

Estimation of Genetic Risk Function with Covariates in the Presence of Missing Genotypes

Annie J. Lee^a, Karen Marder^{bc}, Roy N. Alcalay^{bc}, Helen Mejia-Santana^b, Avi Orr-Urtreger^{de}, Nir Giladi^{fg}, Susan Bressman^h, Yuanjia Wang^{a*}

Appendix A. Additional simulation results

Based on the simulation results from 1,000 replications with 1,000 bootstrap samples in Section 3 and using model (5), we estimated the marginal cumulative risk functions in carriers and non-carriers in Figure S1 and the performance of the estimates at various ages with censoring rate of 40% or 60% in Table S1. The penetrance estimates in male and female relatives were examined separately along with their 95% confidence intervals with the censoring rate of 40% and 60% in Figure 1 and Figure S2, respectively. Moreover, Figure S3 and S4 presents the penetrance estimates in male and female relatives of male and female probands and Table S2 shows the performance of the estimates at various ages with the censoring rate of 40% or 60%.

In Table 2, S1, and S2, we observed the small bias of estimated penetrance over the entire range of age and the estimated standard errors agrees adequately with the empirical standard deviations, and the coverage probabilities were close to the nominal level.

^a Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY, U.S.A.

^b Department of Neurology, College of Physicians and Surgeons, Columbia University, New York, NY, U.S.A.

^c Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY, U.S.A.

^d Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.

^e Genetic Institute, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel.

^f Sackler Faculty of Medicine, Sagol School for Neurosciences, Tel Aviv University, Tel Aviv, Israel.

^g Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel.

^h Department of Neurology, Mount Sinai Beth Israel Medical Center, New York, NY, USA.

* Correspondence to: Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, U.S.A. E-mail: yw2016@columbia.edu

Table S1: Summary results for the marginal penetrance estimates marginalized by relative's sex and proband's sex in the simulation ($\times 10^{-2}$) with censoring rate of 40% or 60%.

Censor	Age	Carrier $\hat{F}_1(\cdot)$					Non-Carrier $\hat{F}_0(\cdot)$				
		True Risk	Bias	SD	SE	CP	True Risk	Bias	SD	SE	CP
40%	60	19.00	0.16	1.69	1.71	94.4	7.97	<0.01	0.56	0.56	95.3
	65	26.91	0.21	2.13	2.13	94.0	11.61	<0.01	0.70	0.69	94.6
	70	36.38	0.24	2.56	2.56	94.4	16.29	<0.01	0.83	0.82	95.1
	75	47.02	0.24	2.91	2.90	94.5	22.06	<0.01	0.97	0.95	94.3
	80	58.16	0.25	3.19	3.11	94.1	28.88	0.04	1.04	1.06	96.1
60%	60	19.00	-0.43	1.81	1.87	94.8	7.97	0.10	0.59	0.58	94.4
	65	26.91	-0.60	2.32	2.38	94.4	11.61	0.14	0.72	0.72	95.6
	70	36.38	-0.73	2.80	2.90	95.3	16.29	0.21	0.89	0.86	94.2
	75	47.02	-1.05	3.20	3.37	95.7	22.06	0.17	1.01	1.01	94.5
	80	58.16	-1.55	3.56	3.70	94.2	28.88	-0.03	1.19	1.16	93.9

Denote the true cumulative risk (True Risk), average estimation bias over 1,000 replications (Bias), empirical standard deviation (SD), average of estimated standard errors from bootstraps (SE), and coverage probability corresponding to nominal 95% confidence intervals (CP). Results are based on 1,000 simulations with sample size $n = 2266$.

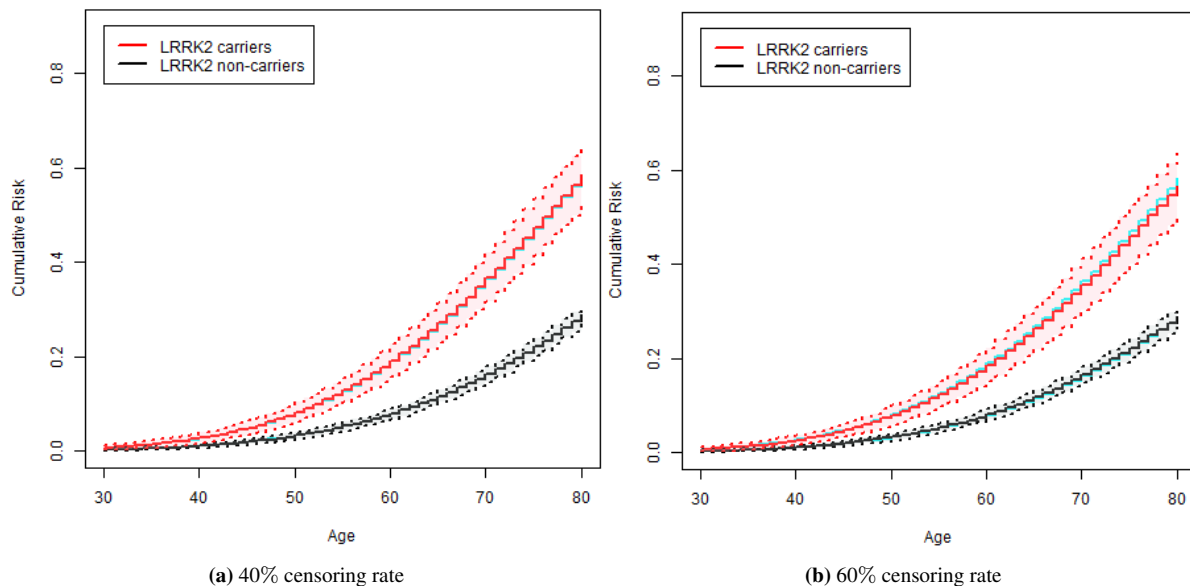


Figure S1: Estimated marginal cumulative risk functions in the simulation with censoring rate of 40% or 60%: Carriers (red solid line) and non-carriers (black solid line) with their 95% confidence intervals (dashed lines) and true cumulative risk functions (blue solid line).

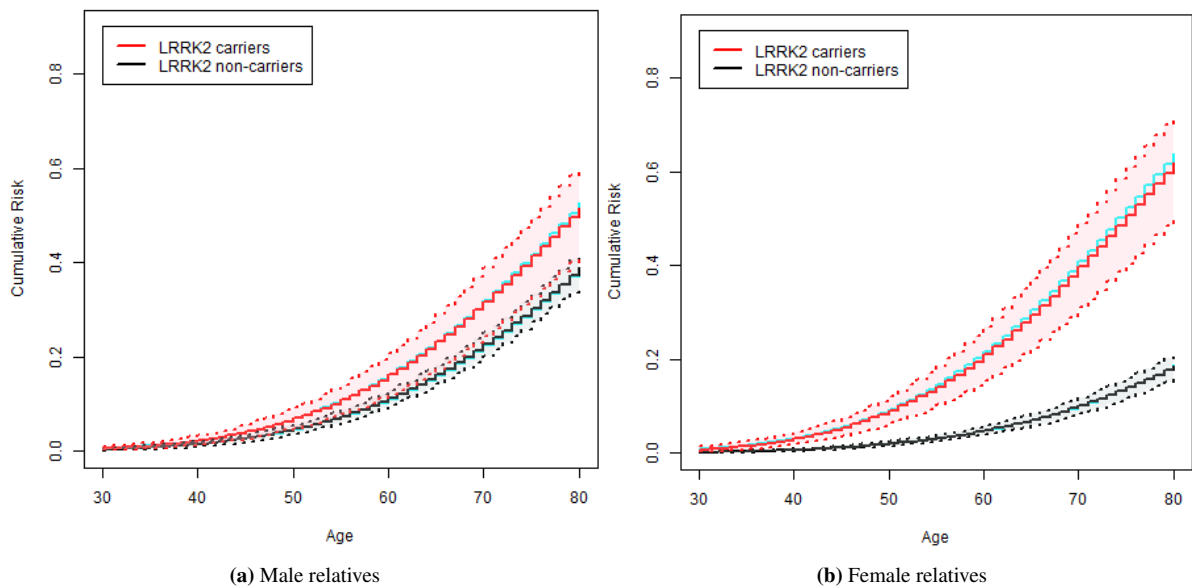


Figure S2: Estimated cumulative risk functions in the simulation with 60% censoring rate: Carriers (red solid line) and non-carriers (black solid line) in male and female relatives marginalized by probands' sex with their 95% confidence intervals (dashed lines) and true cumulative risk functions (blue solid line).

Table S2: Summary results for the penetrance estimates in male and female relatives of male and female probands in the simulation ($\times 10^{-2}$) with censoring rate of 40% or 60%.

Censor	Rel sex	Prob sex	Age	Carrier $\hat{F}_1(\cdot)$					Non-Carrier $\hat{F}_0(\cdot)$				
				True Risk	Bias	SD	SE	CP	True Risk	Bias	SD	SE	CP
40%	Male	Male	60	14.20	0.02	1.77	1.77	94.8	9.56	0.06	0.76	0.75	94.8
			65	20.43	0.04	2.38	2.35	94.5	13.92	0.10	0.96	0.95	94.1
			70	28.18	0.04	3.03	2.99	93.7	19.52	0.14	1.19	1.16	94.9
			75	37.33	0.03	3.68	3.62	94.0	26.41	0.18	1.44	1.39	93.6
			80	47.55	0.06	4.25	4.14	93.5	34.52	0.27	1.62	1.60	94.9
		Female	60	19.49	-0.04	2.36	2.36	94.3	13.26	0.04	1.01	1.03	94.6
			65	27.63	-0.04	3.06	3.05	94.4	19.12	0.07	1.26	1.27	95.3
			70	37.40	-0.08	3.73	3.73	94.7	26.47	0.10	1.49	1.53	95.7
			75	48.39	-0.13	4.30	4.27	94.3	35.21	0.12	1.75	1.78	95.9
			80	59.88	-0.13	4.63	4.55	93.7	45.08	0.20	1.90	1.97	96.1
	Female	Male	60	18.94	0.32	2.34	2.32	94.2	4.12	-0.05	0.39	0.39	93.9
			65	26.90	0.44	3.06	3.01	94.1	6.08	-0.06	0.53	0.52	94.2
			70	36.48	0.53	3.80	3.71	93.9	8.69	-0.09	0.70	0.68	92.7
			75	47.31	0.58	4.40	4.30	94.5	12.04	-0.12	0.90	0.87	93.5
			80	58.72	0.63	4.82	4.65	93.6	16.24	-0.13	1.10	1.09	94.1
		Female	60	25.71	0.33	3.07	3.03	93.9	5.78	-0.08	0.53	0.52	94.0
			65	35.82	0.42	3.84	3.78	93.9	8.50	-0.11	0.71	0.69	93.1
			70	47.39	0.46	4.48	4.40	94.2	12.07	-0.16	0.92	0.90	93.5
			75	59.62	0.42	4.81	4.73	94.3	16.61	-0.21	1.17	1.14	93.5
			80	71.42	0.38	4.78	4.63	94.1	22.18	-0.24	1.41	1.40	93.8
60%	Male	Male	60	14.20	-0.08	2.02	1.97	93.7	9.56	0.17	0.81	0.81	94.7
			65	20.43	-0.14	2.71	2.64	93.5	13.92	0.24	1.04	1.03	94.9
			70	28.18	-0.17	3.46	3.39	94.3	19.52	0.35	1.32	1.29	93.7
			75	37.33	-0.39	4.16	4.13	95.0	26.41	0.33	1.57	1.56	94.2
			80	47.55	-0.83	4.81	4.77	94.3	34.52	0.13	1.88	1.84	94.7
		Female	60	19.49	-0.22	2.61	2.64	94.3	13.26	0.17	1.08	1.11	95.1
			65	27.63	-0.33	3.41	3.43	94.5	19.12	0.22	1.37	1.39	95.2
			70	37.40	-0.40	4.18	4.23	94.8	26.47	0.33	1.71	1.70	95.1
			75	48.39	-0.69	4.77	4.90	95.2	35.21	0.26	1.98	2.02	95.1
			80	59.88	-1.19	5.17	5.30	95.1	45.08	-0.03	2.33	2.32	94.6
	Female	Male	60	18.94	-0.60	2.56	2.60	93.8	4.12	0.04	0.45	0.45	95.0
			65	26.90	-0.83	3.38	3.42	94.3	6.08	0.05	0.61	0.61	95.1
			70	36.48	-1.04	4.22	4.28	94.3	8.69	0.08	0.82	0.81	95.5
			75	47.31	-1.44	4.96	5.03	94.0	12.04	0.05	1.04	1.05	95.7
			80	58.72	-2.01	5.50	5.55	94.0	16.24	-0.08	1.30	1.31	96.1
		Female	60	25.71	-0.90	3.32	3.42	94.3	5.78	0.02	0.60	0.60	94.2
			65	35.82	-1.22	4.22	4.32	93.9	8.50	0.03	0.82	0.81	93.8
			70	47.39	-1.46	5.00	5.12	94.0	12.07	0.05	1.08	1.07	94.1
			75	59.62	-1.86	5.49	5.62	93.8	16.61	-0.01	1.37	1.37	94.5
			80	71.42	-2.32	5.58	5.68	93.8	22.18	-0.21	1.70	1.69	94.1

Denote the true cumulative risk (True Risk), average estimation bias over 1,000 replications (Bias), empirical standard deviation (SD), average of estimated standard errors from bootstraps (SE), and coverage probability corresponding to nominal 95% confidence intervals (CP). Results are based on 1,000 simulations with sample size $n = 2266$.

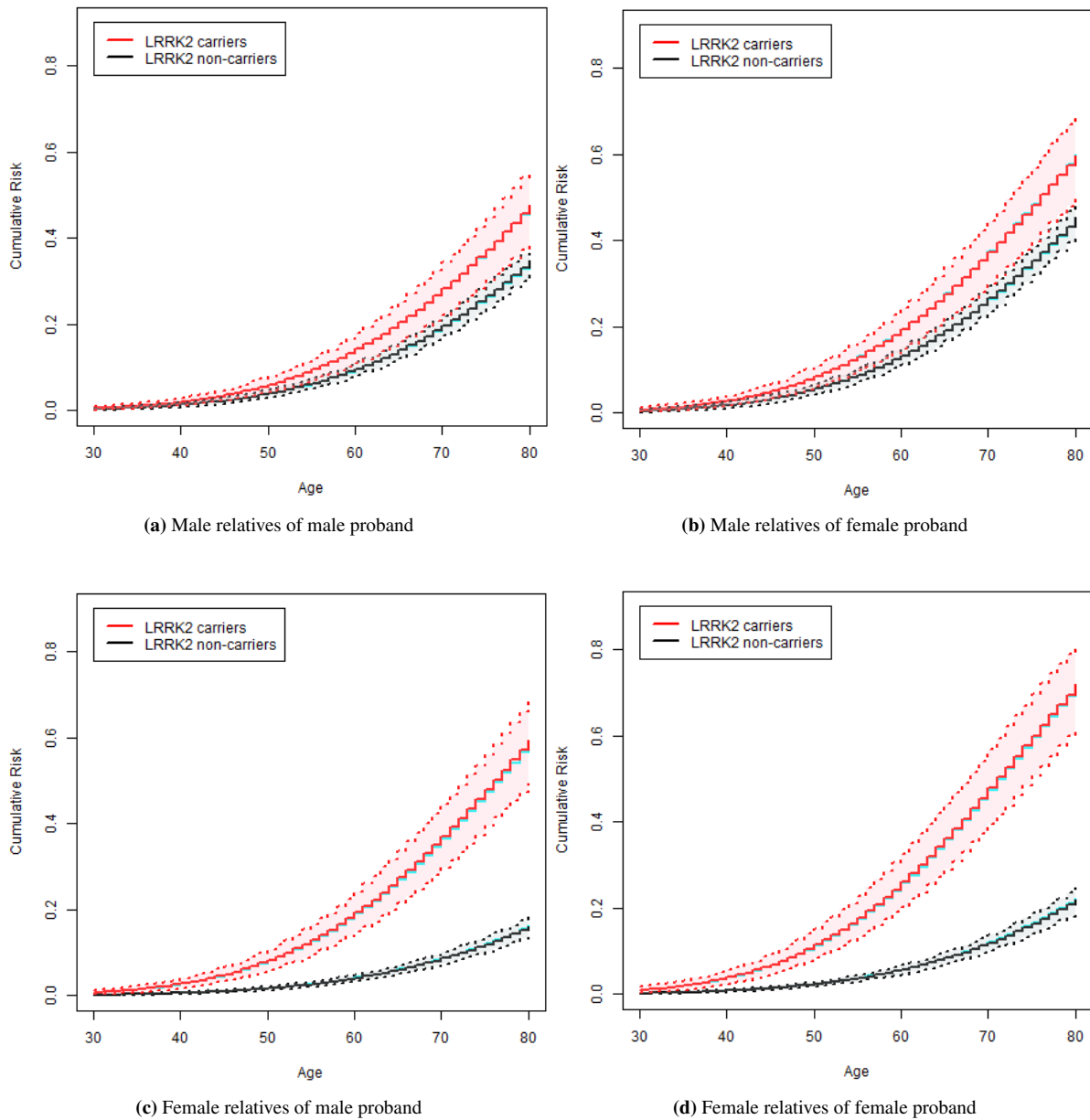


Figure S3: Estimated cumulative risk functions in the simulation with 40% censoring rate: Carriers (red solid line) and non-carriers (black solid line) in male and female relatives of male and female probands with their 95% confidence intervals (dashed lines) and true cumulative risk functions (blue solid line).

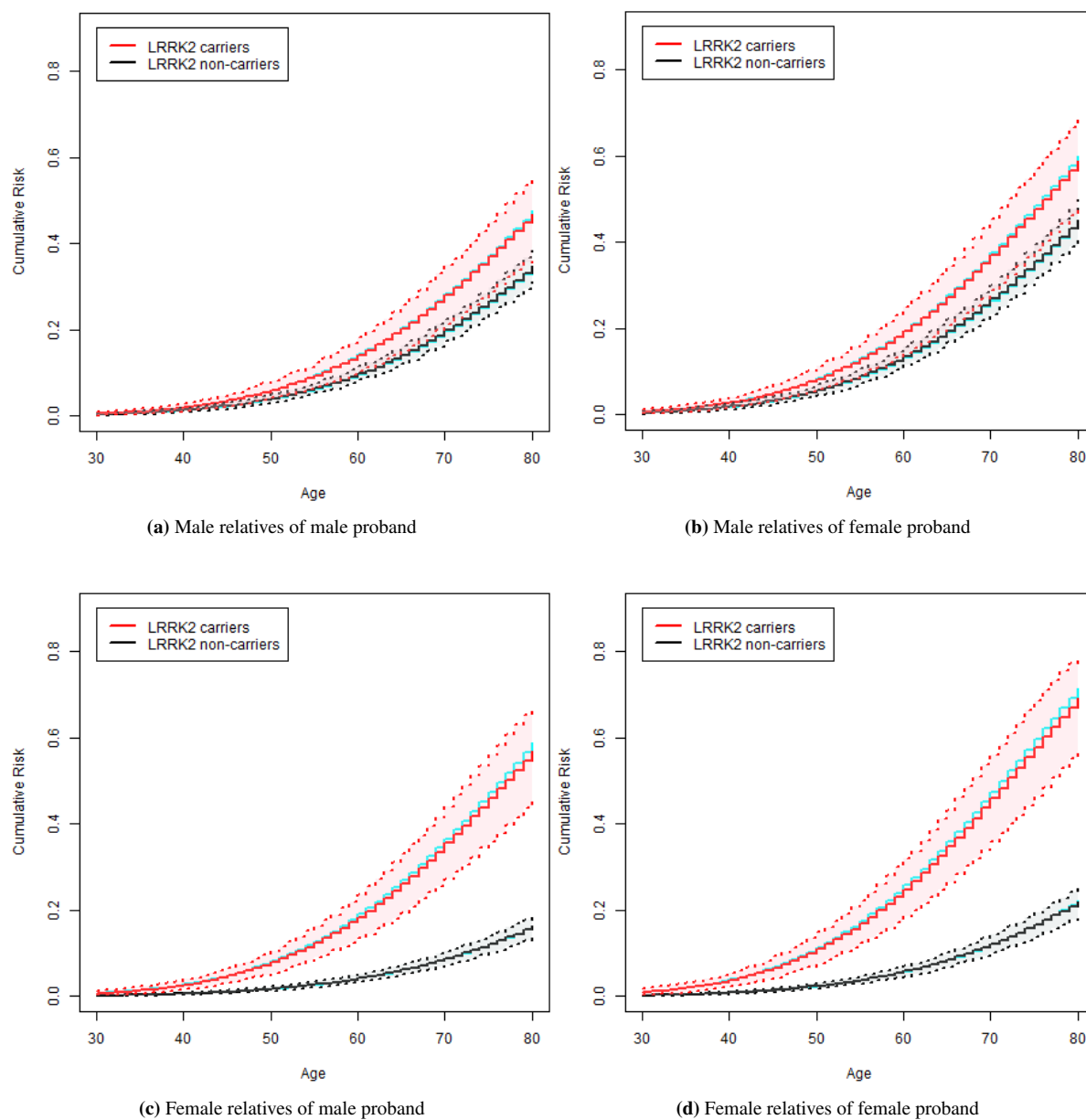


Figure S4: Estimated cumulative risk functions in the simulation with 60% censoring rate: Carriers (red solid line) and non-carriers (black solid line) in male and female relatives of male and female probands with their 95% confidence intervals (dashed lines) and true cumulative risk functions (blue solid line).

Appendix B. Additional data analyses results

Based on the parameter estimates that we obtained in the Ashkenazi Jewish *LRRK2* Consortium study in Section 4, the marginal cumulative risk functions in *LRRK2* G2019S carriers and non-carriers can be estimated using model (5) and the results are in Table S3 and Figure S5. We adjusted for relative's sex and carrier status interaction, proband's sex, and site of enrollment to provide precise risk prediction and we compared the penetrance estimates to the one in Marder et al. (2015) where none of the covariates were controlled.

Table S3: Estimated marginal cumulative risk of Parkinson's disease onset in *LRRK2* carriers and non-carriers in the Ashkenazi Jewish *LRRK2* Consortium study ($\times 10^{-2}$).

	Age	Carrier $\hat{F}_1(\cdot)$			Non-Carrier $\hat{F}_0(\cdot)$		
		Risk	Lower limit	Upper limit	Risk	Lower limit	Upper limit
Adjusted (Proposed)	60	7.27	4.18	11.16	3.01	1.94	4.45
	65	11.44	6.86	16.50	4.80	3.28	6.83
	70	16.74	10.84	23.46	7.15	4.93	9.88
	75	20.81	13.27	28.86	9.01	6.68	11.99
	80	24.75	16.26	34.26	10.87	8.07	14.46
Unadjusted (Marder et al., 2015)	60	7.84	4.61	12.27	2.79	1.80	3.99
	65	12.28	7.59	17.89	4.44	3.11	6.08
	70	17.89	12.05	24.83	6.60	4.70	8.95
	75	22.16	15.03	30.39	8.31	6.13	10.99
	80	26.22	17.94	36.30	10.00	7.31	13.37

The penetrance estimates adjusted for relative's sex and carrier status interaction, proband's sex, and site of enrollment (Proposed) was compared to the penetrance estimates unadjusted for any covariates reported in the previous study (Marder et al., 2015). Denote the estimated age-specific risk (Risk) and 95% confidence intervals (Lower limit and Upper limit).

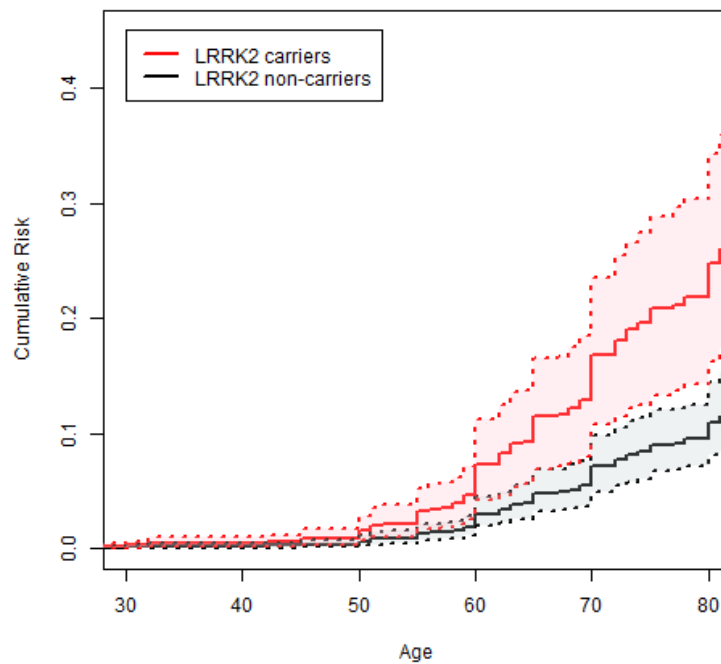


Figure S5: Estimated marginal age-specific risk of Parkinson's disease (PD) in *LRRK2* G2019S carriers (red solid line) and non-carriers (black solid line) with their 95% confidence intervals (dashed lines).

Table S4: BIC for analyses assuming time-invariant genotype effect β (Cox PH model) and time-varying genotype effect $\beta(t)$ estimated by B-splines with various number of knots and degrees in the Ashkenazi Jewish *LRRK2* Consortium study.

Analysis*	Cox PH model	Time-varying genotype effect				
		degree	Number of knots			
			0	1	2	3
Scenario 1	91.6	Linear	105.9	111.5	119.6	125.3
		Quadratic	108.8	120.4	126.8	135.8
		Cubic	120.4	129.0	137.6	144.3
Scenario 2	97.0	Linear	111.6	116.8	125.1	130.5
		Quadratic	114.0	126.0	127.6	139.8
		Cubic	126.0	134.8	143.3	151.5
Scenario 3	114.6	Linear	129.4	134.7	143.0	148.3
		Quadratic	131.9	143.9	145.5	157.5
		Cubic	143.8	152.7	161.2	169.2

*: Scenario 1 adjusted for relative's sex and carrier status interaction. Scenario 2 adjusted for relative's sex and carrier status interaction and proband's sex. Scenario 3 adjusted for relative's sex and carrier status interaction, proband's sex, and site of enrollment.