IISE Transactions on Healthcare Systems Engineering

ISSN: 2472-5579 (Print) 2472-5587 (Online) Journal homepage: https://www.tandfonline.com/loi/uhse21

Optimal population screening policies for Alzheimer's disease*

Zehra Önen, Serpil Sayin & Ibrahim Hakan Gürvit

To cite this article: Zehra Önen, Serpil Sayin & Ibrahim Hakan Gürvit (2019) Optimal population screening policies for Alzheimer's disease*, IISE Transactions on Healthcare Systems Engineering, 9:1, 14-25, DOI: 10.1080/24725579.2018.1543738

To link to this article: https://doi.org/10.1080/24725579.2018.1543738

Accepted author version posted online: 20 Dec 2018.
Published online: 26 Feb 2019.

IISE TRANSACTIONS ON HEALTHCARE SYSTEMS ENGINEERING
2019, VOL. 9, NO. 1, 14–25
https://doi.org/10.1080/24725579.2018.1543738

Optimal population screening policies for Alzheimer's disease*

Zehra Önen[a], Serpil Sayin[b], and Ibrahim Hakan Gürvit[c]

[a]Graduate School of Sciences and Engineering, Koç University, Istanbul, Turkey; [b]College of Administrative Sciences and Economics, Koç University, Istanbul, Turkey; [c]Department of Neurology, Faculty of Medicine, Istanbul University, Istanbul, Turkey

ABSTRACT

Alzheimer's disease (AD) constitutes a serious societal healthcare issue as the proportion of the aging population increases. There are ongoing discussions about the necessity of screening the population for AD. We investigate optimal population screening policies for AD using Markov Decision Processes (MDPs). The objective function combines quality-adjusted life years and costs. The disease states are identified according to Clinical Dementia Rating (CDR) scores. The screening test in the model is the Mini Mental State Examination (MMSE), a cognitive test that is widely used in clinical practice. A numerical implementation of the MDP model is presented based on data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and existing literature. In the baseline case, the optimal outcome is not to employ a population-wide screening program. We conduct extensive sensitivity analyses on several model parameters. Our study reveals that the optimal policy may be sensitive to changes in transition probability estimates. When we focus on transitions that are related to treatment effectiveness, we find that implementing a population screening policy becomes socially optimal when plans that lead to cognitive ability stabilization or improvement become available.

KEYWORDS

Alzheimer's disease; population screening; optimal policy; Markov Decision Process

1. Introduction

According to the *World Alzheimer Report 2015* delivered by Alzheimer's Disease International, currently there are over 46 million people suffering from dementia. This figure is expected to double by year 2030 and more than triple by year 2050 (Prince, 2015). According to this report, the total cost of dementia, including direct medical, social sector and informal care costs, is estimated to be US $818 billion worldwide, and is foreseen to reach a trillion dollars in the next two years. Alzheimer's disease (AD) is believed to be the most common type of dementia, accounting for 50% to 75% of the total with a greater proportion in the higher age ranges (Duthey, 2013).

AD has an insidious onset and its progression is characterized by memory dysfunction and cognitive disturbances, such as problems recalling familiar names and objects and/or behavioral changes. Loss of short-term memory and impaired visuospatial orientation are typical early symptoms of the disease. As AD progresses, there is a general decline of multiple cognitive functions related to daily activities. Ultimately, patients become very dependent on caregivers. The clinical diagnosis of AD is based on the patient's medical history and a neurological assessment, along with neuropsychiatric testing of the patient's cognitive functions. To be able to differentiate from other causes of dementia, neuroimaging techniques may be used. Technically, a definite diagnosis of AD cannot be made until a neuropathological exam of the brain is carried out after the patient's death.

Clinical Dementia Rating (CDR) was developed as part of a memory and aging project at Washington University in 1979. It is a widely used instrument to diagnose and stage AD. CDR involves an individual's assessment in six domains (memory, orientation, judgment and problem solving, community affairs, home and hobbies and personal care) where information is collected from the individual as well as an informant. Scores in these dimensions are then combined to obtain a score that indicates severity of dementia. A CDR score of 0 indicates no dementia, 0.5 is questionable dementia and 1, 2 and 3 are mild, moderate and severe dementia, respectively (Morris, 1993). CDR is found to have high inter-evaluator validity and is a trusted method of AD staging; however, it is not appropriate as a brief screening tool due to the quantity and complexity of information to be collected (Morris, 1997).

Several screening tests have been developed for screening cognition in individuals with a likely cognitive decline. Among these, Mini Mental State Examination (MMSE), developed by Folstein *et al.* (1975), is the one that has widespread use in clinical practice. MMSE is composed of 30 items, each correct response of which is worth one point. Those individuals scoring below a certain cut-off point are likely to be demented with a certain sensitivity and specificity. In a review that studies

CONTACT Serpil Sayin ssayin@ku.edu.tr College of Administrative Sciences and Economics, Koc University, Sariyer, Istanbul 34450, Turkey.

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

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uhse.

validity and applicability of different screening tests, MMSE and its variations are found to be most reliable along a number of dimensions (Cullen *et al.*, 2007). Another study that evaluates the quality of a mapping of MMSE scores to CDR stages in terms of accuracy concludes that it can be used as a reliable tool to diagnose and stage dementia (Perneczky *et al.*, 2006).

AD is not curable once cognitive decline begins. Despite the definite progressive nature of AD, some studies have reported episodes of cognitive improvement in patients under drug therapy (AD2000 Collaborative Group, 2004; Mangialasche *et al.*, 2010; Neumann *et al.*, 2001). There are four registered drugs that are given to treat symptoms of AD. These are the three acetylcholinesterase inhibitors (AChEIs) donepezil, galantamine and rivastigmine, and the glutamate inhibitor memantine. It has been established that AChEIs have positive effects on cognition, behavior and activities of daily living. According to the National Institute for Health and Care Excellence (NICE), AChEIs are recommended for patients with mild to moderate dementia (NICE *et al.*, 2011). There are several cost-effectiveness studies for drug treatment. For example Touchon *et al.* (2014) showed that a combination of memantine and one AChEI can significantly delay institutionalization time and increase life quality. The reader may refer to the reviews of Kaduszkiewicz *et al.* (2005) and Kirby *et al.* (2006) for further details about the effectiveness of AChEIs and memantine, respectively.

Screening tests aim to sort out individuals who probably have a disease from those who probably do not (World Health Organization, 2012). In general, screening tests may generate false positives and false negatives with a certain likelihood. A screening policy is expected to balance the trade-off between the benefits of early intervention and risks of unnecessary screening test applications and implications of false positives. Currently, there are no population screening policies in any country for AD, although its importance has been stated at various occasions, such as in the Leon Thal Symposium series (Khachaturian *et al.*, 2010, 2011) and the National Alzheimer's Project Act (NAPA) in the United States (Khachaturian *et al.*, 2012). The government proposal in the UK for screening the elderly population for dementia during their routine health checks set off a discussion among health professionals who are in favor and those who are against such a policy (Kmietowicz, 2012). In January 2015, the UK National Screening Committee upheld its recommendation against screening everyone aged 65 and over for dementia. Their decision was mainly based on poor accuracy of the suggested cognitive test and lack of a cure (UK National Screening Committee, 2015). A survey-based study conducted in 2009 in multiple European countries found that a smaller proportion of physicians (42%) and payers (44%) than members of the general public (81%) or caregivers (80%) agreed that a routine screening for AD starting at age 65 would be beneficial (Bond *et al.*, 2010). Participants who were not in favor of screening cited reasons such as the inaccuracy of available tests, high costs, absence of cure, negative impact on the individual and late visibility of symptoms in AD.

There have been a number of studies that find AD screening cost-effective. Weimer and Sager (2009) used a Monte Carlo cost-benefit simulation framework using parameter estimates from the literature. Getsios *et al.* (2012) used a discrete event simulation framework where some data were taken from donepezil treatment research conducted in the UK. Barnett *et al.* (2014) developed a cohort model with two different treatment scenarios depending on MMSE course over time. These studies support the idea that screening for AD is cost-effective for society as well as the individual. Dixon *et al.* (2014) compared the results of a hypothetical one-time screening program with a non-screening program on people aged 75 or older in England and Wales and found that a screening program could be cost-effective if treatments and social care interventions were to be more effective. Finally, Yu *et al.* (2015) conducted cost-effectiveness research on a screening program in Korea and found cost-effectiveness to be very sensitive to treatment effectiveness. None of these studies sought an optimal screening policy for AD.

In this study, our goal is to investigate the viability of employing a population-wide screening program for AD by modeling the decision problem as a Markov Decision Process (MDP). Markov models have been used widely to model AD progression (Cohen and Neumann, 2008; Green, 2007; Green *et al.*, 2011). However, to the best of our knowledge, they have not been used to investigate a population screening decision for AD, although they have been utilized in addressing screening decisions in different health conditions. Pioneering studies on screening date back to the 1990s (Özekici and Pliska, 1991; Parmigiani, 1993). Screening models for several types of cancer, diabetes and infectious diseases have been built using MDPs and partially observable MDPs. (Ayer *et al.*, 2012; Gillies *et al.*, 2008; Harper and Jones, 2005; Kirkizlar *et al.*, 2010; Kurt *et al.*, 2011; Maillart *et al.*, 2008; Underwood *et al.*, 2012; Zhang *et al.*, 2012). Our model is built so as to employ a cognitive screening test and treat individuals who are identified in the process as cognitively deficient with an appropriate drug therapy. We implement the model using data from the literature. Data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study is used to estimate transition probabilities based on a multistate Markov model. Our screening model is a simple MDP that can be solved using standard techniques. This allows us to conduct extensive sensitivity analyses on the parameters of the model. Our studies direct us to focus on the impact of improvements in treatment effectiveness on the optimal policy and quantify expected cost and QALY gains that become possible. As such, we obtain an indication as to at which levels of treatment effectiveness a population-wide screening action would be desirable, as well as a description of the associated screening policy. The remainder of this article is organized as follows. In Section 2 we provide a mathematical description of our model. Section 3 contains our numerical implementation as well as results of sensitivity analyses. A summary of our findings with future research directions is given in the last section.

2. The model

To find an optimal screening policy for a given population, we develop a finite horizon discounted MDP where an optimal action with respect to maximizing a predetermined

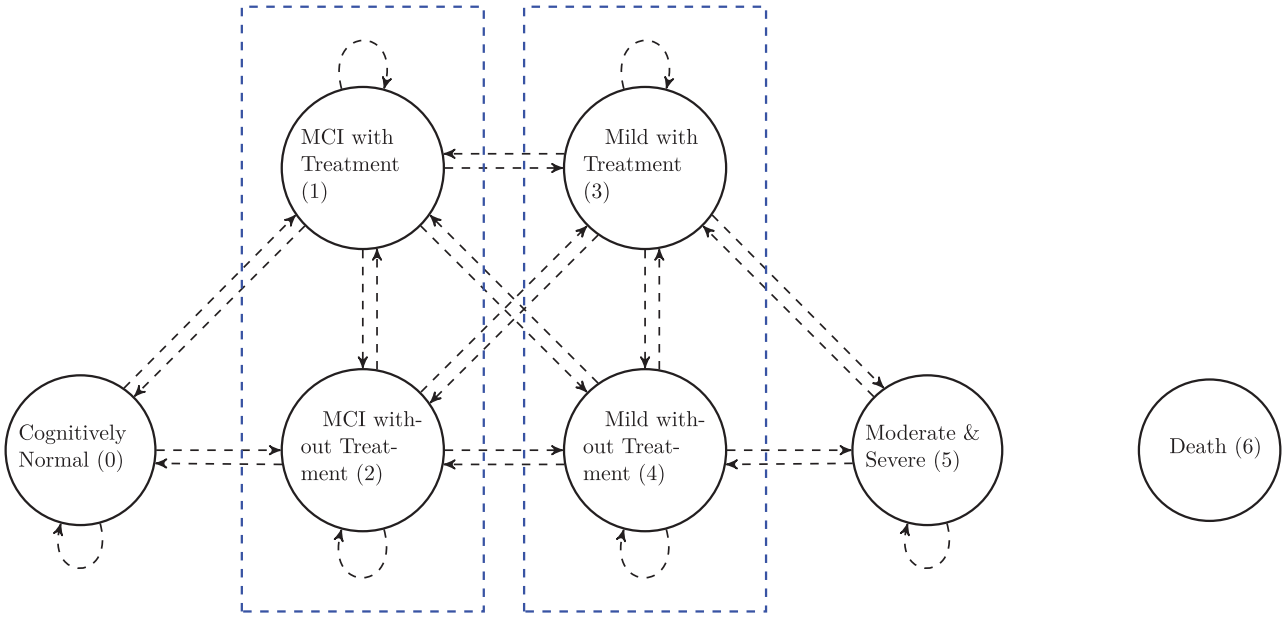

Figure 1. The model framework.

objective function is determined in every period. The problem is modeled from the perspective of a policymaker whose possible actions are to implement or not to implement a screening action in a given period. Thus, the collection of optimal actions across time constitutes a population-wide optimal screening policy. In the model, t denotes the decision epochs or periods, $t = 0, 1, ..., T<\infty$, and our cycle time is one year.

S denotes the state space, $S = \{0, 1, 2, 3, 4, 5, 6\}$ where $s_t \in S$ represents the state of an individual in period t. The stages typically used in AD disease models, namely cognitively normal, Mild Cognitive Impairment (MCI), mild, moderate and severe stages, constitute the basis of our model. We have two states for each of MCI and mild stages, differentiating between treatment and no treatment options. In moderate and severe stages, we do not make this distinction. This is mainly due to transition probability estimation issues, because our data set contains few instances of untreated patients in these disease stages. We pool these stages where the disease clearly manifests itself and include treatment costs when estimating our model data.

In our model, state 0 is the cognitively normal stage, corresponding to CDR score of 0. In this state, AD is not detectable by a cognitive test. Our MCI states, 1 with treatment and 2 without treatment, both correspond to a CDR score of 0.5. Likewise, state 3 corresponds to mild AD stage with treatment whereas state 4 is the no treatment mild stage, both corresponding to a CDR score of 1. State 5 corresponds to moderate and severe stages of AD with CDR score higher than 1, regardless of treatment. State 6 is the absorbing state of death due to AD and all other causes. These states have been defined with respect to the natural course of the disease as presented in the literature. Although several studies using Markov models have defined the states with respect to mild, moderate and severe stages of AD (López-Bastida *et al.*, 2009; Neumann *et al.*, 2001), recent literature has shown the existence of MCI to be a potential early stage of AD. We include MCI as one of the states of the model (Association *et al.*, 2017; McKhann *et al.*, 1984; Sperling *et al.*, 2011).

Figure 1 illustrates the model states where the dashed arrows represent the instant state transitions that are allowed in the multi-state Markov (MSM) model, which is used to estimate transition probabilities. MSM models are widely used to model disease progression (Jackson *et al.*, 2003). An MSM model describes how an individual moves among a number of states in continuous time. This is translated into movement in the discrete state space by transition intensities. An MSM model requires observing the state of individuals at particular points in time from which a maximum likelihood transition intensity matrix is estimated. MSM models may include covariates. The estimation of transition probabilities is, in turn, conducted based on intensity rates. More detailed information on MSM methodology is given in the Appendix. Despite the progressive nature of the disease, our model allows some of the reverse transitions because it has been reported that cognition may improve temporarily during the course of the disease (AD2000 Collaborative Group, 2004; Koepsell and Monsell, 2012; Mangialasche *et al.*, 2010; Neumann *et al.*, 2001). Transitions to death state are possible from any state and arrows are not drawn to avoid clutter. The yearly transition matrices for the MDP model are then derived from the estimated transition intensities. The transition probability matrices are also given in Section 3.

When a screening action takes place, individuals identified as positive in states 0 and 2 move to state 1. An individual in state 1 may stay there, move to state 0 since her cognition improves, to 3, 4 and 5 because of a cognitive decline, or to 2 due to treatment nonadherence.

We need the following additional notation for a full mathematical description of the model:

- a_t: Action chosen at period t; i.e., $a_t \in A_t = \{Y, N\}$ where Y represents the action to screen the patient using MMSE test and N represents the action not to screen.
- $p_t^{ss'}(a)$: State transition probabilities. It is the probability that the individual will be in state $s' \in S$ at decision epoch $t+1$ given that he/she is in state $s \in S$ at decision epoch t and action a is taken. For example, $p_t^{21}(N)$ represents the probability that the individual in state 2 at period t will be in 1 in period $t+1$ when the chosen action at t is N. The transition probability matrix in period t when action a is taken is denoted as $P_t(a)$.
- $R_t(s,a)$: Expected reward between epochs t and $t+1$ when the individual is in state $s \in S$ and action a is taken. The reward is a function of an individual's QALY at that state and the associated costs of the state.
- $V_t(s)$: Value function that gives the expected remaining reward when the individual is in state s in period t.
- γ: Discount factor, $0<\gamma \leq 1$.

To find the optimal solution, the following set of recursive equations needs to be solved. These equations are known as Bellman's equations (Puterman, 1994).

$$\begin{aligned} V_T(s) &= 0, \forall s \in S, \\ V_t(s) &= \max\Bigg\{ R_t(s,Y) + \sum_{i=0}^{6} p_t^{si}(Y)\gamma V_{t+1}(i), \\ &\qquad R_t(s,N) + \sum_{i=0}^{6} p_t^{si}(N)\gamma V_{t+1}(i) \Bigg\}, \forall s \in S, \ \ \forall t<T \end{aligned} \tag{1}$$

For a finite horizon model like ours, any solution to Bellman's equations is an optimal solution. Also, since the state space and action set are finite, any policy that satisfies Bellman's equations is an optimal policy (Puterman, 1994).

3. A numerical implementation of the model

The time horizon we choose for our screening model is between the ages 60 and 100; i.e., $T=40$ where $t=0$ corresponds to age 60. For an implementation of the model, we need to estimate transition probabilities, accuracy of the screening test, and establish rewards and costs. We use a combination of Quality Adjusted Life Years (QALYs) and monetary costs as rewards. The QALY of a particular patient is the equivalent time in full health that matches her one year in her health state (Brazier, 2007). Cost and QALY figures are taken from the literature. We estimate transition probabilities based on ADNI data. Detailed explanation for each of the estimated parameters are presented below.

3.1. Transition probabilities

ADNI is a longitudinal study that has been collecting data on clinical, imaging, genetic and biospecimen biomarkers through the process of normal aging to MCI and to dementia. The main objective of ADNI is currently stated as tracking the progression of the disease using biomarkers. The initial phase of the ADNI study, which is labeled as ADNI1, had the goal of creating uniform standards for acquiring longitudinal, multi-site magnetic resonance imaging and positron emission tomography data on patients with AD and MCI as well as a control group. ADNI1 is a multi-center study in which about 800 subjects were recruited over 50 different sites across the United States and Canada. The key eligibility criteria were being between the ages 55 to 90 and having a study partner who is able to provide an independent evaluation of functioning (Alzheimer's Disease Neuroimaging Initiative, 2006). The participants were followed for a maximum duration of eight years between 2005 and 2014. Each participant was asked to undergo follow-up exams at six months. Medication information of participants was updated annually. Some follow-ups may have missing values due to reasons such as inaccessibility of the participant at that date or unwillingness to participate. Observations may be truncated by death.

There are 229 Normal participants, 398 MCI patients and 192 AD patients in our ADNI1 data set whose ages, CDR scores, MMSE scores and treatment information are available. The transition probabilities among model states are estimated by means of an MSM model. We used R's MSM package (Jackson, 2011) to estimate the transition intensities where age was introduced as a covariate. We generated two sets of transition probabilities based on age, the first set covering ages 60 to 74 and the second set covering ages 75 to 100. In the latest Alzheimer's Association report (2017), age is mentioned as being one of the most important factors that have an impact on the prevalence of the disease. Seventy-five is used as one of the cutoff values in this report and also in Neumann *et al.* (2011).

The resulting probability transition matrices among our states estimated from ADNI data are presented in Table 1. $P_t(a_t)$ denotes the estimated probability of transitioning from one state to another in one year, t being broken into two intervals between 0 to 14 and 15 to 40 with the chosen action a_t being either N for no screening or Y for screening. With rows and columns of each matrix numbered in increasing order of the states starting from state 0, the ij^{th} entry of a matrix in Table 1 corresponds to the transition probability $p_t^{ij}(a)$ where i is the outgoing state and j is the incoming state. The ADNI observations are used to estimate the transitions when the chosen action is not to screen. The transition probabilities when the chosen action is to screen is built on these estimates as follows. When a_t = Y, transitions from states 1, 3, 4 and 5 will not change as these individuals are already diagnosed. Individuals in states 0 and 2 who are identified as cognitively deficient will be moved to a treatment state 1. Therefore, the transitions from state 2 to 2 and from state 0 to 0 of $P_t(N)$ need to be reallocated in accordance with the sensitivity (α) and specificity (β) of the test employed. The computation of transition probabilities when action is to screen can be summarized as follows:

$$\begin{aligned} p_t^{00}(Y) &= \beta p_t^{00}(N) \\ p_t^{01}(Y) &= (1-\beta)p_t^{00}(N) + p_t^{01}(N) \\ p_t^{21}(Y) &= p_t^{21}(N) + \alpha p_t^{22}(N) \\ p_t^{22}(Y) &= (1-\alpha)p_t^{22}(N) \\ p_t^{0i}(Y) &= p_t^{0i}(N), \quad for \quad i = 2,3,4,5,6. \\ p_t^{2i}(Y) &= p_t^{2i}(N), \quad for \quad i = 0,3,4,5,6. \end{aligned}$$

Table 1. Estimated transition probabilities.

$P_t(a_t)$	$a_t = N$	$a_t = Y$
$t<15$	$\begin{pmatrix} 0.912 & 0.004 & 0.076 & 0.001 & 0.002 & 0 & 0.004 \\ 0.017 & 0.662 & 0.02 & 0.258 & 0.002 & 0.037 & 0.005 \\ 0.085 & 0.085 & 0.742 & 0.035 & 0.038 & 0.01 & 0.004 \\ 0.001 & 0.072 & 0.002 & 0.703 & 0.007 & 0.196 & 0.02 \\ 0.007 & 0.021 & 0.109 & 0.256 & 0.356 & 0.187 & 0.064 \\ 0 & 0.007 & 0 & 0.139 & 0.001 & 0.705 & 0.147 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.848 & 0.068 & 0.076 & 0.001 & 0.002 & 0 & 0.004 \\ 0.017 & 0.662 & 0.02 & 0.258 & 0.002 & 0.037 & 0.005 \\ 0.085 & 0.612 & 0.215 & 0.035 & 0.038 & 0.01 & 0.004 \\ 0.001 & 0.072 & 0.002 & 0.703 & 0.007 & 0.196 & 0.02 \\ 0.007 & 0.021 & 0.109 & 0.256 & 0.356 & 0.187 & 0.064 \\ 0 & 0.007 & 0 & 0.139 & 0.001 & 0.705 & 0.147 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$
$t \geq 15$	$\begin{pmatrix} 0.891 & 0.008 & 0.086 & 0.001 & 0.004 & 0 & 0.01 \\ 0.003 & 0.7 & 0.013 & 0.22 & 0.003 & 0.042 & 0.019 \\ 0.077 & 0.13 & 0.671 & 0.029 & 0.058 & 0.01 & 0.025 \\ 0 & 0.078 & 0.002 & 0.632 & 0.011 & 0.257 & 0.02\ 7 \\ 0.008 & 0.022 & 0.13 & 0.148 & 0.505 & 0.168 & 0.02 \\ 0 & 0 & 0 & 0 & 0 & 0.872 & 0.128 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.829 & 0.07 & 0.086 & 0.001 & 0.004 & 0 & 0.01 \\ 0.003 & 0.7 & 0.013 & 0.22 & 0.003 & 0.042 & 0.019 \\ 0.077 & 0.606 & 0.195 & 0.029 & 0.058 & 0.01 & 0.025 \\ 0 & 0.078 & 0.002 & 0.632 & 0.011 & 0.257 & 0.02 \\ 0.008 & 0.022 & 0.13 & 0.148 & 0.505 & 0.168 & 0.02 \\ 0 & 0 & 0 & 0 & 0 & 0.872 & 0.128 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

The first equation reflects the fact that transitions from 0 to 0 happen when the test result is a true negative. The second equation indicates that additional transitions from 0 to 1 may happen when the test result is a false positive. The third equation is associated with the case when the test result is a true positive in addition to the normal course of transition from state 2 to state 1. Finally, the fourth equation refers to the case when the test result is a false negative and hence the individual remains in state 2.

We estimate α and β by maximizing the Youden Index, which is equal to $\alpha + \beta - 1$ (Youden, 1950) over ADNI data where cognitively normal individuals are distinguished from others using an MMSE cut-off score. This resulted in an MMSE cut-off score of 27 (MMSE score $\geq$ 28 for cognitively normal) with $\alpha = 0.71$ and $\beta = 0.93$.

3.2. Quality of life data

QALYs of this implementation are computed based on the studies of Ready *et al.* (1999) and Neumann *et al.* (2001). QALY values for different stages of the disease (1, 2, 3, 4, 5) are 0.73, 0.73, 0.68, 0.63 and 0.52, respectively. The QALY value of death is 0. To determine the QALY value of our state 5, we took the weighted average of QALY in moderate and QALY in severe stages of Neumann *et al.* (1999). Neumann *et al.* (1999) do not provide a QALY value for a cognitively normal individual. We use 0.88 for this value, which is the QALY value of the caregiver of a patient in nursing home care (Neumann *et al.*, 1999). Most of the time, the caregiver is the spouse of the patient and is of around the same age, but presumably in a mentally healthy state. Without the burden of home care, the caregiver's QALY is a good approximation of the QALY value of a participant in cognitively normal stage.

3.3. Cost data

Costs associated with different model states are taken from different research studies. The cost of being in state 0 is taken from the work of Alemayehu and Warner (2004). They estimated health care costs for an elderly patient at age 65 as $10,245 in year 2000 dollars. For other states, we use Leon *et al.*'s (1998) study as a basis. In this study, costs (including treatment, formal and informal care) are given across a number of settings and for mild (including MCI patients), moderate and severe stages of the disease. We associate the annual cost of $13,068, $18,408 and $30,699 for our states 1, 3 and 5, respectively, the last one being a weighted average of the moderate and severe stage costs. For our states 2 and 4, which are MCI without treatment and mild without treatment, we subtract the annual cost of treatment, estimated as $1,825 in 2009 year dollars (Leon and Neumann, 1999) after all figures are expressed in 2016 US dollars. MMSE is a screening tool which has no cost other than a regular physician consultation, so we take $c_{Sc} = 0$. All costs used in the model implementation expressed in 2016 dollars are presented in Table 2.

Table 2. Costs associated with states in 2016 dollars ($).

s	$c(s)$
0	14,235
1	20,108
2	18,061
3	28,325
4	26,278
5	47,224
6	0

3.4. Rewards

Our rewards combine costs and QALY values using a cost-effectiveness ratio (r). For all $s \in S \setminus \{6\}$ and $a \in \{Y, N\}$, $R(s,a) = r \times QALY(s) - c(s,a)$ where $c(s,N) = c(s)$ for all s and $c(s,Y) = c(s) + c_{sc}$ for all s. Because screening action is valid for states 0 and 2, the rewards for states 1, 3, 4 and 5 are action independent, hence $R(s,a) = R(s) = r \times QALY(s) - c(s)$. State 6 is assumed to have a zero reward. We take r as $100,000 per QALY per year.

3.5. Results

3.5.1. Baseline results

In the baseline implementation, we use the parameters described earlier and a discount factor γ of 0.98. The optimal policy turns out to be not to screen the population at any time. We then conduct sensitivity analyses on a number

Table 3. Estimated transition probabilities ($a_t = N$)

$P_t(a_t)$	Gender = Female	Gender = Male
$t<15$	$\begin{pmatrix} 0.933 & 0.005 & 0.058 & 0.001 & 0.002 & 0.001 & 0 \\ 0.01 & 0.64 & 0.019 & 0.295 & 0.004 & 0.03 & 0.002 \\ 0.084 & 0.122 & 0.679 & 0.036 & 0.056 & 0.021 & 0.001 \\ 0 & 0.067 & 0.001 & 0.754 & 0.016 & 0.147 & 0.014 \\ 0 & 0.008 & 0 & 0.174 & 0.434 & 0.347 & 0.036 \\ 0 & 0.008 & 0 & 0.177 & 0.002 & 0.674 & 0.139 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.894 & 0.004 & 0.091 & 0.002 & 0.002 & 0 & 0.007 \\ 0.022 & 0.679 & 0.021 & 0.231 & 0 & 0.038 & 0.009 \\ 0.085 & 0.061 & 0.78 & 0.035 & 0.027 & 0.005 & 0.007 \\ 0.001 & 0.076 & 0.001 & 0.673 & 0 & 0.224 & 0.025 \\ 0.023 & 0.038 & 0.349 & 0.313 & 0.209 & 0.062 & 0.007 \\ 0 & 0.006 & 0 & 0.109 & 0 & 0.718 & 0.166 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$
$t \geq 15$	$\begin{pmatrix} 0.908 & 0.007 & 0.071 & 0.001 & 0.003 & 0 & 0.01 \\ 0.001 & 0.615 & 0.023 & 0.291 & 0.004 & 0.063 & 0.003 \\ 0.087 & 0.127 & 0.67 & 0.034 & 0.059 & 0.014 & 0.008 \\ 0 & 0.053 & 0.003 & 0.627 & 0.013 & 0.284 & 0.02 \\ 0.009 & 0.017 & 0.138 & 0.081 & 0.529 & 0.21 & 0.015 \\ 0 & 0 & 0 & 0 & 0 & 0.882 & 0.118 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.873 & 0.009 & 0.101 & 0.001 & 0.004 & 0 & 0.011 \\ 0.005 & 0.732 & 0.009 & 0.195 & 0.002 & 0.034 & 0.023 \\ 0.071 & 0.13 & 0.673 & 0.028 & 0.058 & 0.008 & 0.033 \\ 0 & 0.096 & 0.002 & 0.634 & 0.009 & 0.238 & 0.021 \\ 0.006 & 0.027 & 0.119 & 0.189 & 0.481 & 0.123 & 0.054 \\ 0 & 0 & 0 & 0 & 0 & 0.866 & 0.134 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

of parameters used in the model. An analysis on the cost-effectiveness ratio shows that the solution is not sensitive to this value and not screening remains the optimal policy for r values as large as \$200,000. Similarly, a sensitivity analysis on discount factor γ is conducted. No changes occur when γ varies between 0.70 and 1. We then analyze the effect of the accuracy of the screening test on the result and find that even when the screening test is perfect, i.e. $\alpha = \beta = 1.0$, not screening remains the optimal policy.

In order to interpret the underlying dynamics that lead to the robustness of the no screening solution, a closer look at the transition probabilities reveals the following counter intuitive observation. We note that the transition probabilities from state 2 to 0 are higher than the transition probabilities from state 1 to 0. In other words, individuals with MCI and who do not receive treatment are more likely to show cognitive improvement in a period than individuals who have MCI and are under treatment. Likewise, the probability of transitioning from state 1 to 3 is higher than the probability of transitioning from state 2 to 3. Therefore, between the two MCI states, the one with treatment, state 1, is the less desirable one in terms of disease progression. As a screening action increases the probability of moving from state 2 to 1 and decreases the probability of staying in state 2, a no-screening solution dominates in accordance with the reward structure. The counter intuitive transitions in ADNI data can possibly be attributed to some selection bias. It is possible that, among all MCI individuals, the ones who receive treatment are the ones who are believed to possess a higher risk of AD progression than others who do not receive treatment. This observation motivates us to conduct a more in-depth analysis of the effect of transition probabilities on the optimal solution of our model. In the following, we present our analysis on three factors that can be characterized as gender-specific response, disease progression and treatment effectiveness.

3.5.2. Gender-specific transition probability matrices

As a good indication of the reaction of our model to changes in transition probability matrices, we run our baseline analysis with matrices estimated separately for males and females using the same age cut-off value. The estimated matrices for the no-screening action can be seen in Table 3. The transition probability estimates for the screening case were obtained as described in Section 3.1. When transition probabilities are compared on a one-on-one basis, we observe that gender-specific estimates do not deviate from each other or from the original estimates in big magnitudes in general. However, there are a few transitions where the deviation may be relatively large, the maximum difference being observed in the transition from state 4 to 2 as 0.35. Still, our model finds a never-screen policy optimal for both cases under the baseline assumptions.

3.5.3. Investigating the effect of disease progression

In order to analyze the effect of changes in transition probability estimates that correspond to disease progression, we establish parameters for transitions into next disease stage. In the model built according to the baseline description, let μ, θ and ϵ denote incremental probabilities of moving from cognitively normal to MCI, from MCI to mild and from mild to moderate and severe stages respectively. As we decrease the probability of staying in a current stage, we increase the probability of moving into the next stage using the associated parameter. For instance, as the probability of staying in the cognitively normal state decreases by μ, the probability of moving into MCI increases by μ. This increase is allocated to two MCI states in proportion to their original probability estimates. For transitions from MCI to mild, as we decrease the probability of staying in MCI without treatment with θ, we increase the probability of transitioning to mild without treatment with the same quantity. ϵ captures transitions from mild without treatment to moderate and severe in a similar way. We change one parameter at a time in its valid range at increments of 0.01. For values of μ between 0 and 0.891, we always find a never screen optimal policy. When we investigate values of θ up to 0.671, we find a never-screen policy optimal until $\theta = 0.28$. From then on, the optimal policy starts to include screenings at certain periods. At $\theta = 0.29$, the optimal policy is to screen at ages 60, 61 and 62. The number of screenings increases with θ, suggesting screening every year from 60 to 71 at $\theta = 0.36$

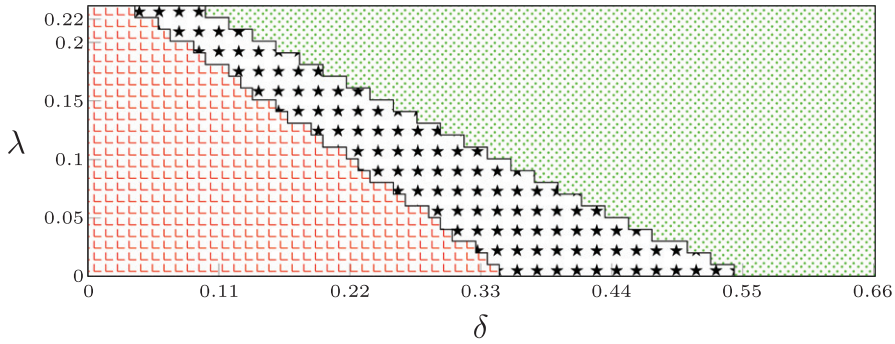

Figure 2. Screening policies with respect to δ and λ. Never screen, Screen yearly until age $d \in [94, 97]$, Screen yearly between $[74, b]$ with $b \leq 97$ or screen yearly between [a,b] and [c,d] with $60 \leq a<b<c<d \leq 97$.

and from 60 to 97 for values of $\theta \geq 0.46$. Finally, when probability of staying in mild without treatment is decreased by ϵ as the probability of transitioning into moderate and severe state increases, a never-screen policy remains optimal for all possible values of ϵ between 0 and 0.356. When all parameters are changed simultaneously, we find that the never-screen policy does not change for values less than 0.20. Therefore, we conclude that the no-screening policy is not sensitive to reasonable changes in probability estimates related to disease progression rates.

3.5.4. Investigating the effect of treatment effectiveness

We characterize the impact of treatment effectiveness through certain transition probabilities in our baseline model. A screening action results in an increased likelihood of moving to state 1, the MCI with treatment state. We now want to explore what happens if, when receiving treatment, it is more likely for individuals to transition into state 0 or less likely to transition into state 3 while in state 1. Let δ denote the incremental probability of moving from state 1 to state 0 instead of staying in state 1. Let λ denote the incremental probability of staying in state 1 instead of moving from state 1 to state 3. Note that δ captures the capability of the treatment applied in state 1 in terms of improving cognitive ability and λ corresponds to the capability of the treatment applied in state 1 in terms of keeping cognitive ability stable. Therefore, these parameters can be considered as indicators of effectiveness of the treatment along two dimensions that are not necessarily independent from each other. The higher these values are, the more likely it is for the patients to improve their cognitive ability or remain stable.

Based on our transition probability matrices, the bounds on these parameters should be as follows:

$$\begin{aligned} 0 \leq \delta \leq p_t^{11}(a_t) = \min(0.662, 0.7) = 0.662 \\ 0 \leq \lambda \leq p_t^{13}(a_t) = \min(0.22, 0.258) = 0.22 \end{aligned}$$

We solve our model by changing δ and λ from initial values of 0 to their respective upper bounds at increments of 0.01. Figure 2 summarizes the resulting optimal policies based on values of δ and λ.

Three main regions can be observed in Figure 2. Recalling that the origin corresponds to current treatment effectiveness levels, we observe that no screening remains as the optimal policy around a significant neighborhood of current values. If we assume that δ and λ increase at the same level, a value of $\delta = \lambda = 0.15$ would take the optimal policy out of the no-screening zone. When δ and λ are large, we observe optimal policies where screening is conducted annually almost until the end of the horizon. Again, if we assume that they change in the same way, a value of $\delta = \lambda = 0.18$ would result in an optimal policy where screening is recommended every year between the ages 60 to 95. To exemplify what happens in the third zone in between, we start tracing the path from 0.14 keeping $\delta = \lambda$. At 0.14, the optimal policy is not to screen, at 0.15 to screen every year between the ages 74 and 84, at 0.16 to screen every year between the ages 74 and 94, at 0.17 to screen every year between the ages 60 and 66 and 74 and 95, and at 0.18 to screen every year between the ages 60 to 95. To generalize the patterns we observe in this region, we can state that for moderate to high values of λ and low to moderate values of δ, the optimal policy is to screen every year between the ages 74 and 97. For low to moderate values of λ and moderate to high values of δ, the optimal policy is to screen every year starting at age 74 until an age before 97 or to screen every year for two distinct periods of time $[a, b]$ and $[c, d]$ where $60 \leq a<b<c<d \leq 97$.

The majority, but not all, of the policies we observe in our experimentation are policies where the optimal action changes only once in the horizon and remains the same afterwards. These policies are easy to implement in the sense that they can be expressed via simple rules that depend on the age of individuals. There are some optimal policies where the screening schedule covers a number of periods at the beginning of the horizon, skips some periods and covers a number of periods again. In addition to the $\delta = \lambda = 0.17$ case described earlier, another example is with $\delta = 0.25$ and $\lambda = 0.12$, where the optimal policy is to screen yearly between ages 62–66 and 74–96. We believe this type of policy is related to the two separate transition probability matrices we use based on age.

3.5.5. A closer look at costs and QALYs associated with optimal policies

So far, our focus has been on the nature of the optimal policy delivered by our model. We now want to have a closer look at the cost and QALY components of the value function associated with the optimal policies we presented

Table 4. Optimal values for the cognitively normal state (QALY, cost ($), $V_0^*(0)$($)).

		δ												
		0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.66
λ	0.22	16.55	16.9	**17.25**	**17.58**	**17.84**	**18.04**	**18.21**	**18.34**	**18.46**	**18.55**	**18.64**	**18.71**	**18.78**	**18.83**	**18.84**
		413,008	412,978	**419,258**	**414,504**	**410,556**	**407,395**	**404,695**	**402,406**	**400,441**	**398,734**	**397,239**	**395,918**	**394,743**	**393,690**	**393,492**
		1,242,298	1,281,722	**1,319,336**	**1,356,985**	**1,385,845**	**1,408,662**	**1,427,158**	**1,442,451**	**1,455,306**	**1,466,264**	**1,475,714**	**1,483,948**	**1,491,187**	**1,497,601**	**1,498,797**
	0.2	16.41	16.73	17.02	**17.32**	**17.59**	**17.81**	**17.99**	**18.13**	**18.26**	**18.37**	**18.46**	**18.54**	**18.62**	**18.68**	**18.69**
		414,747	409,424	412,564	**417,294**	**413,473**	**410,294**	**407,550**	**405,198**	**403,159**	**401,375**	**399,800**	**398,400**	**397,148**	**396,020**	**395,808**
		1,226,220	1,264,100	1,295,070	**1,328,108**	**1,358,578**	**1,382,977**	**1,402,962**	**1,419,626**	**1,433,735**	**1,445,835**	**1,456,325**	**1,465,508**	**1,473,613**	**1,480,820**	**1,482,166**
	0.15	16.14	16.46	16.69	16.89	17.08	**17.31**	**17.51**	**17.68**	**17.83**	**17.96**	**18.07**	**18.17**	**18.26**	**18.33**	**18.35**
		417,366	412,424	409,014	413,646	412,216	**415,408**	**412,820**	**410,496**	**408,436**	**406,598**	**404,949**	**403,460**	**402,110**	**400,880**	**400,647**
		1,196,822	1,233,614	1,260,132	1,281,268	1,303,009	**1,328,826**	**1,351,172**	**1,370,154**	**1,386,478**	**1,400,667**	**1,413,114**	**1,424,122**	**1,433,926**	**1,442,714**	**1,444,364**
	0.1	15.95	16.26	16.48	16.66	16.81	16.98	17.13	17.31	**17.47**	**17.61**	**17.73**	**17.84**	**17.94**	**18.03**	**18.05**
		418,649	414,248	411,044	408,608	406,693	413,770	415,022	414,101	**412,144**	**410,368**	**408,750**	**407,269**	**405,909**	**404,656**	**404,417**
		1,177,173	1,211,742	1,237,770	1,258,062	1,274,321	1,291,364	1,309,377	1,329,333	**1,346,949**	**1,362,435**	**1,376,158**	**1,388,402**	**1,399,394**	**1,409,317**	**1,411,187**
	0.05	15.82	16.1	16.32	16.5	16.65	16.77	16.9	17.03	17.16	17.31	**17.44**	**17.56**	**17.66**	**17.76**	**17.78**
		419,300	415,418	412,469	410,152	408,282	406,740	414,818	413,137	414,026	413,133	**411,597**	**410,174**	**408,853**	**407,624**	**407,388**
		1,163,258	1,195,369	1,220,382	1,240,413	1,256,814	1,270,490	1,283,295	1,297,330	1,313,578	1,329,693	**1,344,222**	**1,357,286**	**1,369,096**	**1,379,825**	**1,381,855**
	0	15.72	15.99	16.2	16.37	16.51	16.64	16.74	16.83	16.96	17.05	17.18	**17.3**	**17.42**	**17.52**	**17.54**
		419,635	416,204	413,508	411,331	409,535	408,027	406,743	406,006	414,001	413,761	413,379	**412,418**	**411,162**	**409,981**	**409,754**
		1,152,937	1,182,688	1,206,500	1,225,989	1,242,235	1,255,985	1,267,775	1,278,003	1,289,991	1,301,896	1,316,450	**1,329,957**	**1,342,317**	**1,353,608**	**1,355,750**

earlier. Table 4 presents optimal QALY and cost values associated with the cognitively normal state along with the value of $V_0(0)$ as a function of treatment levels experimented with in Section 3.5.4. Values in bold correspond to screening yearly until age $d \in [94, 97]$ type of optimal policy. The values associated with the baseline model are $V_0(0) = 1,152,937$ where a decomposition into QALYs and costs yields 15.72 for expected QALY and $\$419,635$ for expected cost. Not surprisingly, QALYs increase and costs decrease as δ and λ increase. We observe that expected costs may decline by as much as approximately 6% and expected QALYs may increase by as much as 20% as treatment effectiveness increases. In that regard, it appears that higher levels of treatment effectiveness have a higher impact on optimal expected QALYs than expected costs. Both changes are observed to be monotone along both parameter dimensions.

Finally, to isolate gains resulting from a screening policy at varying levels of treatment effectiveness, we compare values associated with our optimal policies with values associated with a never-screen policy. Table 5 contains comparative information between optimal policies and a never-screen policy at the associated levels of the parameters. The reported figures are differences in expected QALYs, costs and values associated with the cognitively normal state. As in Table 4, values in bold are associated with screening yearly until age $d \in [94, 97]$ type of optimal policy. We note that, for treatment effectiveness levels where we find screening every year optimal, the incremental expected cost per expected QALY gain is quite low, leading to significant improvements in optimal value.

4. Conclusion and future work

This is a first attempt to compute optimal population screening policies for AD. Our model is a simple MDP model that uses MMSE as a screening test, which is a widely used tool in practice. We use data from the literature and ADNI project to implement the model. The objective function is based on a combination of QALYs and costs. We estimate transition probability matrices using age as a covariate in a multistate Markov model. We find that a no-screening policy is optimal in our baseline implementation. When the sensitivity of the policy to various model parameters is investigated, the policy appears to be quite robust to changes in discount factor, cost-effectiveness ratio and screening test accuracy. We also analyze sensitivity of our results to changes in the transition probability estimates. Our baseline finding of optimality of a never-screen policy does not change when gender-specific transition probability estimates are used. In addition, we study sensitivity to changes in transitions related to disease progression and drug effectiveness in more detail. Overall, we observe that the baseline result that finds a never-screen policy optimal does not change when small changes take place in the transition probability estimates.

Through a parametric analysis, we observe that the optimal policy may be very sensitive to treatment effectiveness. When treatment plans that lead to higher cognitive ability

Table 5. Differences in expected QALYs, costs ($) and value functions ($) for the cognitively normal state: optimal policy versus never-screen policy.

								δ							
	0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.66
0.22	0	0.01	**0.1**	**0.24**	**0.35**	**0.43**	**0.51**	**0.57**	**0.62**	**0.66**	**0.7**	**0.73**	**0.76**	**0.78**	**0.79**
	0	3,942	**13,277**	**11,198**	**9,425**	**7,926**	**6,635**	**5,520**	**4,547**	**3,692**	**2,935**	**2,260**	**1,655**	**1,109**	**1,006**
	0	1,817	**14,249**	**34,065**	**49,663**	**62,239**	**72,591**	**81,254**	**88607**	**94,927**	**100,414**	**105,224**	**109,474**	**113,256**	**113963**
0.2	0	0	0.03	**0.13**	**0.24**	**0.34**	**0.42**	**0.47**	**0.53**	**0.58**	**0.62**	**0.65**	**0.69**	**0.72**	**0.72**
	0	0	4376	**11,739**	**10,129**	**8,727**	**7,499**	**6,421**	**5,467**	**4,620**	**3,862**	**3,181**	**2,565**	**2,006**	**1,901**
	0	0	5,011	**19,118**	**35,272**	**48,476**	**59,466**	**68,748**	**76,690**	**83,560**	**89,561**	**94,847**	**99,539**	**103,731**	**104,517**
λ0.15	0	0	0	0.01	0.03	**0.12**	**0.21**	**0.28**	**0.34**	**0.4**	**0.45**	**0.49**	**0.53**	**0.56**	**0.57**
	0	0	0	4,757	4,612	**9,688**	**8,695**	**7,786**	**6,954**	**6,191**	**5,491**	**4,847**	**4,254**	**3,707**	**3,602**
	0	0	0	1,110	6,954	**19,995**	**31,854**	**42,076**	**50,972**	**58,783**	**65,694**	**71,852**	**77,373**	**82,350**	**83,287**
0.1	0	0	0	0	0	0.02	0.05	0.12	**0.19**	**0.25**	**0.3**	**0.35**	**0.39**	**0.43**	**0.44**
	0	0	0	0	0	4,999	7,850	8,333	**7,661**	**7,025**	**6,424**	**5,859**	**5,327**	**4,827**	**4,730**
	0	0	0	0	0	3,551	10,215	20,547	**29,900**	**38,216**	**45,655**	**52,349**	**58,403**	**63,904**	**64,944**
0.05	0	0	0	0	0	0	0.01	0.03	0.06	0.13	**0.18**	**0.23**	**0.27**	**0.32**	**0.33**
	0	0	0	0	0	0	5,233	4,949	7,129	7,396	**6,906**	**6,433**	**5,977**	**5,540**	**5,455**
	0	0	0	0	0	0	1,176	5,070	12,515	20,918	**28,634**	**35,636**	**42,018**	**47,857**	**48,966**
0	0	0	0	0	0	0	0	0	0.02	0.02	0.07	**0.12**	**0.17**	**0.21**	**0.22**
	0	0	0	0	0	0	0	76	5,137	6,048	6,714	**6,712**	**6,336**	**5,968**	**5,896**
	0	0	0	0	0	0	0	7	2,901	6,735	14,106	**21,182**	**27,747**	**33,792**	**34,944**

improvement or stabilization become available, implementing a population screening policy may become socially optimal. This places our study more towards the studies of Dixon *et al.* (2014), who examined a hypothetical screening program for individuals over 75 in England and Wales by means of a static decision model, and Yu *et al.* (2015), who studied an opportunistic screening program in Korea using simulation. Although our study is not directly comparable to these studies because of the different assumptions dictated by different modeling environments and choices, we provide additional support to demonstrate the impact of treatment effectiveness on outcomes.

As we see treatment effectiveness as the most dynamic factor among the ones we study, we suggest that transition probability estimation should be conducted carefully when better treatment effectiveness levels are observed due to emerging options. Our parametric analysis on treatment effectiveness with gender-specific transition probability matrices suggests that it may be worthwhile to consider gender-specific policies if and when higher treatment effectiveness levels are reached. We also note that a careful reassessment of other model parameters, such as costs and rewards, should be taken into account using the characteristics of the population in focus, as these may show significant variations in different settings and across time.

An analysis of the optimal values of the value function associated with the cognitively normal state reveals the magnitude of possible gains at varying levels of treatment effectiveness. A decomposition of the optimal value function with respect to expected QALYs and costs through the horizon reveals that QALY improvements may be more significant than cost improvements. A comparison of the optimal screening policy to a no-screening policy at experimented levels of treatments effectiveness demonstrates benefits of screening. We suggest that policymakers should assess the effectiveness of emerging treatment alternatives periodically and consider revising their population screening decisions.

Findings in medical literature indicate that AD seems irreversible when cognitive decline becomes detectable via tools like MMSE. Therefore, treatments that are highly effective at cognitively evident stages of AD may not be highly probable. The focus in the medical literature has thus shifted to diagnosing AD at a pre-clinical stage. In 2011, the criteria and guidelines for diagnosis of AD were revised and updated by the National Institute on Aging (NIA) and the Alzheimer's Association (AA). The most important differences in criteria were the introduction of the preclinical aspect of the disease. The NIA and AA propose criteria for the preclinical stages of AD and introduce the biomarker tests for Aβ and MAP-τ as possible clue providers for the onset of AD (Albert *et al.*, 2011; Dubois *et al.*, 2007; Sperling *et al.*, 2011). Research is ongoing on treatment plans that target Aβ elimination whose accumulation is thought to result in the neurodegeneration that leads to AD. A natural extension of our study would therefore be to build a model that suggests screening using biomarkers with the assumption that positive individuals are treated at a pre-clinical stage. There are several challenges associated with such a model. First of all, there is little information on how individuals progress from pre-clinical to MCI and finally to AD (Hampel *et al.*, 2008; Henriksen *et al.*, 2014). Magnetic resonance imaging scans and positron emission tomography scans are costly means of biomarker detection. The alternative method of cerebrospinal fluid collection causes discomfort and comes with some risks. There are no universal guidelines for cut-off levels to be applied to biomarker specimens for AD diagnosis. Large-scale data for treatment effectiveness and cost of treatment are not available on experimental therapies. Still, a partially observable MDP may be better suited to build such a model where a cognitively normal state is replaced with a collection of partially observable pre-clinical states, including a disease-free state. The results of the biomarker tests can be modeled to reveal partial information that leads to belief updates with respect to states of the model. The initial belief regarding an individual's being in one of the pre-clinical states of AD may incorporate their characteristics and risk factors, thus making personalized screening policy recommendations possible. As more data is collected in ongoing studies, a realistic implementation of such a model would become more plausible.

Acknowledgement

The authors would like to thank anonymous reviewers and the associate editor whose constructive feedback and suggestions have improved the quality of the article significantly.

References

AD2000 Collaborative Group (2004) Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): Randomised double-blind trial. *The Lancet*, **363**(9427), 2105–2115.

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., et al. (2011) 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. *Alzheimer's & Dementia*, **7**(3), 270–279.

Alemayehu, B. and Warner, K. E. (2004) The lifetime distribution of health care costs. *Health Services Research*, **39**(3), 627–642.

Alzheimer's Disease Neuroimaging Initiative (2006). ADNI Procedures Manual. Available at: http://adni.loni.usc.edu/wp-content/uploads/2010/09/ADNI_GeneralProceduresManual.pdf.

Andersen, P. K. (1988) Multistate models in survival analysis: A study of nephropathy and mortality in diabetes. *Statistics in Medicine*, **7**(6), 661–670.

Alzheimer's Association. (2017) 2017 Alzheimer's disease facts and figures. Alzheimer's & Dementia, **13**(4), 325–373.

Ayer, T., Alagöz, O., and Stout, N. K. (2012) OR forum: A POMDP approach to personalize mammography screening decisions. *Operations Research*, **60**(5), 1019–1034.

Barnett, J. H., Lewis, L., Blackwell, A. D., and Taylor, M. (2014) Early intervention in Alzheimer's disease: A health economic study of the effects of diagnostic timing. *BMC Neurology*, **14**(1), 101.

Bond, J., Graham, N., Padovani, A., Mackell, J., Knox, S., and Atkinson, J. (2010) Screening for cognitive impairment, Alzheimer's disease and other dementias: Opinions of European caregivers, payors, physicians and the general public. *The Journal of Nutrition, Health & Aging*, **14**(7), 558–562.

Brazier, J. (2007) Measuring and Valuing Health Benefits for Economic Evaluation. Oxford University Press, Oxford.

Cohen, J. T. and Neumann, P. J. (2008) Decision analytic models for Alzheimer's disease: State of the art and future directions. *Alzheimer's & Dementia*, **4**(3), 212–222.

Cullen, B., O'Neill, B., Evans, J. J., Coen, R. F., and Lawlor, B. A. (2007) A review of screening tests for cognitive impairment. *Journal of Neurology, Neurosurgery & Psychiatry*, **78**(8), 790–799.

Dixon, J., Ferdinand, M., D'Amico, F., and Knapp, M. (2014) Exploring the cost-effectiveness of a one-off screen for dementia (for people aged 75 years in England and Wales). *International Journal for Geriatric Psychiatry*, **30**(5), 446–452.

Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., et al. (2007) Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS-ADRDA criteria. *The Lancet Neurology*, **6**(8), 734–746.

Duthey, B. (2013) Background paper 6.11: Alzheimer disease and other dementias. *A Public Health Approach to Innovation*, 1–74.

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975) Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, **12**(3), 189–198.

Getsios, D., Blume, S., Ishak, K. J., Maclaine, G., and Hernández, L. (2012) An economic evaluation of early assessment for Alzheimer's disease in the United Kingdom. *Alzheimer's & Dementia*, **8**(1), 22–30.

Gillies, C. L., Lambert, P. C., Abrams, K. R., Sutton, A. J., Cooper, N. J., Hsu, R. T., Davies, M. J., Khunti, K., et al. (2008) Different strategies for screening and prevention of type 2 diabetes in adults: Cost effectiveness analysis. *British Medical Journal*, **336**(7654), 1180–1185.

Green, C. (2007) Modelling disease progression in Alzheimer's disease. *Pharmacoeconomics*, **25**(9), 735–750.

Green, C., Shearer, J., Ritchie, C. W., and 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**(5), 621–630.

Hampel, H., Bürger, K., Teipel, S. J., Bokde, A. L., Zetterberg, H., and Blennow, K. (2008) Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. *Alzheimer's & Dementia*, **4**(1), 38–48.

Harper, P. and Jones, S. (2005) Mathematical models for the early detection and treatment of colorectal cancer. *Health Care Management Science*, **8**(2), 101–109.

Henriksen, K., O'Bryant, S. E., Hampel, H., Trojanowski, J. Q., Montine, T. J., Jeromin, A., Blennow, K., Lönneborg, A., Wyss-Coray, T., Soares, H., et al. (2014) The future of blood-based biomarkers for Alzheimer's disease. *Alzheimer's & Dementia*, **10**(1), 115–131.

Jackson, C. (2011) Multi-state models for panel data: The MSM package for r. *Journal of Statistical Software*, **38**(1), 1–28. .

Jackson, C. H., Sharples, L. D., Thompson, S. G., Duffy, S. W., and Couto, E. (2003) Multistate Markov models for disease progression with classification error. *Journal of the Royal Statistical Society: Series D (The Statistician)*, **52**(2), 193–209.

Kaduszkiewicz, H., Zimmermann, T., Beck-Bornholdt, H.-P., and van den Bussche, H. (2005) Cholinesterase inhibitors for patients with Alzheimer's disease: Systematic review of randomised clinical trials. *British Medical Journal*, **331**(7512), 321–327.

Khachaturian, Z. S., Barnes, D., Einstein, R., Johnson, S., Lee, V., Roses, A., Sager, M. A., Shankle, W. R., Snyder, P. J., Petersen, R. C., et al. (2010) Developing a national strategy to prevent dementia: Leon Thal Symposium 2009. *Alzheimer's & Dementia*, **6**(2), 89–97.

Khachaturian, Z. S., Khachaturian, A. S., and Thies, W. (2012) The draft national plan to address Alzheimer's disease: National Alzheimer's Project Act (NAPA). *Alzheimer's & Dementia*, **8**(3), 234–236.

Khachaturian, Z. S., Petersen, R. C., Snyder, P. J., Khachaturian, A. S., Aisen, P., De Leon, M., Greenberg, B. D., Kukull, W., Maruff, P., Sperling, R. A., et al. (2011) Developing a global strategy to prevent Alzheimer's disease: Leon Thal Symposium 2010. *Alzheimer's & Dementia*, **7**(2), 127–132.

Kirby, J., Green, M. C., Loveman, E., Clegg, A., Picot, J., Takeda, A., and Payne, E. (2006) A systematic review of the clinical and cost-effectiveness of memantine in patients with moderately severe to severe Alzheimer's disease. *Drugs & Aging*, **23**(3), 227–240.

K irk izlar, E., Faissol, D. M., Griffin, P. M., and Swann, J. L. (2010) Timing of testing and treatment for asymptomatic diseases. *Mathematical Biosciences*, **226**(1), 28–37.

Kmietowicz, Z. (2012) Cameron launches challenge to end "national crisis" of poor dementia care. British Medical Journal, **344**, e2347.

Koepsell, T. D. and Monsell, S. E. (2012) Reversion from mild cognitive impairment to normal or near-normal cognition risk factors and prognosis. *Neurology*, **79**(15), 1591–1598.

Kurt, M., Denton, B. T., Schaefer, A. J., Shah, N. D., and Smith, S. A. (2011) The structure of optimal statin initiation policies for patients with type 2 diabetes. *IIE Transactions on Healthcare Systems Engineering*, **1**(1), 49–65.

Leon, J., Cheng, C.-K., and Neumann, P. J. (1998) Alzheimer's disease care: costs and potential savings. *Health Affairs*, **17**(6), 206–216.

Leon, J. and Neumann, P. J. (1999) The cost of Alzheimer's disease in managed care: A cross-sectional study. *American Journal of Managed Care*, **5**(7), 867–877.

López-Bastida, J., Hart, W., García-Pérez, L., and Linertová, R. (2009) Cost-effectiveness of donepezil in the treatment of mild or moderate Alzheimer's disease. *Journal of Alzheimer's Disease*, **16**(2), 399–407.

MacDonald, I. L. and Zucchini, W. (1997) *Hidden Markov and other Models for Discrete-valued Time Series*, volume **110**. CRC Press, Boca Raton, FL.

Maillart, L. M., Ivy, J. S., Ransom, S., and Diehl, K. (2008) Assessing dynamic breast cancer screening policies. *Operations Research*, **56**(6), 1411–1427.

Mangialasche, F., Solomon, A., Winblad, B., Mecocci, P., and Kivipelto, M. (2010) Alzheimer's disease: Clinical trials and drug development. *The Lancet Neurology*, **9**(7), 702–716.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. (1984) Clinical diagnosis of Alzheimer's disease report of the NINDCS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, **34**(7), 939–939.

Morris, J. C. (1993) The clinical dementia rating (CDR): Current version and scoring rules. *Neurology*, **43**(11), 2412–2414.

Morris, J. C. (1997) Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *International Psychogeriatrics*, **9**(S1), 173–176.

National Institute for Clinical Excellence (NICE) et al. (2011) Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease. NICE Technology Appraisal Guidance, **217, 1-76**.

Neumann, P., Araki, S., Arcelus, A., Longo, A., Papadopoulos, G., Kosik, K. A., Kuntz, K., and Bhattacharjya, A. (2001) Measuring Alzheimer's disease progression with transition probabilities estimates from CERAD. *British Medical Journal*, **57**(6), 957–964.

Neumann, P. J., Kuntz, K. M., Leon, J., Araki, S. S., Hermann, R. C., Hsu, M.-A., and Weinstein, M. C. (1999) Health utilities in Alzheimer's disease: A cross-sectional study of patients and caregivers. *Medical Care*, **37**(1), 27–32.

Özekici, S. and Pliska, S. R. (1991) Optimal scheduling of inspections: A delayed Markov model with false positives and negatives. *Operations Research*, **39**(2), 261–273.

Parmigiani, G. (1993) On optimal screening ages. *Journal of the American Statistical Association*, **88**(422), 622–628.

Perneczky, R., Wagenpfeil, S., Komossa, K., Grimmer, T., Diehl, J., and Kurz, A. (2006) Mapping scores onto stages: Mini-mental state examination and clinical dementia rating. American Journal of Geriatric Psychiatry, **14**(2), 139-144.

Prince, M. J. (2015) *World Alzheimer Report 2015: The Global Impact of Dementia: An Analysis of Prevalence, Incidence, Cost and Trends*. Alzheimer's Disease International, London, England.

Puterman, M. L. (1994) Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons, Hoboken, New Jersey.

Satten, G. A. and Longini Jr., I. M. (1996) Markov chains with measurement error: Estimating the true course of a marker of the progression of human immunodeficiency virus disease. Applied Statistics, 45(3), 275–309.

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., Iwatsubo, T., Jack, C. R., Kaye, J., Montine, T. J., et al. (2011) 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. *Alzheimer's & Dementia*, **7**(3), 280–292.

Touchon, J., Lachaine, J., Beauchemin, C., Granghaud, A., Rive, B., and 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. *European Journal of Health Economics*, **15**(8), 791–800.

UK National Screening Committee (2015). *Screening in the UK: Making Effective Recommendations 2014 to 2015*. Public Health England, London.

Underwood, D., Zhang, J., Denton, B., Shah, N., and Inman, B. (2012) Simulation optimization of PSA-threshold based prostate cancer screening policies. *Health Care Management Science*, **15**(4), 293–309.

Weimer, D. L. and Sager, M. A. (2009) Early identification and treatment of Alzheimer's disease: Social and fiscal outcomes. *Alzheimer's & Dementia*, **5**(3), 215–226.

World Health Organization. (2012) Dementia: A Public Health Priority. World Health Organization, Geneva.

Youden, W. J. (1950) Index for rating diagnostic tests. *Cancer*, **3**(1), 32–35.

Yu, S.-Y., Lee, T.-J., Jang, S.-H., Han, J. W., Kim, T. H., and Kim, K. W. (2015) Cost-effectiveness of nationwide opportunistic screening program for dementia in South Korea. *Journal of Alzheimer's Disease*, **44**(1), 195–204.

Zhang, J., Denton, B. T., Balasubramanian, H., Shah, N. D., and Inman, B. A. (2012) Optimization of prostate biopsy referral decisions. *Manufacturing & Service Operations Management*, **14**(4), 529–547.

Appendix. The multistate model

A multi-state model (MSM) is a model for a stochastic process allowing individuals to move among a finite number of states. States can be either transient or absorbing, if no transitions can occur from that state. In the disease progression framework, multistate Markov models based on the Markovian assumption in discrete and continuous time are generally used. Typically, in this framework, the stages of the disease form a homogeneous continuous time Markov process and the individuals may go forward to or backward from adjacent states or transit to an absorbing state (usually death, but not necessarily). MSM can be adopted in a more general setting where transitions from any state to another can happen. The state of the condition of a particular individual k, recorded at arbitrary times t, is denoted as $s_k(t)$. Here, the states of the disease are modeled as a homogeneous continuous time Markov process and q^{ij} represents the transition rate from state i to state j. The corresponding transition intensity matrix denoted by $\mathbf{Q}$ where state n is an absorbing state is:

$$\mathbf{Q} = q^{ij} = \begin{pmatrix} q^{11} & q^{12} & \cdots & q^{1n} \\ q^{21} & q^{22} & \cdots & q^{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q^{n-1,1} & q^{n-1,2} & \cdots & q^{n-1,n} \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

When no instant transitions are allowed as in an irreversible disease progression framework, then respective q_{ij} values might be set equal to zero. The rows of transition intensity matrix must add up to 0, hence for the diagonal entries we have:

$$q^{ii} = -\sum_{i \neq j} q^{ij}, \forall i = 1, \ldots, n$$

Maximum likelihood estimates for MSM models can be computed via the transition probability matrix P_t at time t that has the following entries for any individual

$$p_t^{ij} = Pr\{s(u) = j | s(t) = i\} \tag{2}$$

where $u \geq t$. Assuming $s(0) = i$, the forward equations can be written for $\Delta t > 0$ as:

$$p_{(t+\Delta t)}^{ik} = p_t^{ik}\left(1 + q^{kk}\Delta t\right) + \sum_{j \neq k} p_t^{ij} q^{jk} \Delta t + o(\Delta t)$$

where $o(\Delta t) \to 0$ as $\Delta t \to 0$. From here and a similar backward equation, Kolmogorov differential equations are derived (Jackson *et al.*,2003). The solution of these equations with initial condition $\mathbf{P}_0 = \mathbf{I}$ is

$$\mathbf{P}_t = e^{\mathbf{Q}t} \tag{3}$$

These models have been used in the medical literature for different types of diseases such as cancer, diabetes, HIV (Andersen, 1988; Jackson et al., 2003; Satten and Longini Jr., 1996).

The likelihood associated with a transition rate matrix $\mathbf{Q}$ is given by the product of probabilities of transition between observed states, over all individuals k and observation times j and a maximum likelihood estimate of $\mathbf{Q}$ is obtained by optimization.

The estimation process may incorporate covariates. There are various ways of relating covariates to dependent variables. Traditionally, generalized regressions are used with proportional hazard model to relate transition intensities q_t^{ij} at time t to covariates $z(t)$ at that time via

$$q_{(t,z(t))}^{ij} = q^{ij} e^{\beta_{ij}^T z(t)}$$

The likelihood function is updated with respect to this $\mathbf{Q}$ and the optimization is solved to find estimates of β as well. This may be a difficult optimization problem to solve numerically depending on problem size and characteristics. One algorithm employed is the expectation-maximization (EM) algorithm. This algorithm alternates between two steps: the first one computes the expectation of the likelihood function given parameters; the second one updates the values of the parameters in an effort to improve the likelihood function. The two steps are repeated until convergence to a local maximum of L. The adaptation of the EM algorithm for different settings of models is described in MacDonald and Zucchini (1997) and Satten and Longini Jr. (1996).