

eAppendix

eTable 1: Mean hazard ratios and empirical type I error rates ( $\alpha = 0.05$ ) for cohorts simulated under a **constant** exposure intensity for **complete** follow-up (through death or loss to follow-up), by hire year/age method, hazard ratio, and lag period <sup>a</sup>

Hire year/age method	Mean hazard ratio							Empirical type I error rate ( $\alpha = 0.05$ ) <sup>b</sup>				
	Hazard ratio	Lag (years)	Includes lagged-out	Restricted-			Restricted-					
				Unadjusted	Category-of- birth adjusted	Year-of-birth adjusted	cubic-spline adjusted	Unadjusted	Category-of- birth adjusted	Year-of-birth adjusted	cubic-spline adjusted	
1:Independent	1	0	---	1.01	1.00	1.01	1.00	0.050	0.054	0.052	0.050	
	2	0	---	2.00	2.00	2.00	2.01	0.042	0.042	0.042	0.044	
	2	10	Yes	2.01	2.00	2.01	2.00	0.050	0.054	0.052	0.052	
			No	2.01	2.00	2.01	2.01	0.048	0.054	0.048	0.052	
	2	20	Yes	2.02	2.01	2.02	2.01	0.044	0.038	0.044	0.036	
			No	2.02	2.01	2.02	2.01	0.042	0.046	0.040	0.042	
2:Negative	1	0	---	1.01	1.00	1.00	1.00	0.048	0.042	0.042	0.044	
	2	0	---	2.02	2.01	2.01	2.01	0.052	0.048	0.044	0.048	
	2	10	Yes	2.03	2.01	2.02	2.00	0.054	0.032	0.040	0.030	
			No	2.02	2.01	2.01	2.01	0.044	0.034	0.042	0.032	
	2	20	Yes	<b>2.06</b>	2.01	<b>2.03</b>	2.00	<b>0.086</b> <sup>c</sup>	0.066	<b>0.074</b> <sup>c</sup>	0.066	
			No	2.03	2.01	2.02	2.01	0.068	0.068	0.068	0.064	

3:Reading	1	0	---	<b>1.02</b>	1.01	1.01	1.01	<b>0.074</b> <sup>c</sup>	0.064	0.066	0.064
	2	0	---	2.01	2.00	2.01	2.00	0.048	0.046	0.046	0.052
	2	10	Yes	<b>2.06</b>	2.02	2.03	2.02	<b>0.114</b> <sup>c</sup>	0.058	0.070	0.050
			No	<b>2.05</b>	2.02	2.03	2.02	<b>0.080</b> <sup>c</sup>	0.058	0.054	0.052
	2	20	Yes	<b>2.10</b>	2.01	2.03	2.00	<b>0.150</b> <sup>c</sup>	0.036	0.050	0.028
			No	2.05	2.02	2.02	2.01	0.054	0.036	0.044	0.034

---

<sup>a</sup> Five hundred cohorts were generated for each combination of hire year/age method and risk scenario. Each cohort contained 4000 (hire year/age methods 1 and 2) or 3569 (hire year/age method 3) workers. Exposure intensity was assigned to each worker-year of exposure using a log-normal distribution with a geometric mean of one unit and a geometric standard deviation of 2.2.

<sup>b</sup> The empirical type I error rate ( $\alpha = 0.05$ ) is the fraction of cohorts for which the 95% confidence interval for the log hazard ratio does not contain the true value.

<sup>c</sup> Significantly elevated (based on the z-test for proportions) empirical type I error rates (i.e., Type I error rate greater than 5% with a two-sided p-value  $< 0.05$ ) are indicated in bold-face type. Actual p-values for flagged results are 0.0041, 0.040, 0.040,  $<0.0001$ , 0.013, and  $<0.0001$  for type I error rates of 0.086, 0.074, 0.074, 0.114, 0.080, and 0.150, respectively.

eTable 2: Mean hazard ratios and empirical type I error rates ( $\alpha = 0.05$ ) for cohorts simulated under an **increasing** exposure intensity for **complete** follow-up (through death or loss to follow-up), by hire year/age method, hazard ratio, and lag period <sup>a</sup>

Hire year/age method	Mean hazard ratio							Empirical type I error rate ( $\alpha = 0.05$ ) <sup>b</sup>			
	Hazard ratio	Lag (years)	Includes lagged-out	Restricted-			Restricted-				
				Unadjusted	Category-of- birth adjusted	Year-of-birth adjusted	cubic-spline adjusted	Unadjusted	Category-of- birth adjusted	Year-of-birth adjusted	cubic-spline adjusted
1:Independent	1	0	---	1.00	0.99	0.99	1.00	0.052	0.052	0.048	0.044
	2	0	---	2.00	1.99	1.99	2.00	0.048	0.052	0.048	0.048
	2	10	Yes	2.00	1.99	1.99	2.01	0.046	0.052	0.058	0.056
			No	2.00	1.99	1.99	2.01	0.050	0.052	0.058	0.048
	2	20	Yes	2.00	1.98	1.99	2.00	0.046	0.050	0.040	0.032
			No	1.99	1.99	1.99	2.00	0.042	0.048	0.040	0.036
2:Negative	1	0	---	1.01	0.99	0.99	1.00	0.058	0.042	0.050	0.038
	2	0	---	2.02	1.99	1.98	2.01	0.044	0.042	0.048	0.046
	2	10	Yes	2.02	1.99	1.98	2.00	0.068	0.050	0.052	0.048
			No	2.01	1.99	1.98	2.01	0.054	0.050	0.054	0.046
	2	20	Yes	2.03	1.99	1.99	2.01	0.058	0.038	0.042	0.046
			No	2.01	2.00	2.00	2.02	0.046	0.040	0.040	0.054

3:Reading	1	0	---	<b>1.02</b>	1.00	0.98	1.00	<b>0.088</b> <sup>c</sup>	0.062	0.064	0.066
	2	0	---	<b>2.05</b>	2.01	1.98	2.01	<b>0.080</b> <sup>c</sup>	0.042	0.048	0.038
	2	10	Yes	<b>2.08</b>	2.02	1.99	2.02	<b>0.128</b> <sup>c</sup>	0.050	0.056	0.050
			No	<b>2.07</b>	2.02	1.99	2.02	<b>0.096</b> <sup>c</sup>	0.050	0.056	0.052
	2	20	Yes	<b>2.09</b>	2.00	1.98	2.00	<b>0.126</b> <sup>c</sup>	0.050	0.046	0.044
			No	<b>2.05</b>	2.00	1.98	2.00	<b>0.074</b> <sup>c</sup>	0.052	0.044	0.048

---

<sup>a</sup> Five hundred cohorts were generated for each combination of hire year/age method and risk scenario. Each cohort contained 4000 (hire year/age methods 1 and 2) or 3569 (hire year/age method 3) workers. Exposure intensity was assigned to each worker-year of exposure using a log-normal distribution a geometric standard deviation of 2.2 and a geometric mean that increased 2.5% each year.

<sup>b</sup> The empirical type I error rate ( $\alpha = 0.05$ ) is the fraction of cohorts for which the 95% confidence interval for the log hazard ratio does not contain the true value.

<sup>c</sup> Significantly elevated (based on the z-test for proportions) empirical type I error rates (i.e., Type I error rate greater than 5% with a two-sided p-value  $< 0.05$ ) are indicated in bold-face type. Actual p-values for flagged results are 0.0027, 0.013,  $<0.0001$ , 0.0005,  $<0.0001$ , and 0.040 for type I error rates of 0.088, 0.080, 0.128, 0.096, 0.126, and 0.074, respectively.

eTable 3: Mean hazard ratios and empirical type I error rates ( $\alpha = 0.05$ ) for cohorts simulated under a **constant** exposure intensity for **incomplete** follow-up (through the earlier of death, loss to follow-up or administrative censoring in 1990), by hire year/age method, hazard ratio, and lag period <sup>a</sup>

Hire year/age method	Mean hazard ratio							Empirical type I error rate ( $\alpha = 0.05$ ) <sup>b</sup>			
	Hazard ratio	Lag (years)	Includes lagged-out	Restricted-			Year- of- Restricted-				
				Unadjusted	Category-of- birth adjusted	Year-of-birth adjusted	cubic- spline adjusted	Unadjusted	Category- of- birth adjusted	birth adjusted	cubic-spline adjusted
1:Independent	1	0	---	1.00	1.00	1.00	1.00	0.056	0.062	0.060	0.056
	2	0	---	2.01	2.00	2.00	2.00	0.030	0.028	0.028	0.030
	2	10	Yes	2.05	2.02	2.01	2.01	0.064	0.060	0.060	0.054
			No	2.04	2.03	2.02	2.02	0.070	0.056	0.056	0.054
	2	20	Yes	<b>2.09</b>	2.04	2.03	2.01	<b>0.080</b> <sup>c</sup>	0.054	0.054	0.056
			No	<b>2.08</b>	2.06	2.05	2.04	<b>0.074</b> <sup>c</sup>	0.056	0.050	0.048
2:Negative	1	0	---	1.02	1.01	1.00	1.00	0.048	0.046	0.046	0.048
	2	0	---	2.05	2.02	2.01	2.01	0.066	0.048	0.042	0.044
	2	10	Yes	<b>2.10</b>	2.03	2.01	2.00	<b>0.098</b> <sup>c</sup>	0.038	0.038	0.032
			No	<b>2.07</b>	2.03	2.02	2.01	<b>0.074</b> <sup>c</sup>	0.048	0.048	0.040
	2	20	Yes	<b>2.20</b>	2.06	2.04	2.01	<b>0.162</b> <sup>c</sup>	0.044	0.036	0.040
			No	2.11	2.07	2.05	2.04	0.064	0.046	0.042	0.034

3:Reading	1	0	---	1.03	1.01	1.01	1.01	0.048	0.042	0.042	0.044
	2	0	---	2.05	2.02	2.01	2.01	0.070	0.052	0.046	0.046
	2	10	Yes	<b>2.13</b>	2.04	2.03	2.02	<b>0.142</b> <sup>c</sup>	0.062	0.048	0.048
			No	<b>2.10</b>	2.04	2.03	2.03	<b>0.102</b> <sup>c</sup>	0.056	0.054	0.054
	2	20	Yes	<b>2.21</b>	2.05	2.01	2.00	<b>0.220</b> <sup>c</sup>	0.052	0.044	0.044
			No	<b>2.12</b>	2.05	2.03	2.02	<b>0.090</b> <sup>c</sup>	0.056	0.046	0.046

<sup>a</sup> Five hundred cohorts were generated for each combination of hire year/age method and risk scenario. Each cohort contained 4000 (hire year/age methods 1 and 2) or 3569 (hire year/age method 3) workers. Exposure intensity was assigned to each worker-year of exposure using a log-normal distribution with a geometric mean of one unit and a geometric standard deviation of 2.2.

<sup>b</sup> The empirical type I error rate ( $\alpha = 0.05$ ) is the fraction of cohorts for which the 95% confidence interval for the log hazard ratio does not contain the true value.

<sup>c</sup> Significantly elevated (based on the z-test for proportions) empirical type I error rates (i.e., Type I error rate greater than 5% with a two-sided p-value  $< 0.05$ ) are indicated in bold-face type. Actual p-values for flagged results are 0.013, 0.040, 0.0003, 0.022,  $<0.0001$ ,  $<0.0001$ , 0.0001,  $<0.0001$ , and 0.0018 for type I error rates of 0.080, 0.074, 0.098, 0.074, 0.162, 0.142, 0.102, 0.220, and 0.090, respectively.

eTable 4: Mean hazard ratios and empirical type I error rates ( $\alpha = 0.05$ ) for cohorts simulated under an **increasing** exposure intensity for **incomplete** follow-up (through the earlier of death, loss to follow-up or administrative censoring in 1990), by hire year/age method, hazard ratio, and lag period <sup>a</sup>

Hire year/age method	Mean hazard ratio							Empirical type I error rate ( $\alpha = 0.05$ ) <sup>b</sup>			
	Hazard ratio	Lag (years)	Includes lagged-out	Restricted-				Restricted-			
				Unadjusted	Category-of- birth adjusted	Year-of-birth adjusted	cubic-spline adjusted	Unadjusted	Category-of- birth adjusted	Year-of-birth adjusted	cubic-spline adjusted
1:Independent	1	0	---	1.01	1.00	0.98	0.99	0.068	0.052	0.056	0.058
	2	0	---	2.05	2.02	1.99	2.01	0.072	0.042	0.026	0.036
	2	10	Yes	2.08	2.03	1.99	2.01	0.070	0.056	0.038	0.054
			No	<b>2.07</b>	2.03	2.00	2.01	<b>0.074</b> <sup>c</sup>	0.066	0.054	0.064
	2	20	Yes	2.14	2.05	2.01	2.01	0.068	0.048	0.046	0.046
			No	2.13	2.07	2.04	2.04	0.068	0.062	0.054	0.054
2:Negative	1	0	---	<b>1.04</b>	1.00	0.98	0.99	<b>0.086</b> <sup>c</sup>	0.062	0.054	0.058
	2	0	---	2.07	2.01	1.96	2.00	0.070	0.042	0.050	0.038
	2	10	Yes	<b>2.15</b>	2.04	1.99	2.01	<b>0.146</b> <sup>c</sup>	0.054	0.038	0.036
			No	<b>2.12</b>	2.04	1.99	2.02	<b>0.102</b> <sup>c</sup>	0.054	0.038	0.040
	2	20	Yes	<b>2.30</b>	2.07	2.01	2.00	<b>0.204</b> <sup>c</sup>	0.054	0.052	0.052
			No	<b>2.19</b>	2.09	2.05	2.05	<b>0.110</b> <sup>c</sup>	0.066	0.066	0.058

3:Reading	1	0	---	<b>1.04</b>	1.01	0.99	1.00	<b>0.076</b> <sup>c</sup>	0.058	0.052	0.056
	2	0	---	<b>2.08</b>	2.02	1.98	2.01	<b>0.082</b> <sup>c</sup>	0.046	0.036	0.044
	2	10	Yes	<b>2.17</b>	2.05	2.00	2.02	<b>0.148</b> <sup>c</sup>	0.062	0.048	0.054
			No	<b>2.13</b>	2.05	2.01	2.02	<b>0.118</b> <sup>c</sup>	0.062	0.046	0.050
	2	20	Yes	<b>2.31</b>	2.06	2.01	2.00	<b>0.238</b> <sup>c</sup>	0.058	0.048	0.048
			No	<b>2.18</b>	2.07	2.03	2.03	<b>0.098</b> <sup>c</sup>	0.054	0.044	0.046

---

<sup>a</sup> Five hundred cohorts were generated for each combination of hire year/age method and risk scenario. Each cohort contained 4000 (hire year/age methods 1 and 2) or 3569 (hire year/age method 3) workers. Exposure intensity was assigned to each worker-year of exposure using a log-normal distribution with a geometric standard deviation of 2.2 and a geometric mean that increased 2.5% each year.

<sup>b</sup> The empirical type I error rate ( $\alpha = 0.05$ ) is the fraction of cohorts for which the 95% confidence interval for the log hazard ratio does not contain the true value.

<sup>c</sup> Significantly elevated (based on the z-test for proportions) empirical type I error rates (i.e., Type I error rate greater than 5% with a two-sided p-value  $< 0.05$ ) are indicated in bold-face type. Actual p-values for flagged results are 0.040, 0.0041,  $<0.0001$ , 0.0001,  $<0.0001$ ,  $<0.0001$ , 0.028, 0.0091,  $<0.0001$ ,  $<0.0001$ ,  $<0.0001$ , and 0.0003 for type I error rates of 0.074, 0.086, 0.146, 0.102, 0.204, 0.110, 0.076, 0.082, 0.148, 0.118, 0.238, 0.098, respectively.