Lifetime risks, prevalence, and incidence estimates from a multistate model for Alzheimer's disease

Tommy Gibson

5/21/2021

Contents

1 Methods

A multistate model for Alzheimer's disease (AD) dementia using the dichotomous factors amyloidosis (A), tauopathy (T), neurodegeneration (N) and mild cognitive impairment (MCI) is shown in Figure 1. We refer to this model as the *ATN model*. The transitions $r_{ij}(a)$ are the one-year transition probabilities from state i to state j, conditional on surviving to age $a+1$. To calculate *lifetime risks* of AD dementia, we estimate the transition rates $r_{ij}(a)$ using available information.

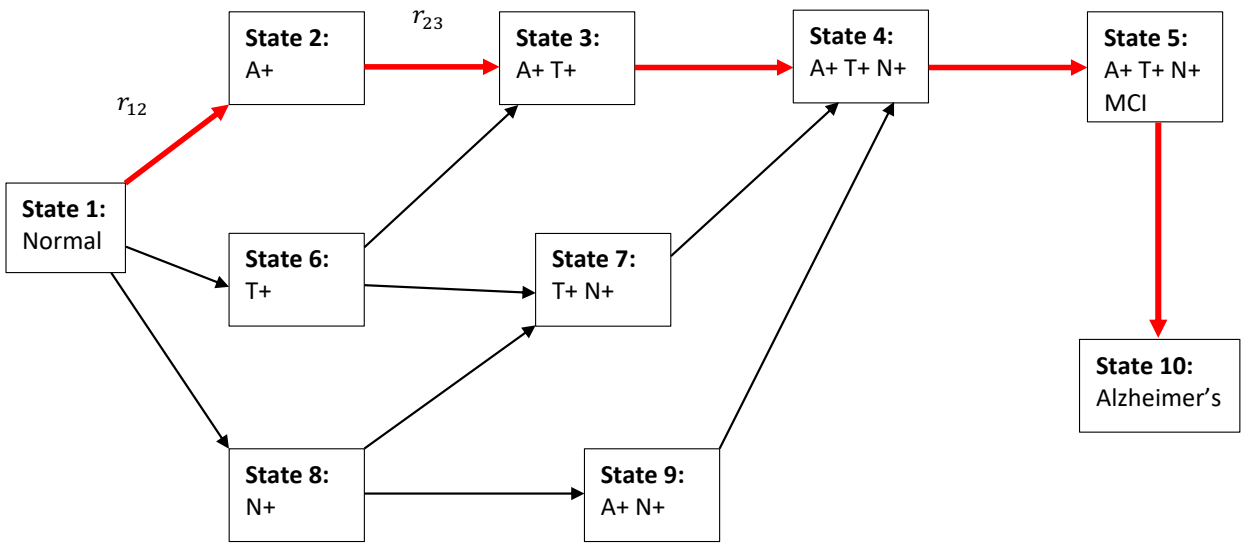


Figure 1: Multistate model for AD dementia with Amyloid (A+), Tauopathy (T+), Neurodegeneration (N+), and mild cognitive impairment (MCI)

1.1 Data and notation

We incorporate three sources of Alzheimer's disease (AD) dementia information into our analysis. These include age- and state-specific conditional prevalence rates $Pc_i(a)$ (Jack Jr et al. 2017), age-specific incidence rates $I(a)$ of AD dementia (Brookmeyer et al. 2007), and age-specific probabilities $C(a)$ of being in a preclinical state (Brookmeyer and Abdalla 2019). The prevalences $Pc_i(a)$ are conditional on being preclinical, i.e. no MCI has developed. The probabilities $C(a)$ were calculated using a multistate model for AD dementia using only the factors A, N, and MCI, which we refer to as the *AN model*.

We will use the following notation throughout:

- $r_{ij}(a)$ is the one-year transition probability of a person at age a going from state i to state j;
- $Pc_i(a)$ is to the prevalence of state i for individuals age $a, a = 50, \ldots, 95$, conditional on being preclinical and alive (Jack Jr et al. 2017),
- $Pu_i(a)$ is the prevalence of state i for individuals age $a, a = 50, \ldots, 95$, conditional only on being alive. These were calculated as $Pu_i(a) = Pc_i(a) * C(a)$, where $C(a)$ is the probability of being preclinical as calculated using the AN model (Brookmeyer and Abdalla 2019);
- $I(a)$ is the incidence of AD dementia for individuals age $a, a = 65, \ldots, 90$ (Brookmeyer et al. 2007); its functional form is $I(a) = .00117 \exp(0.126a)$.
- k_{0ij} and k_{1ij} are transition parameters from state i to state j such that $r_{ij}(a) = k_{0ij} \exp(k_{1ij}a)$

- $\boldsymbol{k}$ is the vector of *preclinical* transition rate parameters
- $\widehat{Pc}_i(\boldsymbol{k}, a)$, $\widehat{Pu}_i(\boldsymbol{k}, a)$, and $\widehat{I}(\boldsymbol{k}, a)$, are estimates of conditional prevalence rates, unconditional prevalence rates, and incidence rates of AD dementia from the multistate model for a given set of transition parameters $\boldsymbol{k}$

Jack Jr et al. (2017) provides estimates of *sex-specific* prevalence. The sex-specific functions $\widehat{Pu}_i(\boldsymbol{k}, a)$ from the multistate model output almost exactly the same values, so we take the average of the male and female prevalence estimates $Pu_i(a)$ and only use female prevalence and incidence functions in optimization.

1.2 Optimizing transition rate parameters

There are 14 total transitions r_{ij}, including 12 preclinical transitions $(r_{12}, \ldots, r_{94})$, r_{45}, and $r_{5,10}$. Each transition has log-linear form $r_{ij}(a) = k_{0ij} \exp(k_{1ij} a)$, and we constrain each pair (k_{0ij}, k_{1ij}) with

$$\begin{aligned} \log(k_{0ij}) &< 0 \\ k_{1ij} &> 0 \\ \log(k_{0ij}) + k_{1ij} \times 95 &< 0 \end{aligned} \tag{1}$$

such that every transition $r_{ij}(a) \in (0, 1)$ for all ages $a \in (50, \ldots, 95)$.

1.2.1 Transitions r_{45} $r_{5,10}$

We assume $r_{5,10}(a) = 0.3$ for all ages a, i.e. $k_{05,10} = 0.3$ and $k_{15,10} = 0$. To estimate the parameters k_{045} and k_{145} we assume that

$$r_{45}(a) = \lambda(a) \widetilde{r}_{45}(a), \tag{2}$$

where $\lambda(a) = \frac{P_4(a)}{P_4(a)+P_9(a)}$ is the proportion of those with both A+N+ that also have T+, and $\widetilde{r}_{45}(a)$ is the transition from State 4 (A+N+) to State 5 (A+N+ MCI) **in the AN model**. The ATN model assumes that only those with A+T+N+ can develop MCI, so we assume that those transitioning to MCI in the AN model were only coming from the subpopulation of A+N+ subjects that also had T+. Having calculated $r_{45}(a)$ using (2), we assume a log-linear form and perform a simple constrained optimization to find the parameters k_{045} and k_{145} using the base `constrOptim` function in R.

1.2.2 Preclinical Transitions

Here we use $\boldsymbol{k}$ to represent the vector of 24 parameters $(\log(k_0), k_1)_{(ij)}$ for preclinical transitions $(r_{12}, \ldots, r_{94})$. Optimization is with respect to the parameters $\boldsymbol{k}$. We use the COBYLA optimization algorithm (Powell 1994) in the R package `nloptr` (Johnson, n.d.), which accepts both linear and nonlinear constraints. In addition to the constraints (1) we impose constraints such that the sum of transitions out of state i for age a is less than 1 up to age 95

$$\log\big(\sum_{j \in R_i} \exp(\log(k_{0ij}) + k_{1ij} \times 95)\big) \leq 0 \quad \forall i \in I \tag{3}$$

where R_i is the set of states j one can transition to out of state i and $I = (1, 2, 3, 4, 6, 7, 8, 9)$ is the set of preclinical states.

The loss function f to be minimized has two components – one for prevalence and one for incidence

$$f(\boldsymbol{k}) = \frac{1}{n_1} \sum_{a=65}^{90} \Big(\log(\widehat{I}(\boldsymbol{k}, a)) - \log(I(a))\Big)^2 + \frac{1}{n_2} \sum_{i \in I} \sum_{a=50}^{95} \Big(\log(\widehat{Pu}_i(\boldsymbol{k}, a)) - \log(Pu_i(a))\Big)^2. \tag{4}$$

where $n_1 = 26$ and $n_2 = 8 \times 46$ are the number of elements contributing to the incidence portion and prevalence portion of the loss function, respectively, and I is the set of preclinical states. Because prevalence and incidence are on different scales, we use squared log-ratios instead of squared absolute differences. We weight each portion with the reciprocal of the number of contributing elements to give "equal" weight to the incidence and prevalence portions of the function.

2 Results

The optimization algorithm converged after 18881 iterations, with an optimal value of the loss function of 0.3441428. Optimal values of parameters $(\log(k_{0ij}), k_{1ij})$ are shown in Table 1.

Table 1: Optimal value of transition rate parameters

	log(k0)	k1
r12	-9.543083	0.0726038
r16	-8.394046	0.0716778
r18	-8.347971	0.0615139

	log(k0)	k1
r23	-9.137298	0.0868288
r29	-10.429581	0.0274339
r34	-4.045851	0.0425879
r63	-7.969481	0.0829260
r67	-7.167907	0.0498023
r74	-5.789055	0.0347527
r87	-8.358605	0.0335239
r89	-11.111297	0.1038721
r94	-7.995824	0.0665064

2.1 Lifetime Risks

Tables 2 and 3 show the lifetime risks of developing AD dementia for women and men, respectively, for ages $a \in (60, 65, \ldots, 90)$ and each pre-AD state $i \in (1, \ldots, 9)$.

Table 2: Lifetime risk of AD for women by age and disease state

Age	Normal	A	A+T	A+T+N	A+T+N + MCI	T	T+N	N	A+N
60	17.8 %	30.07 %	54.75 %	58.37 %	95.63 %	38.8 %	27.26 %	10.06 %	28.69 %
65	16.49 %	28.4 %	51.96 %	56.18 %	93.57 %	36.72 %	24.09 %	9.46 %	26.7 %
70	14.66 %	25.9 %	47.92 %	52.79 %	90.06 %	33.63 %	20.32 %	8.57 %	23.97 %
75	12.25 %	22.41 %	42.29 %	47.85 %	84.71 %	29.32 %	16.07 %	7.35 %	20.42 %
80	9.28 %	17.8 %	34.74 %	40.83 %	76.19 %	23.59 %	11.52 %	5.76 %	16.02 %
85	6.07 %	12.39 %	25.38 %	31.6 %	63.81 %	16.7 %	7.16 %	3.94 %	11.07 %
90	3.24 %	7.16 %	15.56 %	20.98 %	46.68 %	9.85 %	3.66 %	2.24 %	6.42 %

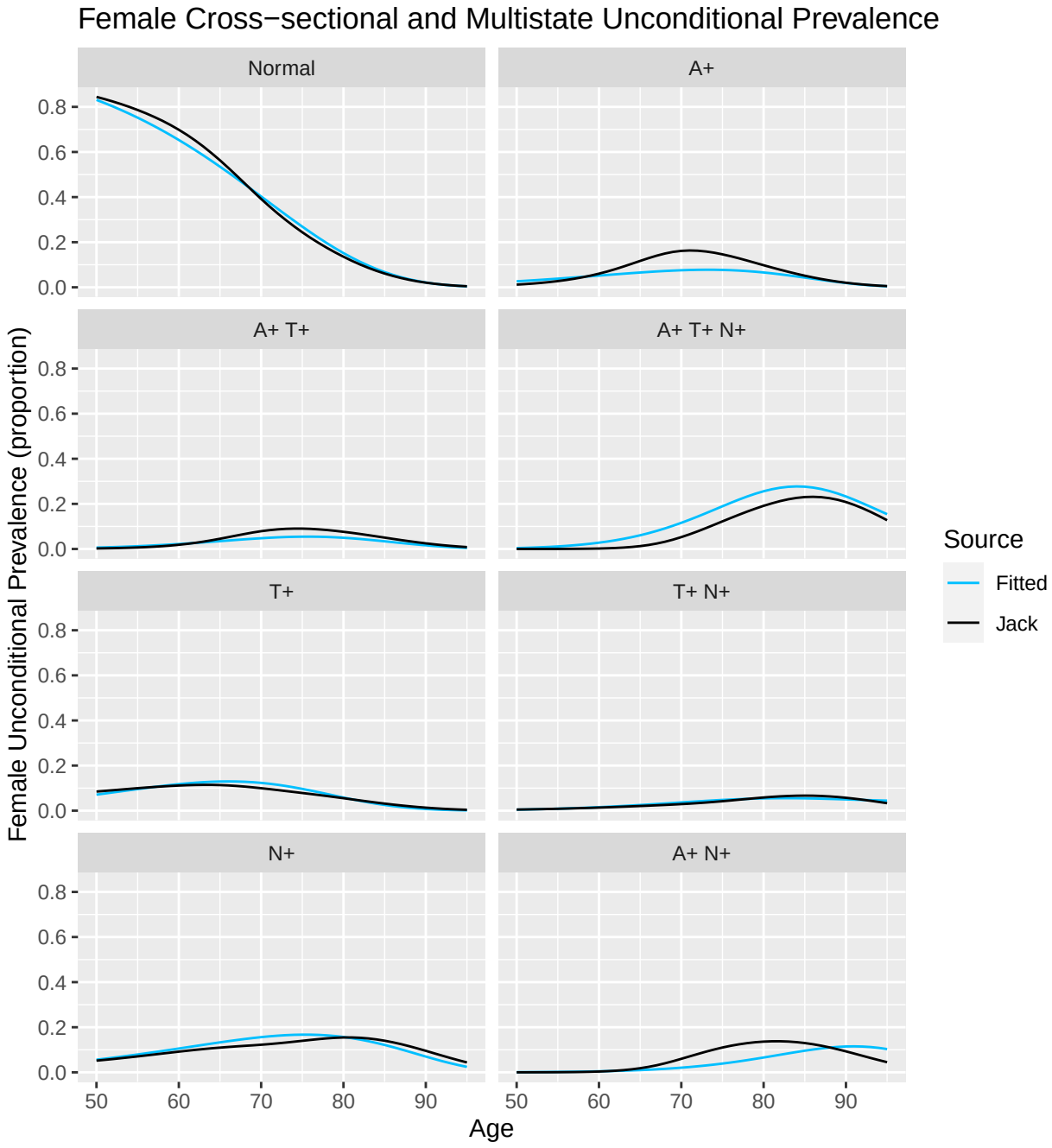
Table 3: Lifetime risk of AD for men by age and disease state

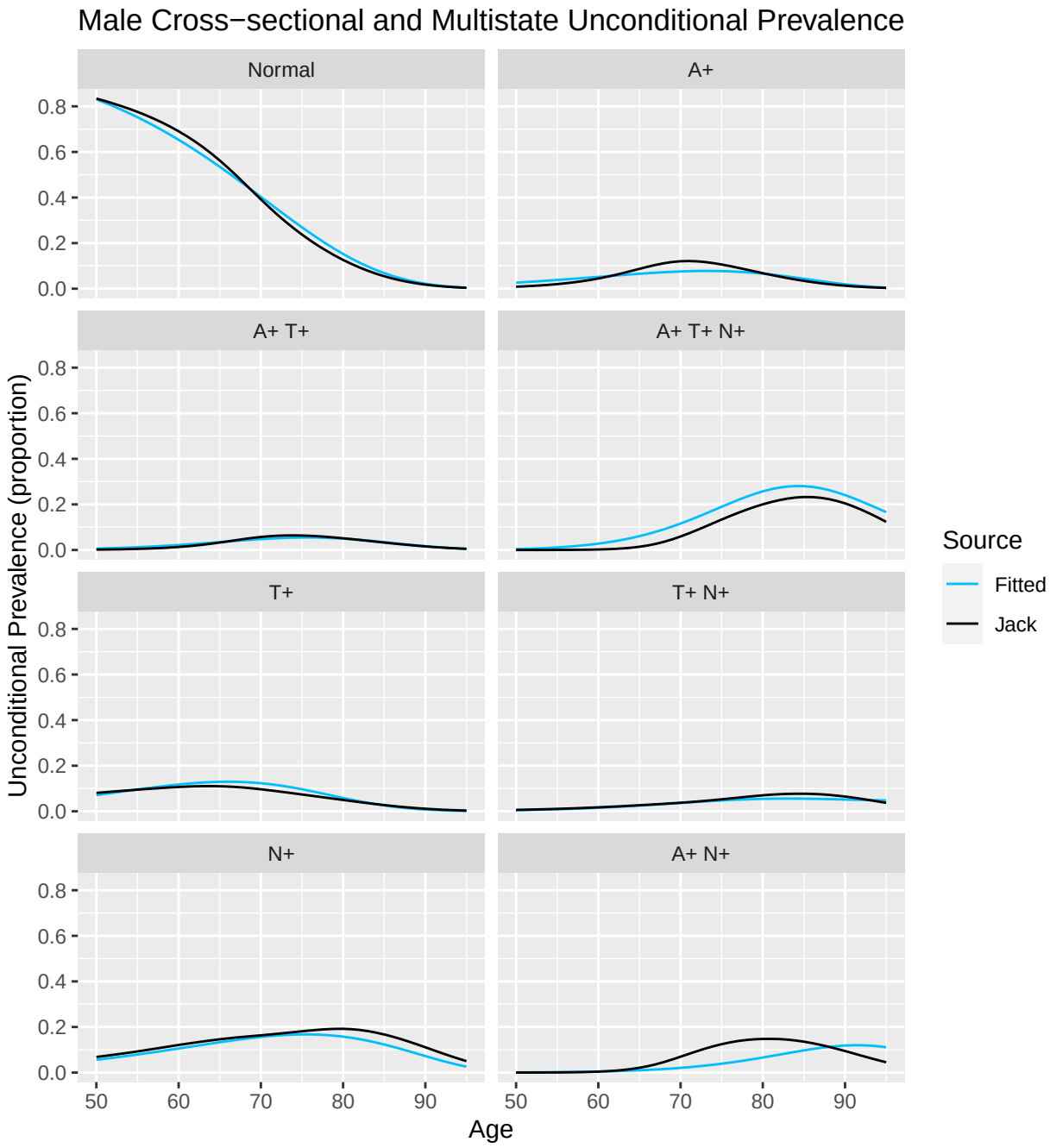
Age	Normal	A	A+T	A+T+N	A+T+N + MCI	T	T+N	N	A+N
60	12.09 %	21.94 %	45.32 %	49.39 %	92.93 %	29.87 %	20.65 %	6.42 %	21.35 %
65	11.17 %	20.71 %	42.84 %	47.49 %	90.37 %	28.19 %	18.1 %	6.05 %	19.83 %

Age	Normal	A	A+T	A+T+N	A+T+N + MCI	T	T+N	N	A+N
70	9.82 %	18.72 %	39.04 %	44.27 %	85.99 %	25.55 %	15.04 %	5.45 %	17.63 %
75	8.05 %	15.94 %	33.82 %	39.59 %	79.54 %	21.89 %	11.64 %	4.61 %	14.79 %
80	5.9 %	12.31 %	26.98 %	33.02 %	69.91 %	17.11 %	8.09 %	3.53 %	11.3 %
85	3.66 %	8.21 %	18.97 %	24.79 %	56.66 %	11.62 %	4.81 %	2.31 %	7.53 %
90	1.77 %	4.38 %	10.91 %	15.73 %	40.15 %	6.35 %	2.29 %	1.21 %	4.08 %

2.2 Unconditional prevalence

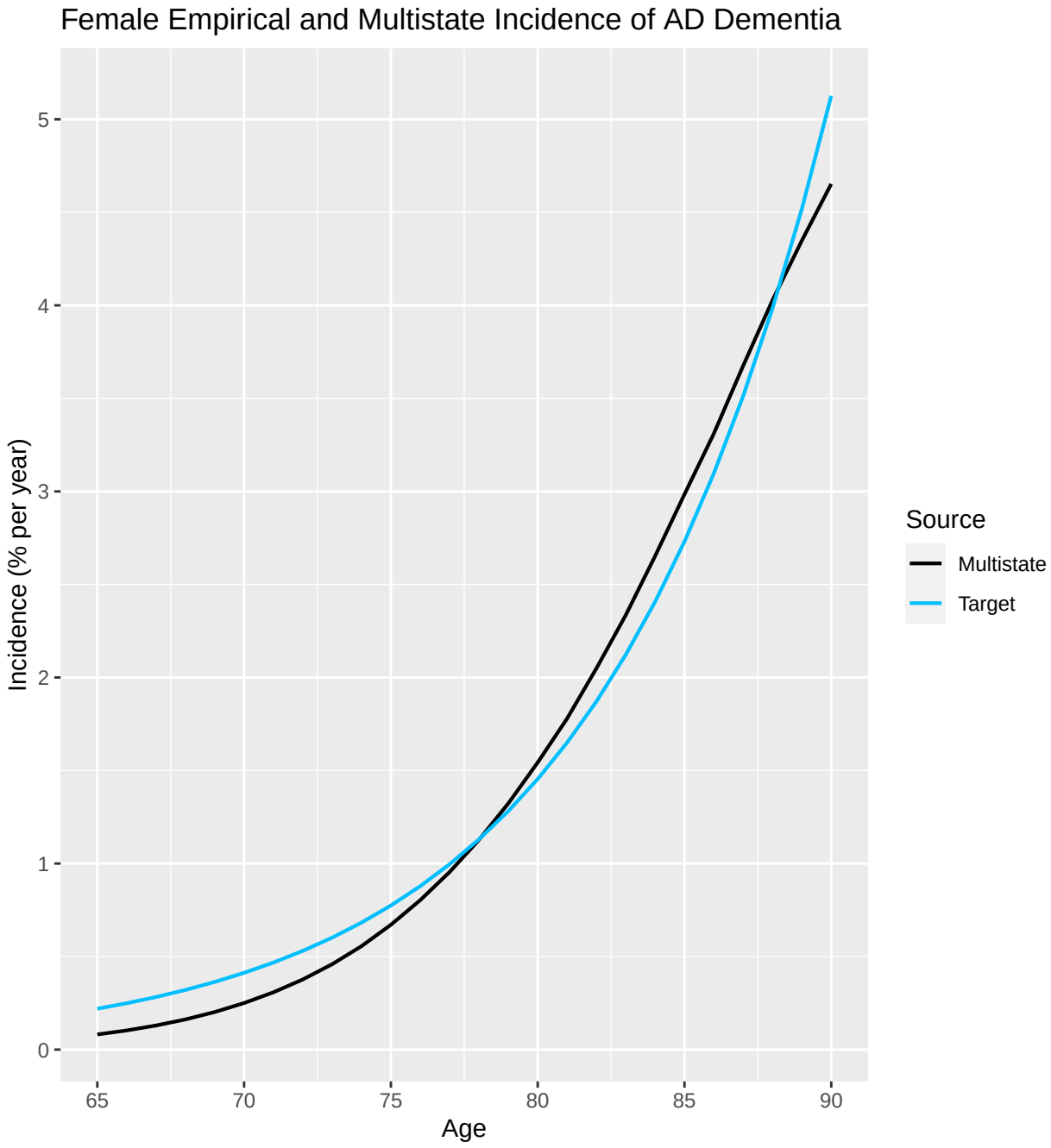
Figures @ref{fig:uncondf} and @ref{fig:uncondm} show the male and female prevalence estimates $\widehat{Pu}_i(a)$ as compared with $Pu_i(a)$. The plots mainly differ in $Pu_i(a)$, as the multistate functions for prevalence $\widehat{Pu}_i(a)$ are almost exactly equal.

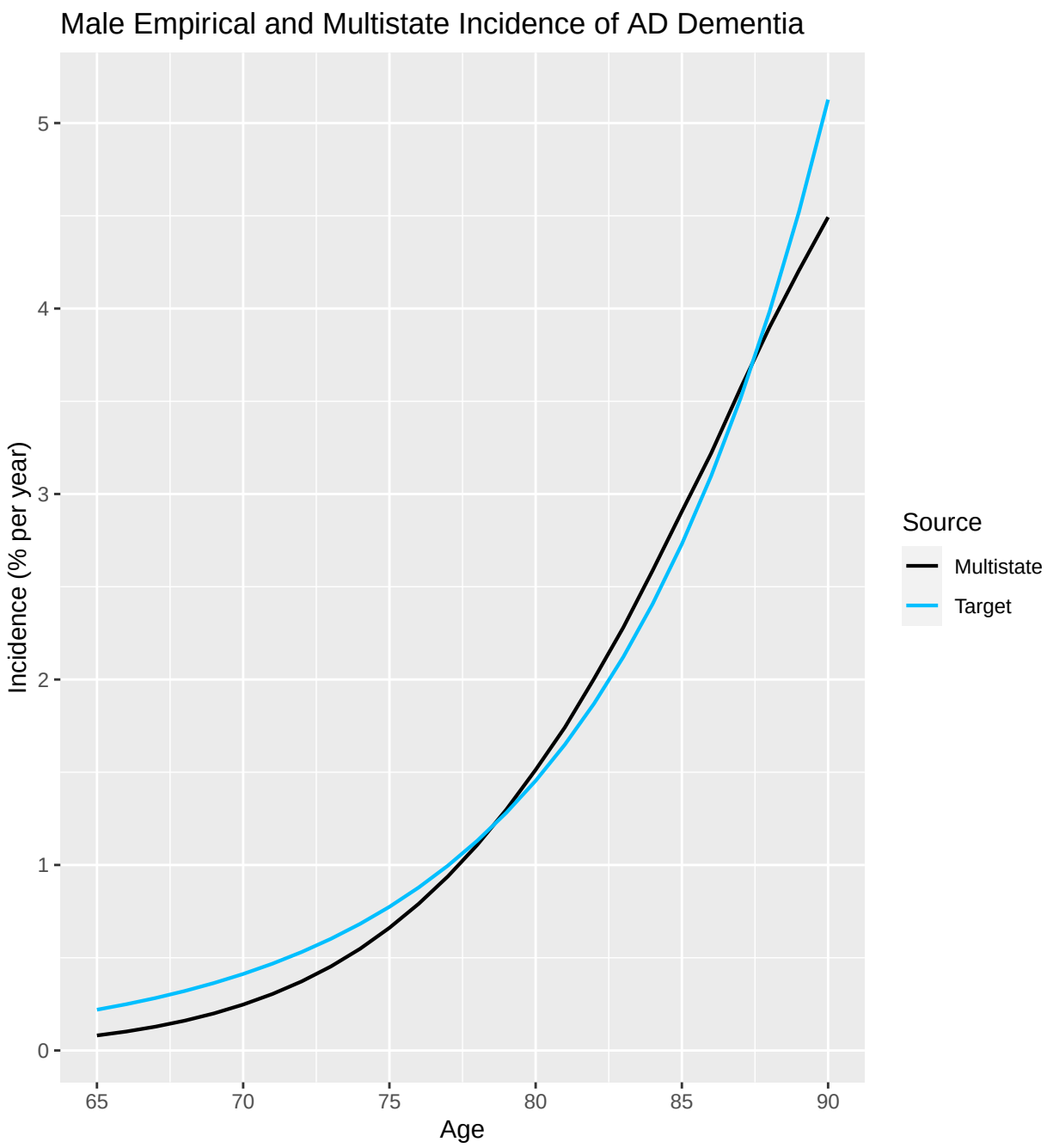



Male Cross−sectional and Multistate Unconditional Prevalence
Normal
A+
A+ T+
A+ T+ N+
T+
T+ N+
N+
A+ N+
Unconditional Prevalence (proportion)
0.0
0.2
0.4
0.6
0.8
50
60
70
80
90
Age
Source
Fitted
Jack

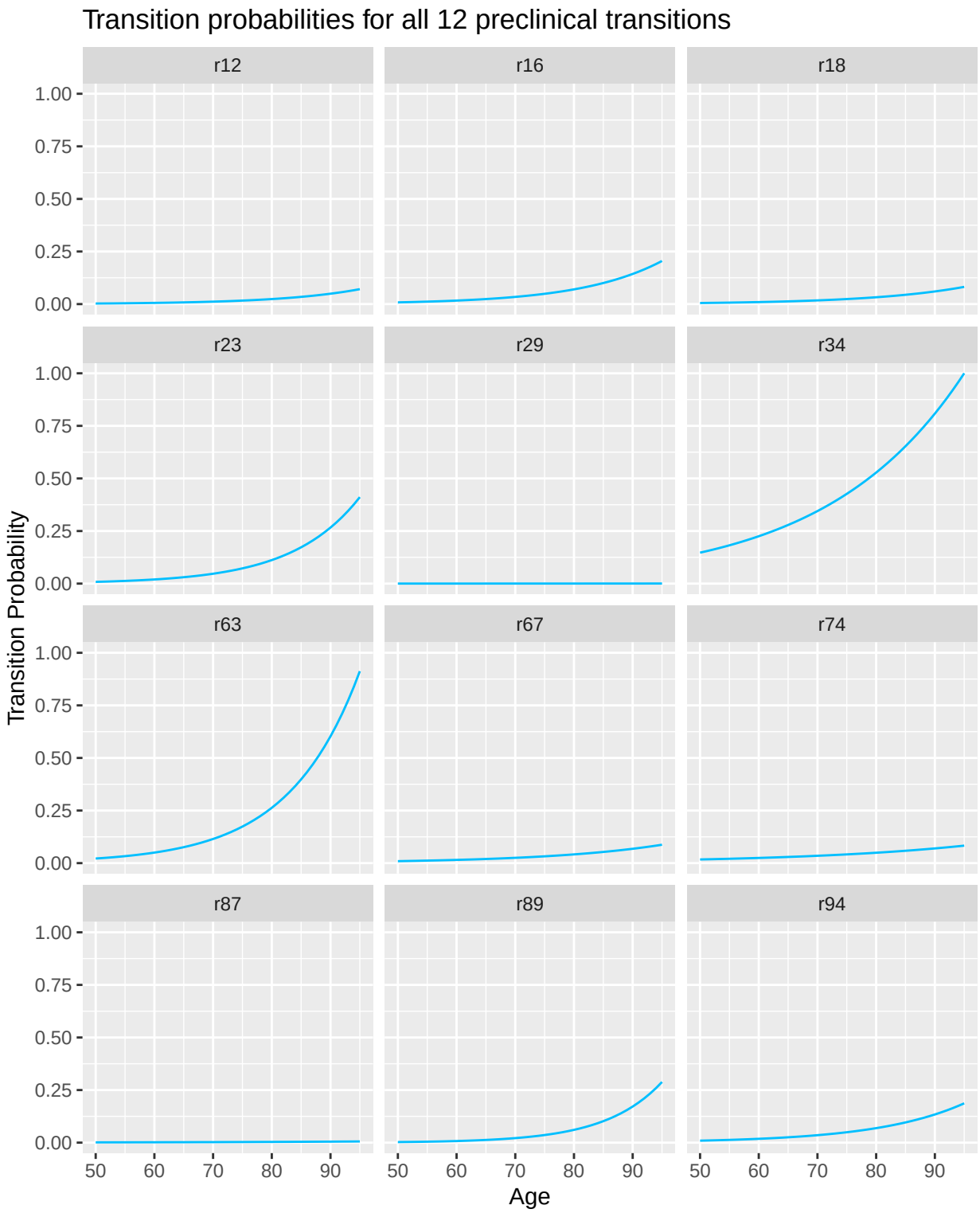
2.3 Incidence of AD

Female and male incidence rates as compared with the empirically estimated rate. The plots are nearly identical, as output from the multistate functions for incidence, $\widehat{I}(\boldsymbol{k}, a)$ are very close.




Male Empirical and Multistate Incidence of AD Dementia
Incidence (% per year)
0
1
2
3
4
5
65
70
75
80
85
90
Age
Source
Multistate
Target

2.4 Individual transitions $r_{ij}(a)$



References

Brookmeyer, Ron, and Nada Abdalla. 2019. "Multistate Models and Lifetime Risk Estimation: Application to Alzheimer's Disease." *Statistics in Medicine* 38 (9): 1558–65.

Brookmeyer, Ron, Elizabeth Johnson, Kathryn Ziegler-Graham, and H Michael Arrighi. 2007. "Forecasting the Global Burden of Alzheimer's Disease." *Alzheimer's & Dementia* 3 (3): 186–91.

Jack Jr, Clifford R, Heather J Wiste, Stephen D Weigand, Terry M Therneau, David S Knopman, Val Lowe, Prashanthi Vemuri, et al. 2017. "Age-Specific and Sex-Specific Prevalence of Cerebral β-Amyloidosis, Tauopathy, and Neurodegeneration in Cognitively Unimpaired Individuals Aged 50–95 Years: A Cross-Sectional Study." *The Lancet Neurology* 16 (6): 435–44.

Johnson, Steven G. n.d. "The NLopt Nonlinear-Optimization Package." http://github.com/stevengj/nlopt.

Powell, Michael JD. 1994. "A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation." In *Advances in Optimization and Numerical Analysis*, 51–67. Springer.