SUPPLEMENTARY MATERIAL

S1. Transition Probabilities

Below are the annual transition probabilities (per year) used in the multistate model in Figure 1. The references on which they are based are provided. Here $r_{ij}(a)$ is the probability of transitioning from state *i* at age *a* to stage *j* at age *a*+1. Also, $d_i\,(a, g)$ are the annual death rates for a person of age *a* and gender *g* who is in state *i*.

	Transition rate per year	References
$r_{12}(a)$	$.000149e^{.086a}$ for $a \le 65$ $.0010527e^{.05596a}$ for $65 < a \le 75$ $.07$ for $a > 75$	[1]
$r_{13}(a)$	$7.53\times10^{-6}e^{0.117a}$	[1]
$r_{24}(a)$	$.0011e^{0.061a}$	[1]
$r_{34}(a)$	$.00012e^{0.082a}$	[1]
$r_{45}(a)$	$2.5958\times10^{-5}\,\mathrm{e}^{.096a}$	[1,2,3]
$r_{36}(a)$	$.00327e^{.0198a}$ *for* $a \le 75$ $.00327e^{.0198(75)}$ *for* $a \ge 75$	[1,2,3]
r_{57}	0.30	[3]
r_{67}	0.09	[3]
$d_i(a,g)$ States *i*=1 to 4	U.S death rates (2014)	[4]
$d_i(a,g)$ States *i*=5 & 6	1.65 x U.S death rates(2014)	[4-9]

S.2 Calculation of $P_k(a, g, I, J)$

The probability that an individual in study arm *k* (control =1 or intervention=2) who is age *a* (in years), gender *g*, and is initially in preclinical state *I* at enrollment reaches state *J* during *T* years of follow-up is

called $P_k(a,g,I,J)$ where state J is the primary study endpoint (e.g. MCI state 5 or Alzheimer's dementia state 7). Here we use the notation $P_k(a,g,I,J)$ rather than simply P_k to emphasize that the probability depends on enrollment age, gender, the initial state (I) and the final primary endpoint (J). This probability, $P_k(a,g,I,J)$, is the summation of probabilities of entering state J at exactly age $a+n(.25)$ for values of n ranging from 1 to $4T$ (the factor .25 arises because 1 time unit=.25 years; the term $4T$ arises because there are $4T$ time units over T years if one time unit=.25 years); that is,

$$P_k(a,g,I,J) = \sum_{n=1}^{4T} \phi(n \mid a,g,I,J) \quad (1)$$

where $\phi(n \mid a,g,I,J)$ is the probability of entering state J at exactly age $a+n(.25)$ for a person who at enrollment is in disease state I at age a. It turns out that the terms $\phi(n \mid a,g,I,J)$ can be calculated by matrix multiplication of one step transition matrices as described next.

We assume the disease process illustrated in Figure 1 in the paper is a discrete nonhomogeneous multistate Markov model. We use a discrete time model in which transitions into and out of states occur only at discrete time points. We allow only one transition in a time unit. The one step transition matrix $\mathbf{T}(a,g)$ gives the probabilities for an individual who is age a and gender g of transitioning from one state to another state in one time unit (one time unit=0.25 years). The entry in the i^{th} row and j^{th} column of the matrix $\mathbf{T}(a,g)$, called $p_{ij}(a,g)$, is the probability that a person who age a (in years), gender g and is in state i will transition to state j at age a+.25. We wish to calculate $\phi(n \mid a,g,I,J)$ which is the probability a person who is age a and in disease state I enters state J at exactly age $a+n(.25)$.

The n-step transition matrix $\mathbf{T}^{(n)}(a,g)$ is a matrix where the entry in the i^{th} row and j^{th} column is the probability that an individual who is age a and is in state i will be in state j at age $a+n(.25)$. The n-step transition matrix $\mathbf{T}^{(n)}(a,g)$ can be calculated by matrix multiplication, and is the product of n one step transition matrices; specifically, the n-step transition matrix is

$$\mathbf{T}^{(n)}(a,g) = \prod_{i=1}^{n} \mathbf{T}(a+i-1,g)$$

The above result follows from the Chapman-Kolmogorov equation. The entry in the I^{th} row and J^{th} column of $\mathrm{T}^n(a,g)$ is the probability that an individual who is age *a* and in state *I* is in state *J* at age *a+n(.25)* . Now, that probability includes the possibilities that the individual first entered state *J* at any of the ages between (or equal to) *a*+.25 and *a+n(.25)*. However, the quantity we need is $\phi(n \mid a,g,I,J)$ which refers to the probability that a person enters state *J* at *exactly* age *a+n*(*.25*). Fortunately, we can also obtain $\phi(n \mid a,g,I,J)$ through matrix multiplication using the following adjustment. We define an adjusted one step transition matrix **W** where **W** is the same as the one step transition matrix **T** with the following important difference: the probability of death in one time unit for persons in states *J* is always set to 1. By making that adjustment, we remove the possibility that a person remains in state *J* for more than one time unit. Let

$$\mathbf{W}^n(a,g) = \prod_{i=1}^{n} \mathbf{W}(a+i-1,g).$$

Then the entry in the I^{th} row and J^{th} column of $\mathbf{W}^n(a,g)$ is $\phi(n \mid a,g,I,J)$ which we use in equation to obtain $P_k(a,g,I,J)$

S.3 Impact of Ages at Enrollment

We considered the impact of ages at enrollment on $P_k(a,g,I,J)$ and thus on sample size requirements. We assumed the numbers of males and females trial participants are approximately equal and that the distributions of ages at trial enrollment are the same for males and females. Let *w*(*a)* represent the proportion of trial participants who are enrolled at age *a*. Then, the probability of a primary endpoint occurring in group *k* is

$$P_k(I,J) = \sum_{ages\ a} \frac{w(a)\left\{P_k(a,\text{male},I,J) + P_k(a,\text{female},I,J)\right\}}{2}$$

We considered three age distributions: a young age distribution where *w*(*a*)= 0.25 at ages *a*=60, 65, 70 and 75 (and 0 at other ages) which has a mean age of 67.5; a uniform age distribution where *w*(*a*)= 0.167 at ages *a*=60, 65, 70, 75, 80 and 85 which has a mean age of 72.5; an old age distribution where *w*(*a*) =0.25 at ages a=70, 75, 80 and 85 which has a mean age of 77.5.

S.4 Impact of Preclinical Disease Enrollment Criteria

We also considered the impact of preclinical disease state at enrollment on $P_k(a,g,I,J)$ and thus on sample size requirements. For example, we considered secondary prevention trials that enrolled persons with amyloidosis who are in state 2 (I=2). We considered primary prevention trials that enrolled persons without any brain pathology who are in the normal state 1 (I=1). We also considered trials that did not screen with biomarkers for the specific preclinical disease state and instead enrolled a random sample (or cross-section) of persons in preclinical states (I=1, 2, 3 or 4). In the case of random sampling of persons who are preclinical, the probability of a primary event (state J) is

$$P_k(J)=\sum_{ages\ a}\sum_{I=1}^{4}\frac{w(a)\,f(I,a)\left\{P_k(a,\text{male},I,J)+P_k(a,\text{female},I,J)\right\}}{2}$$

where $f(I,a)$ is the proportion in individuals at age a who are preclinical (states1 through 4) that are in state I. These proportions can be obtained from estimates of the prevalence in each disease state.[1] It turns out that these proportions are nearly identical for males and females and so we use only one set of values for $f(I,a)$ rather than gender specific values. The proportion of persons at age a who are in state I among those who are preclinical (states 1, 2, 3 or 4) is denoted by $f(I,a)$. These proportions are used in the calculation of sample sizes when a random sample of persons in the preclinical states (states1-4) is enrolled. The numbers in the Table below are based on state specific prevalence rates by age.[2]

Table S1: Values of $f(I,a)$

age	I=1	I=2	I=3	I=4
60	.692	.189	.056	.063
65	.550	.241	.085	.124
70	.390	.266	.120	.224
75	.236	.249	.151	.364
80	.121	.185	.165	.529
85	.049	.108	.157	.686

The values *f(I, a)* given in Table S1 can also be used to calculate the factor to inflate sample sizes to account for numbers of persons who would need to be screened with biomarkers to reach the target

sample sizes. For example, consider a prevention trial that enrolls persons with amyloidosis (state 2). The prevalence of amyloidosis (state 2) among person without clinical disease assuming a uniform age distribution between ages 60 and 85 is approximately

$$\begin{aligned}&\sum_a w(a) f\left(I, a\right)\\&= .167(.189 + .241 + .266 + .249 + .185 + .108)\\&= .206\end{aligned}$$

Thus, the numbers of preclinical individuals who would need to be screened for biomarkers in order to enroll the required sample sizes is approximately (1/.206)=4.85 times the numbers shown in Table 1

Supplementary Material References

1. Jack, C. R., Therneau, T. M., Wiste, H. J., Weigand, S. D., Knopman, D. S., Lowe, V. J., ... & Senjem, M. L. Transition rates between amyloid and neurodegeneration biomarker states and to dementia: a population-based, longitudinal cohort study *The Lancet Neurology*. 2016; *15*: 56-64.
2. Brookmeyer R, Abdalla N, Kawas CH, Corrada MM. Forecasting the prevalence of preclinical and clinical AD in the United States. *Alzheimer's & Dementia* 2018; 14(2):121-9.
3. Vos S.J., Verhey F., Frölich L, Kornhuber J., Wiltfang J., Maier W., Peters O., Rüther E., Nobili F., Morbelli S., Frisoni G.B. Prevalence and prognosis of Alzheimer's disease at the mild cognitive impairment stage. *Brain* 2015; 138:1327-1338.
4. Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org (downloaded on 4/1/2016).
5. Vassilaki M, Cha RH, Aakre JA, Therneau TM, Geda YE, Mielke MM, Knopman DS, Petersen RC, Roberts RO. Mortality in mild cognitive impairment varies by subtype, sex, and lifestyle factors: the Mayo Clinic Study of Aging. *Journal of Alzheimer's Disease* 2015; 45:1237-45
6. Bennett DA, Wilson RS, Schneider JA, Evans DA, Beckett LA, Aggarwal NT, Barnes LL, Fox JH, Bach J. Natural history of mild cognitive impairment in older persons. *Neurology* 2002; 59:198-205.
7. Wilson RS, Aggarwal NT, Barnes LL, Bienias JL, de Leon CF, Evans DA. Biracial population study of mortality in mild cognitive impairment and Alzheimer disease. *Archives of Neurology.* 2009; 66:767-72.
8. Contador I, Bermejo-Pareja F, Mitchell AJ, Trincado R, Villarejo A, Sánchez-Ferro Á, Benito-León J. Cause of death in mild cognitive impairment: a prospective study (NEDICES). *European Journal of Neurology*. 2014; 21:253.
9. Luck T, Riedel-Heller SG, Roehr S, Wiese B, Leeden C, Heser K, Bickel H, Pentzek M, König HH, Werle J, Mamone S. Mortality in Incident Cognitive Impairment: Results of the Prospective AgeCoDe Study. *Journal of the American Geriatrics Society.* 2017; 65:738-46.