., Schweikert, B., Hampel, H., Happich, M., 2007. Long-term cost-effectiveness of donepezil for the treatment of Alzheimer's disease. Eur. Arch. Psychiatry Clin. Neurosci. 257, 330–6. https://doi.org/10.1007/s00406-007-0727-1

57. Thibault, C., Özer Stillman, I., Chen, S., Getsios, D., Proskorovsky, I., Hernandez, L., Dixit, S., 2015. Cost-utility analysis of memantine extended release added to cholinesterase inhibitors compared to cholinesterase inhibitor monotherapy for the treatment of moderate-to-severe dementia of the Alzheimer's type in the US. J. Med. Econ. 18, 930–943. https://doi.org/10.3111/13696998.2015.1063501

58. Touchon, J., Lachaine, J., Beauchemin, C., Granghaud, A., Rive, B., Bineau, S., 2014. The impact of memantine in combination with acetylcholinesterase inhibitors on admission of patients with Alzheimer's disease to nursing homes: cost-effectiveness analysis in France. Eur. J. Health Econ. HEPAC Health Econ. Prev. Care 15, 791–800. https://doi.org/10.1007/s10198-013-0523-y

59. Weinstein, M.C., O'Brien, B., Hornberger, J., Jackson, J., Johannesson, M., McCabe, C., Luce, B.R., 2003. 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 Health 6, 9–17. https://doi.org/10.1046/j.1524-4733.2003.00234.x

60. Wong, C.L., Bansback, N., Lee, P.E., Anis, A.H., 2009. Cost-Effectiveness: Cholinesterase Inhibitors and Memantine in Vascular Dementia. Can J Neurol Sci 36, 735–739.

61. Yang, K.-C., Chen, H.-H., 2016. Probabilistic cost-effectiveness analysis of vaccination for mild or moderate Alzheimer's disease. Curr. Alzheimer Res. 13, 809–816.

62. Yu, S.-Y., Lee, T.-J., Jang, S.-H., Han, J.W., Kim, T.H., Kim, K.W., 2015. Cost-Effectiveness of Nationwide Opportunistic Screening Program for Dementia in South Korea. J. Alzheimers Dis. 44, 195–204.

63. Zhang, Y., Kivipelto, M., Solomon, A., Wimo, A., 2011. Cost-Effectiveness of a Health Intervention Program with Risk Reductions for Getting Demented: Results of a Markov Model in a Swedish/Finnish Setting. J. Alzheimers Dis. 26, 735–744. https://doi.org/10.3233/JAD-2011-110065

List of figures and tables

Figure 1: Some characteristics of decision-analytic models used in economic evaluation of dementia interventions

Figure 2: Links among cost-effectiveness models of dementia interventions

Figure 3: Changes of quality reporting, by percentage of items scored "Yes" (sufficient evidence) and "No" (no or unclear evidence)

Table 1: Changes in modelling and reporting practice and quality (based on the Philips checklist scoring results)