

1 Log files from Stata (external links)

[SHARE-1-harmon \(reshape data\)](#)

[SHARE-2-harmon-Part5 \(generation of disease variables\)](#)

2 Sample Selection

```
name: log
log: C:/Users/User/Documents/GitHub/2-projectMM-SHARE/files/logs/log-sampleselection.txt
log type: text
opened on: 11 Jan 2024, 17:51:16
```

. tab hacohort wave,m

hacohort: Sample cohort	Survey Wave							Total
	2004 wave	2006/07 w	2011/12 w	2013 wave	2015 wave	2017 wave	2019/20 w	
1.Original sample for	51,312	51,312	51,312	51,312	51,312	51,312	51,312	359,184
2.2006 Refreshment sa	5,836	5,836	5,836	5,836	5,836	5,836	5,836	40,852
Total	57,148	57,148	57,148	57,148	57,148	57,148	57,148	400,036

. tab iwstatr wave,m

r interview status	Survey Wave							Total
	2004 wave	2006/07 w	2011/12 w	2013 wave	2015 wave	2017 wave	2019/20 w	
0.inap.	37,079	25,054	17,649	20,009	14,567	1,898	1,730	117,986
1.resp, alive	20,069	25,176	26,853	22,511	23,322	32,289	19,780	170,000
4.nr, alive	0	4,828	7,138	6,779	8,805	8,662	10,489	46,701
5.nr, died this wv	0	279	468	593	603	958	1,061	3,962
6.nr, died prev wv	0	0	660	1,096	1,895	2,799	3,983	10,433
9.nr, dk if alive or	0	1,811	4,380	6,160	7,956	10,542	20,105	50,954
Total	57,148	57,148	57,148	57,148	57,148	57,148	57,148	400,036

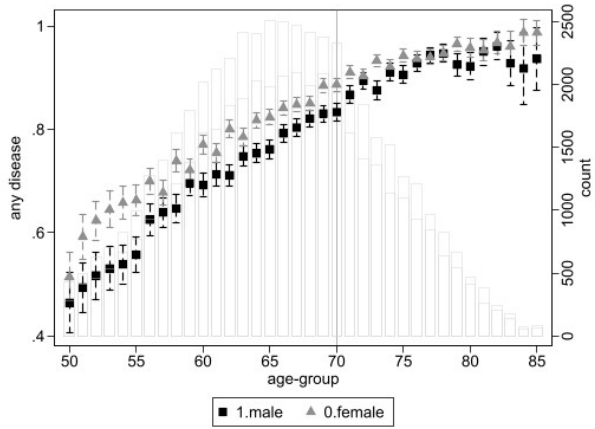
```
. *tab countryID wave,m
. sum agemin
```

Variable	Obs	Mean	Std. dev.	Min	Max
age	400,036	59.13131	6.417423	14	70

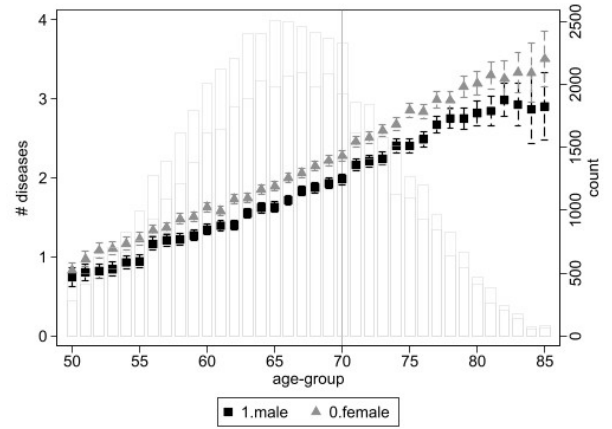
```
. log close log
name: log
log: C:/Users/User/Documents/GitHub/2-projectMM-SHARE/files/logs/log-sampleselection.txt
log type: text
closed on: 11 Jan 2024, 17:51:16
```

Figures and Tables

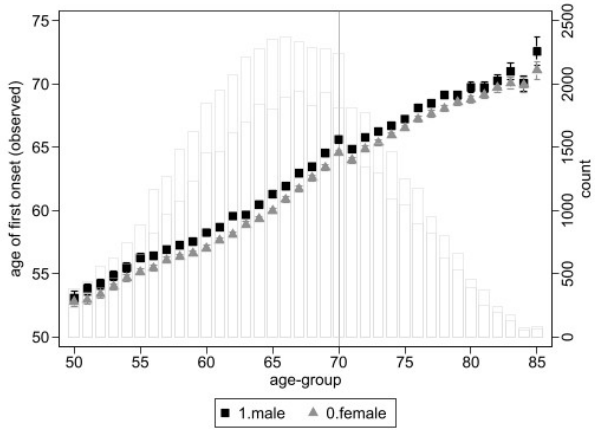
Figure 1: Outcome variable by age group



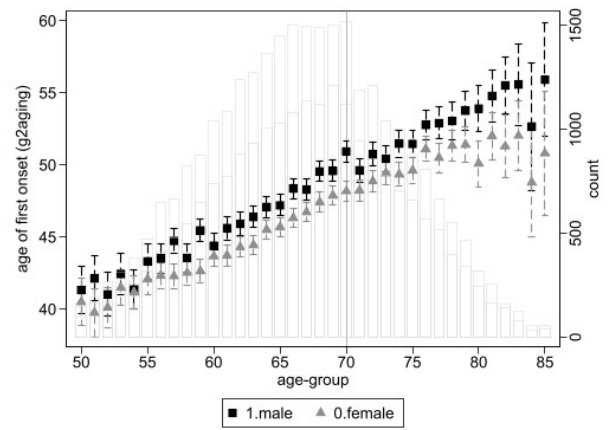
(a) at least one disease by age



(b) number of diseases by age



(c) age at first onset observed by age



(d) age at first onset self reported by age

Note for figure c: For individuals who already have at least one disease at baseline, the age of onset is the baseline age

(a) Count of Diseases by age at baseline

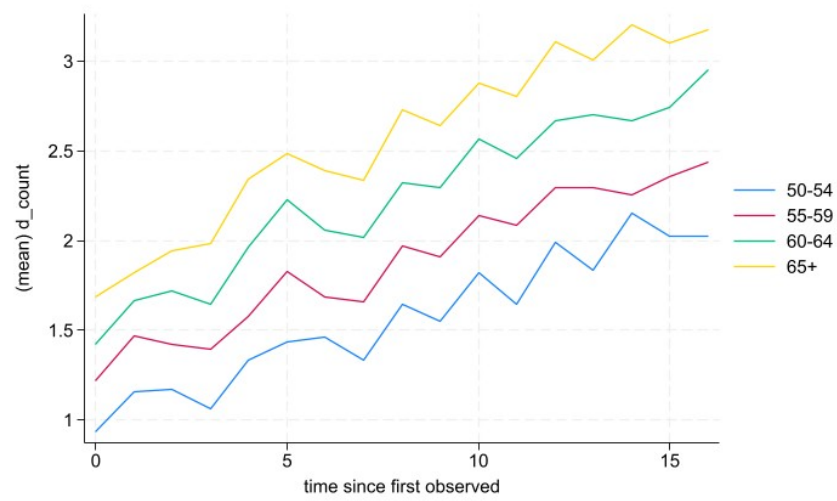
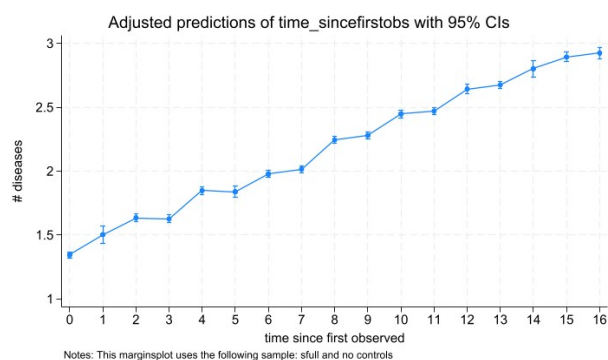
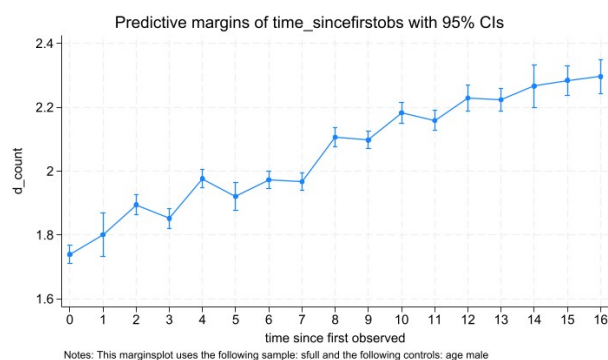


Figure 3: Predictions of disease count adjusted by time since baseline without (left) and with (right - age, gender) control variables

Panel A: Full Sample

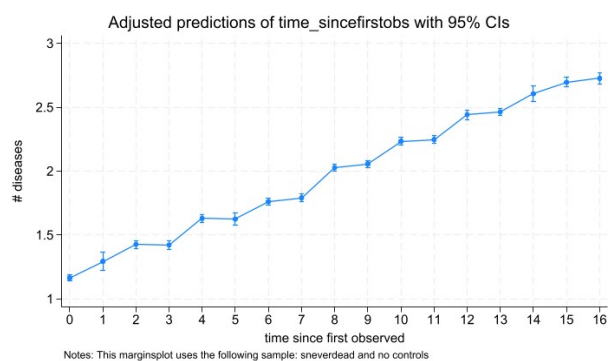


(a) without controls
number of diseases by time of follow-up (years)

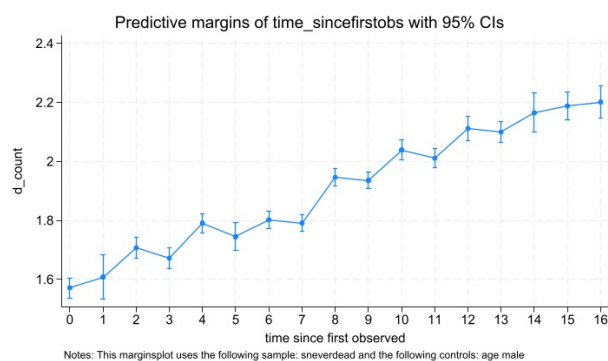


(b) with controls
number of diseases by time of follow-up (years) - adjusted by age and sex

Panel B: Full sample, accounting for mortality (only people who never die between 2004-2019)

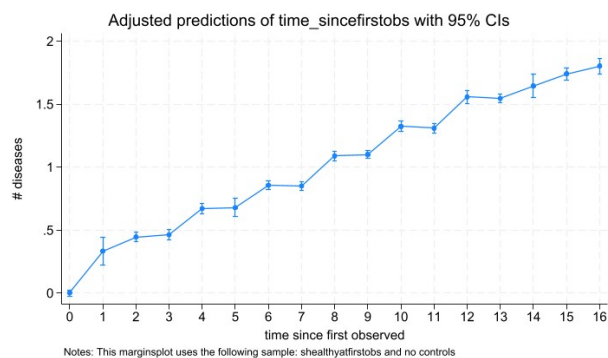


(c) without controls

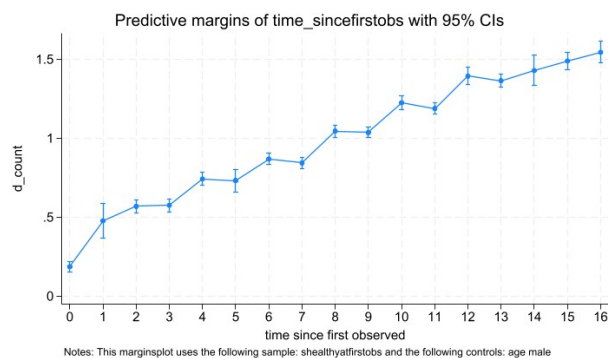


(d) with controls

Panel C: Full sample, but only people who are healthy-at-baseline



(e) without controls

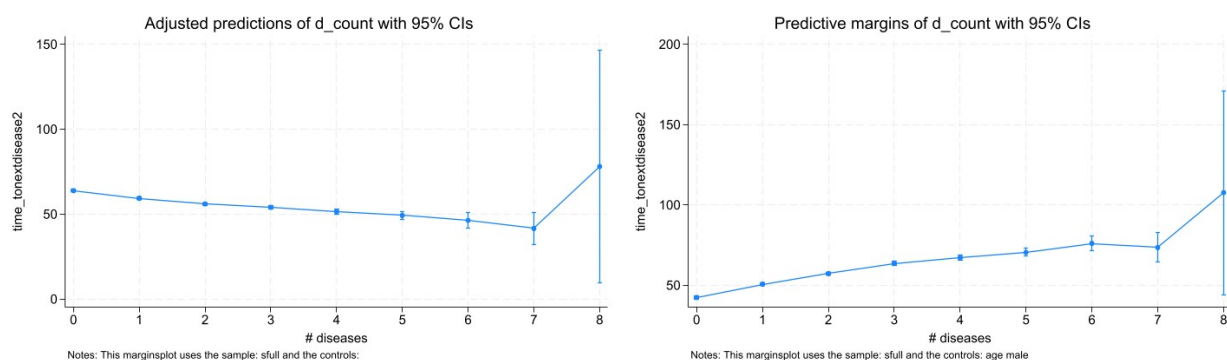


(f) with controls

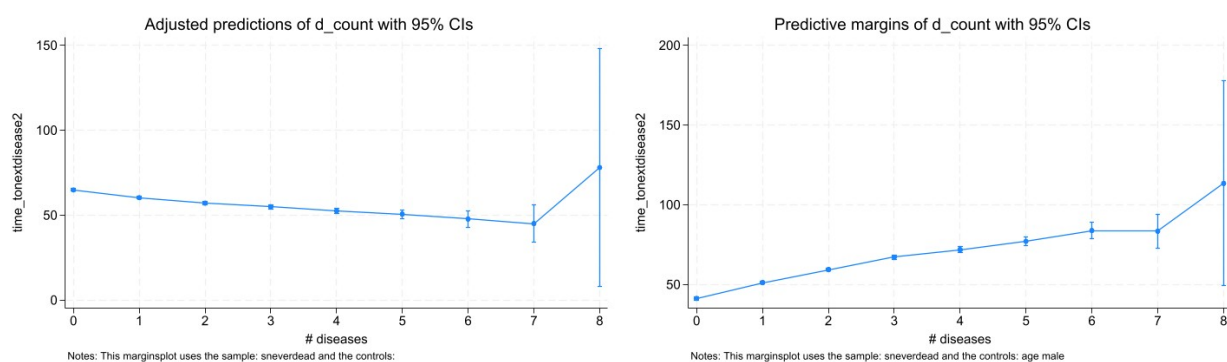
Notes: time_sincefirstobs refers to years (year at interview rather than survey year/wave)

Figure 4: Duration from count to higher count, by disease count

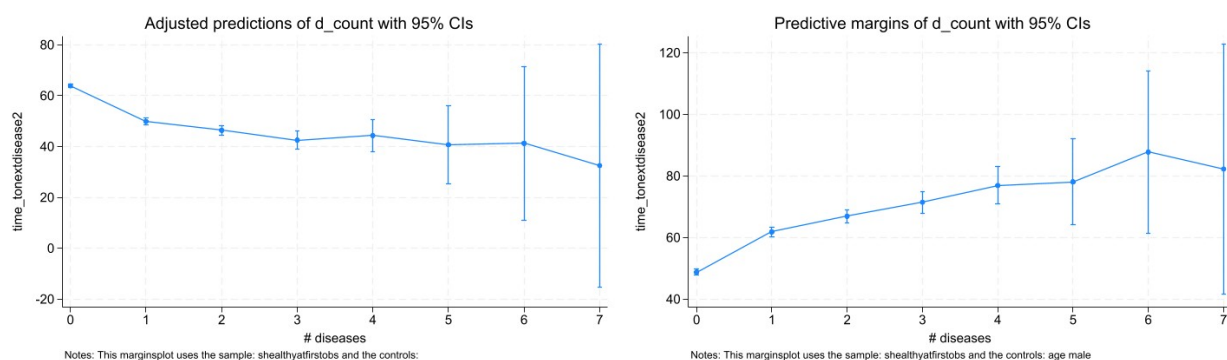
Panel A: Full Sample



Panel B: Full sample, accounting for mortality (only people who never die between 2004-2019)



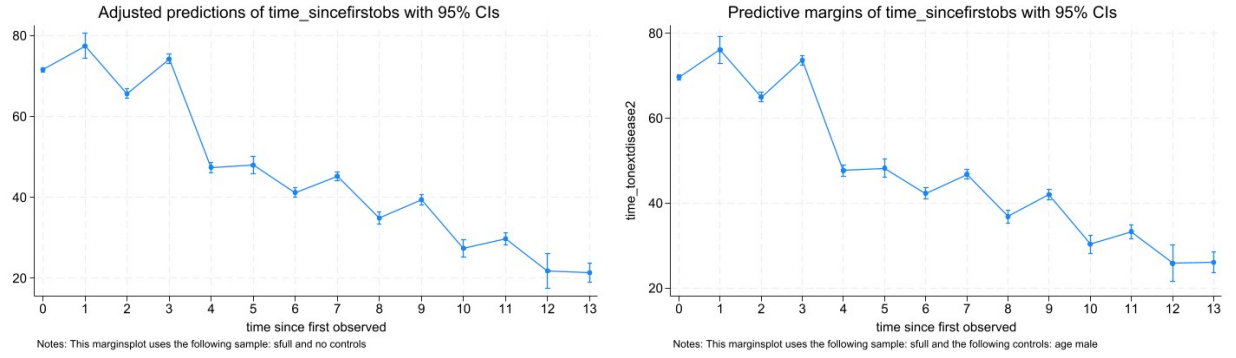
Panel C: Full sample, but only people who are healthy-at-baseline



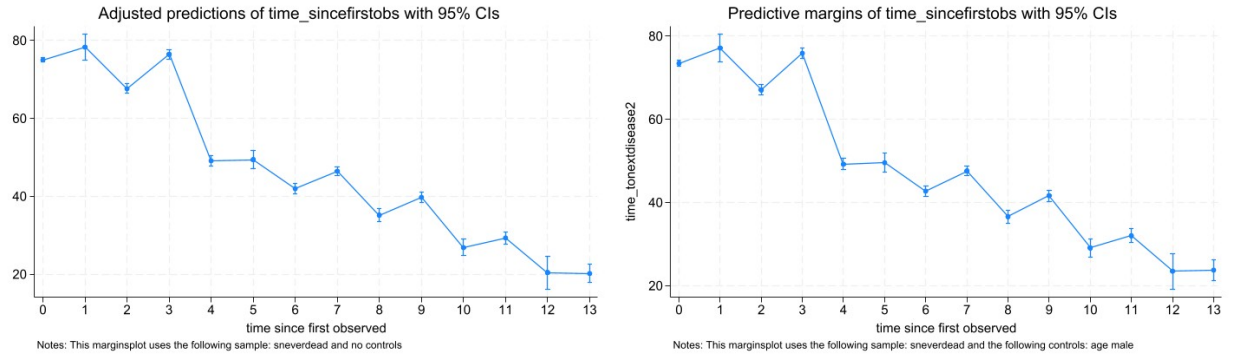
Notes: Only observations with positive changes c to $c+1$ diseases are used, which may cause the plots to be similar

Figure 5: Duration from count to higher count, by time

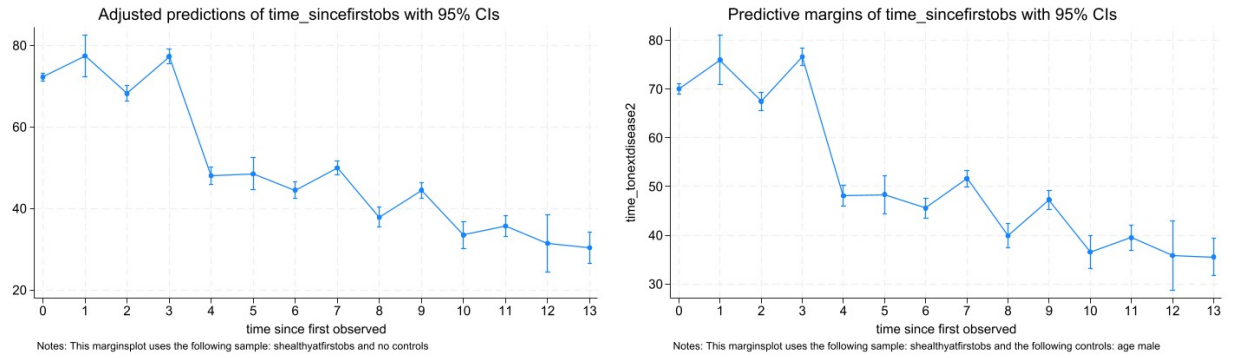
Panel A: Full Sample



Panel B: Full sample, accounting for mortality (only people who never die between 2004-2019)



Panel C: Full sample, but only people who are healthy-at-baseline



Notes: Only observations with positive changes c to $c+1$ diseases are used, which may cause the plots to be similar

Note: summary statistics contain the following: nobs — mean(sd)

Table 1: Full sample: main statistics

	2004 wave	2006/07 wave	2011/12 wave	2013 wave	2015 wave	2017 wave	2019/20 wave
	10,261 (9.6%)	13,859 (13.0%)	16,089 (15.1%)	16,397 (15.4%)	16,945 (15.9%)	17,516 (16.4%)	15,433 (14.5%)
age at interview	10,261 59.40 (6.27)	13,580 60.76 (6.49)	14,961 63.72 (6.58)	14,708 65.75 (6.55)	14,447 67.56 (6.51)	13,759 69.31 (6.53)	10,389 71.72 (6.43)
r interview year	10,261 2004.35	13,580 2006.93	14,961 2011.01	14,708 2013.00	14,447 2015.00	13,759 2017.00	10,389 2019.49
rabyear: r birth year	10,261 1944.57	13,859 1945.54	16,089 1946.38	16,397 1946.25	16,945 1946.30	17,516 1946.41	15,433 1946.41
radyear: r death year	2,196 2012.75	2,836 2013.20	3,518 2013.71	3,197 2013.55	3,632 2013.59	4,332 2013.92	4,564 2014.03
dead	10,261 0.00	13,859 0.02	16,089 0.07	16,397 0.10	16,945 0.15	17,516 0.21	15,433 0.33
already has disease at baseline	10,261 0.65	13,859 0.65	16,089 0.67	16,397 0.66	16,945 0.67	17,515 0.68	15,432 0.68
male	10,261 0.46	13,859 0.45	16,089 0.46	16,397 0.45	16,945 0.45	17,516 0.46	15,433 0.47
any disease	10,261 0.65	13,580 0.70	14,961 0.78	14,708 0.81	14,447 0.85	13,754 0.87	10,389 0.89
# diseases	10,234 1.20 (1.24)	13,554 1.41 (1.35)	14,936 1.74 (1.46)	14,695 1.93 (1.54)	14,435 2.17 (1.60)	13,708 2.25 (1.59)	10,373 2.47 (1.65)
age of first onset (observed)	9,459 61.50 (6.67)	12,816 61.26 (6.63)	14,871 61.37 (6.52)	15,108 61.40 (6.55)	15,635 61.41 (6.50)	16,091 61.42 (6.49)	14,144 61.60 (6.43)
age of first onset (g2aging)	5,635 47.48 (12.88)	7,688 47.46 (12.72)	9,224 46.93 (13.20)	9,209 47.19 (13.17)	9,652 47.26 (13.08)	9,947 47.22 (13.09)	8,891 47.09 (13.28)
years first onset to death (observed)	1,907 7.72 (4.50)	2,522 7.38 (4.37)	3,159 6.47 (4.09)	2,861 6.78 (4.23)	3,231 6.61 (4.13)	3,778 6.39 (4.12)	3,984 6.38 (4.11)
months (observed) 1 to 2+ diseases	7,617 33.53 (46.12)	10,355 31.77 (43.93)	11,925 28.19 (41.21)	12,053 29.61 (42.26)	12,550 28.48 (41.40)	12,900 27.63 (41.16)	11,303 26.80 (41.61)
months (observed) 2 to 3+ diseases	5,273 34.55 (44.85)	7,188 33.12 (42.63)	8,219 30.01 (40.44)	8,275 30.76 (40.96)	8,687 29.91 (40.28)	8,917 29.29 (40.20)	7,809 28.84 (40.81)
months (observed) 3 to 4+ diseases	3,129 33.72 (42.73)	4,280 31.86 (40.79)	4,840 28.97 (38.36)	4,862 29.58 (39.05)	5,158 29.40 (39.03)	5,284 28.76 (38.88)	4,644 28.10 (39.11)
months (observed) 4 to 5+ diseases	1,615 30.28 (39.49)	2,192 29.66 (38.66)	2,447 27.83 (36.33)	2,462 28.46 (37.03)	2,626 27.94 (36.83)	2,691 27.60 (36.82)	2,375 27.24 (37.09)
months (observed) 5 to 6+ diseases	651 27.87 (36.75)	867 27.84 (35.31)	956 27.38 (34.29)	973 27.59 (35.21)	1,043 26.72 (34.11)	1,066 26.34 (34.64)	953 26.57 (34.71)
months (observed) 6 to 7+ diseases	213 25.84 (33.80)	276 25.07 (33.19)	289 23.25 (31.16)	307 22.66 (30.52)	325 23.58 (31.92)	329 22.68 (31.59)	291 22.80 (32.18)
months (observed) 7 to 8+ diseases	43 21.30 (29.89)	59 21.88 (28.12)	59 23.64 (29.15)	65 25.11 (28.32)	67 24.36 (27.47)	70 24.43 (27.94)	58 26.43 (29.06)
months (observed) 8 to 9+ diseases	4 19.50 (39.00)	4 19.50 (39.00)	7 11.14 (29.48)	6 13.00 (31.84)	6 13.00 (31.84)	7 11.14 (29.48)	5 0.00 (0.00)
time_tonextdisease2	7,420 77.11 (45.28)	8,385 76.63 (31.59)	7,840 48.81 (27.26)	5,739 43.62 (22.43)	3,847 37.39 (15.43)	1,931 31.30 (1.90)	0 . (.)
r educ							
1.less than upper secondary	4,599 (44.8%)	5,949 (42.9%)	6,399 (39.8%)	6,479 (39.5%)	6,760 (39.9%)	7,036 (40.2%)	6,044 (39.2%)
2.upper secondary or vocational	3,292 (32.1%)	4,829 (34.8%)	6,249 (38.8%)	6,199 (37.8%)	6,482 (38.3%)	6,743 (38.5%)	6,114 (39.6%)
3.tertiary education	2,370 (23.1%)	3,081 (22.2%)	3,441 (21.4%)	3,719 (22.7%)	3,703 (21.9%)	3,737 (21.3%)	3,275 (21.2%)
cohort							
50-59	4,842 (48.9%)	5,659 (42.7%)	4,098 (27.6%)	2,723 (18.6%)	1,533 (10.6%)	721 (5.2%)	206 (2.0%)
60-69	4,653 (47.0%)	6,219 (46.9%)	7,443 (50.2%)	7,340 (50.2%)	7,247 (50.3%)	6,422 (46.8%)	3,827 (36.9%)
70-79	410 (4.1%)	1,389 (10.5%)	3,282 (22.1%)	4,569 (31.2%)	5,281 (36.7%)	5,693 (41.5%)	5,012 (48.3%)
80+	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (0.0%)	348 (2.4%)	898 (6.5%)	1,333 (12.8%)

Table 2: Full sample: all diseases

	2004 wave	2006/07 wave	2011/12 wave	2013 wave	2015 wave	2017 wave	2019/20 wave
	10,261 (9.6%)	13,859 (13.0%)	16,089 (15.1%)	16,397 (15.4%)	16,945 (15.9%)	17,516 (16.4%)	15,433 (14.5%)
age of first onset (observed) for d_hibp	6,461 63.84 (6.97)	8,799 63.49 (6.88)	10,202 63.38 (6.73)	10,373 63.45 (6.77)	10,810 63.43 (6.71)	11,095 63.41 (6.69)	9,714 63.52 (6.66)
age of first onset (observed) for d_diab	2,397 65.20 (7.25)	3,242 64.99 (7.06)	3,542 64.83 (6.93)	3,713 64.87 (6.97)	3,951 64.89 (6.89)	4,095 64.80 (6.86)	3,631 64.79 (6.85)
age of first onset (observed) for d_heart	3,895 66.67 (7.19)	5,215 66.24 (7.13)	6,006 65.92 (6.96)	6,064 66.11 (6.99)	6,410 66.01 (6.95)	6,638 65.91 (6.92)	5,933 65.92 (6.86)
age of first onset (observed) for d_lung	1,566 65.96 (7.59)	2,198 65.70 (7.47)	2,495 65.24 (7.31)	2,501 65.47 (7.42)	2,609 65.52 (7.30)	2,675 65.44 (7.29)	2,383 65.45 (7.21)
age of first onset (observed) for d_psych	5,792 63.31 (7.97)	7,851 62.99 (7.88)	9,148 62.91 (7.65)	9,203 63.04 (7.73)	9,658 63.05 (7.68)	9,864 63.01 (7.65)	8,576 63.12 (7.58)
age of first onset (observed) for d_osteo	2,153 65.82 (7.67)	2,822 65.52 (7.57)	2,985 65.45 (7.38)	3,182 65.59 (7.42)	3,249 65.58 (7.35)	3,308 65.46 (7.39)	2,732 65.63 (7.34)
age of first onset (observed) for d_cancr	1,813 66.39 (7.27)	2,333 66.27 (7.29)	2,666 65.94 (7.22)	2,739 66.06 (7.19)	2,788 65.99 (7.16)	2,870 65.92 (7.16)	2,633 65.94 (7.10)
age of first onset (observed) for d_strok	1,125 68.15 (7.57)	1,467 67.79 (7.53)	1,680 67.22 (7.39)	1,740 67.43 (7.37)	1,811 67.34 (7.28)	1,886 67.21 (7.23)	1,677 67.32 (7.26)
age of first onset (observed) for d_arthr	4,948 64.81 (7.28)	6,950 64.47 (7.19)	8,037 64.25 (7.10)	7,901 64.44 (7.14)	8,229 64.37 (7.08)	8,362 64.32 (7.08)	7,174 64.51 (6.97)
radiaghibp: R age first diagnosed with high BP	3,080 50.31 (10.98)	4,326 50.23 (10.73)	5,349 49.68 (11.23)	5,325 49.94 (11.18)	5,589 49.98 (11.08)	5,780 49.99 (11.05)	5,224 49.80 (11.30)
radiagdiab: R age first diagnosed with diabetes	1,008 52.39 (10.51)	1,364 52.33 (10.08)	1,583 52.53 (10.39)	1,619 52.51 (10.35)	1,746 52.62 (10.40)	1,844 52.47 (10.41)	1,724 52.52 (10.52)
radiagcancr: R age first diagnosed with cancer	544 53.06 (10.58)	699 53.37 (10.17)	882 53.66 (10.38)	885 53.27 (10.73)	918 53.50 (10.51)	958 53.52 (10.65)	918 53.68 (10.67)
radiaglhung: R age first diagnosed with lung disease	472 43.13 (17.94)	655 44.20 (17.79)	863 43.47 (18.03)	829 42.95 (18.33)	865 43.68 (18.01)	906 43.74 (17.96)	859 43.68 (17.89)
radiagheart: R age first diagnosed with heart problem	958 51.79 (11.72)	1,330 51.32 (11.84)	1,715 50.11 (12.96)	1,656 50.48 (12.84)	1,789 50.42 (12.86)	1,892 50.48 (12.72)	1,807 50.19 (13.13)
radiagstrok: R age first diagnosed with stroke	294 54.05 (9.75)	397 53.39 (10.34)	534 53.38 (10.87)	515 53.55 (10.95)	556 53.41 (10.92)	592 53.60 (10.87)	573 53.81 (10.43)
radiagarthr: R age first diagnosed with arthritis	1,836 46.74 (12.93)	2,600 46.78 (12.87)	3,219 47.08 (12.78)	3,013 47.17 (12.93)	3,224 47.21 (12.80)	3,334 47.19 (12.84)	2,934 47.44 (12.85)
radiaghip: R age first diagnosed with hip fracture	130 45.78 (18.80)	182 44.23 (18.24)	227 43.40 (18.48)	224 44.31 (18.71)	235 43.88 (18.75)	258 43.83 (18.48)	236 44.33 (18.26)
radiagosteo: R age first diagnosed with osteoporosis	655 52.78 (10.35)	879 52.83 (10.28)	782 52.41 (10.55)	823 52.76 (10.44)	860 52.96 (10.34)	888 52.93 (10.50)	724 53.35 (10.37)
radiagpsych: R age first diagnosed with psych disorder	8 38.38 (20.30)	14 40.57 (17.66)	19 41.16 (17.25)	22 42.18 (17.64)	25 43.80 (18.09)	25 43.80 (18.09)	21 44.24 (18.93)
radiagkidney: R age first diagnosed with kidney disease	0 . (.)	0 . (.)	1 65.00 (.)	1 65.00 (.)	9 53.00 (9.62)	9 53.00 (9.62)	9 53.00 (9.62)

Table 3: healthy-at-baseline sample: main statistics

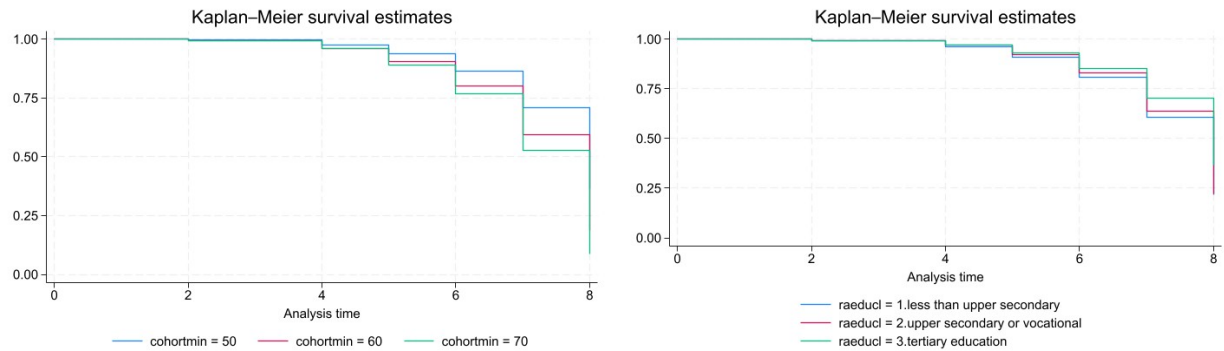
	2004 wave	2006/07 wave	2011/12 wave	2013 wave	2015 wave	2017 wave	2019/20 wave
	3,602 (10.2%)	4,810 (13.6%)	5,250 (14.9%)	5,518 (15.6%)	5,563 (15.7%)	5,711 (16.2%)	4,869 (13.8%)
age at interview	3,602 57.85 (6.24)	4,760 59.19 (6.46)	5,028 62.45 (6.56)	5,140 64.45 (6.52)	5,010 66.23 (6.50)	4,903 67.99 (6.57)	3,790 70.54 (6.38)
r interview year	3,602 2004.31	4,760 2006.91	5,028 2011.01	5,140 2013.00	5,010 2015.00	4,903 2017.00	3,790 2019.48
rubyear: r birth year	3,602 1946.09	4,810 1947.12	5,250 1947.78	5,518 1947.67	5,563 1947.76	5,711 1947.94	4,869 1947.86
rubyear: r death year	502 2013.15	626 2013.63	687 2013.75	674 2013.64	758 2013.66	921 2014.15	951 2014.22
dead	3,602 0.00	4,810 0.01	5,250 0.04	5,518 0.07	5,563 0.10	5,711 0.14	4,869 0.22
already has disease at baseline	3,602 0.00	4,810 0.00	5,250 0.00	5,518 0.00	5,563 0.00	5,663 0.00	4,868 0.00
male	3,602 0.52	4,810 0.51	5,250 0.51	5,518 0.50	5,563 0.51	5,711 0.51	4,869 0.52
any disease	3,602 0.00	4,760 0.23	5,028 0.41	5,140 0.54	5,010 0.64	4,852 0.68	3,790 0.75
# diseases	3,584 0.00 (0.00)	4,743 0.30 (0.63)	5,019 0.62 (0.91)	5,139 0.89 (1.06)	5,008 1.14 (1.16)	4,837 1.30 (1.25)	3,786 1.53 (1.31)
age of first onset (observed)	2,800 64.49 (6.96)	3,767 64.08 (6.84)	4,032 63.99 (6.78)	4,229 64.07 (6.77)	4,253 64.05 (6.73)	4,239 64.04 (6.73)	3,580 64.18 (6.72)
age of first onset (g2aging)	119 53.61 (12.53)	147 53.39 (12.28)	152 52.17 (14.06)	171 52.77 (13.62)	160 52.41 (13.61)	159 52.47 (13.83)	139 52.14 (14.14)
years first onset to death (observed)	288 5.72 (3.81)	382 5.34 (3.57)	389 4.81 (3.46)	396 4.98 (3.55)	427 4.81 (3.44)	463 4.77 (3.41)	471 4.84 (3.44)
months (observed) 1 to 2+ diseases	1,785 37.75 (40.60)	2,385 34.09 (38.28)	2,494 32.82 (36.93)	2,600 32.77 (37.02)	2,644 31.70 (35.96)	2,639 31.59 (36.01)	2,210 33.00 (38.03)
months (observed) 2 to 3+ diseases	936 29.96 (36.01)	1,224 29.01 (34.00)	1,244 27.80 (33.35)	1,304 27.84 (33.39)	1,332 26.89 (32.56)	1,333 26.94 (32.78)	1,119 28.38 (34.33)
months (observed) 3 to 4+ diseases	383 23.98 (31.73)	484 22.37 (29.79)	468 21.46 (29.03)	508 22.14 (29.69)	526 21.62 (29.03)	522 21.61 (29.27)	448 21.10 (29.10)
months (observed) 4 to 5+ diseases	136 22.33 (29.80)	173 19.55 (28.26)	175 20.28 (28.81)	179 20.35 (28.93)	182 18.64 (27.14)	186 19.76 (28.52)	155 20.97 (29.72)
months (observed) 5 to 6+ diseases	32 15.09 (26.34)	40 14.43 (24.12)	40 14.03 (24.29)	43 14.16 (23.61)	45 13.53 (23.25)	46 13.24 (23.08)	38 13.63 (24.34)
months (observed) 6 to 7+ diseases	8 22.38 (25.52)	8 22.38 (25.52)	7 25.57 (25.78)	9 19.89 (25.01)	9 19.89 (25.01)	8 15.38 (22.48)	7 22.00 (27.54)
months (observed) 7 to 8+ diseases	1 32.00 (.)	2 32.50 (0.71)	1 32.00 (.)	2 32.50 (0.71)	2 32.50 (0.71)	2 32.50 (0.71)	2 32.50 (0.71)
months (observed) 8 to 9+ diseases	0 . (.)	0 . (.)	0 . (.)	0 . (.)	0 . (.)	0 . (.)	0 . (.)
time_tonextdisease2	2,786 77.03 (44.98)	3,375 77.21 (32.15)	2,961 50.39 (27.70)	2,301 43.72 (22.46)	1,521 36.89 (15.38)	758 31.25 (1.96)	0 . (.)
r educ							
1.less than upper secondary	1,322 (36.7%)	1,694 (35.2%)	1,735 (33.0%)	1,828 (33.1%)	1,841 (33.1%)	1,911 (33.5%)	1,578 (32.4%)
2.upper secondary or vocational	1,237 (34.3%)	1,779 (37.0%)	2,076 (39.5%)	2,132 (38.6%)	2,201 (39.6%)	2,268 (39.7%)	1,968 (40.4%)
3.tertiary education	1,043 (29.0%)	1,337 (27.8%)	1,439 (27.4%)	1,558 (28.2%)	1,521 (27.3%)	1,532 (26.8%)	1,323 (27.2%)
cohort							
50-59	1,991 (58.5%)	2,392 (52.1%)	1,722 (34.7%)	1,220 (23.9%)	711 (14.3%)	341 (7.0%)	102 (2.7%)
60-69	1,326 (39.0%)	1,887 (41.1%)	2,394 (48.3%)	2,653 (52.0%)	2,722 (54.6%)	2,594 (53.1%)	1,662 (43.9%)
70-79	86 (2.5%)	308 (6.7%)	844 (17.0%)	1,224 (24.0%)	1,466 (29.4%)	1,723 (35.3%)	1,645 (43.5%)
80+	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)	83 (1.7%)	224 (4.6%)	373 (9.9%)

Table 4: healthy-at-baseline sample: all diseases

[illegible]

3 THE PART BELOW CONTAINS PRELIMINARY RESULTS. IT IS SYNCED TO GITHUB AND CAN BE DISREGARDED UNTIL ADDED ABOVE. — For the team to edit/add notes to the part above, use “7-graphsAndTables.tex”

Figure 6: Survival curves



Notes:Analysis time is in waves