

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, 18:07:39
```

. 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

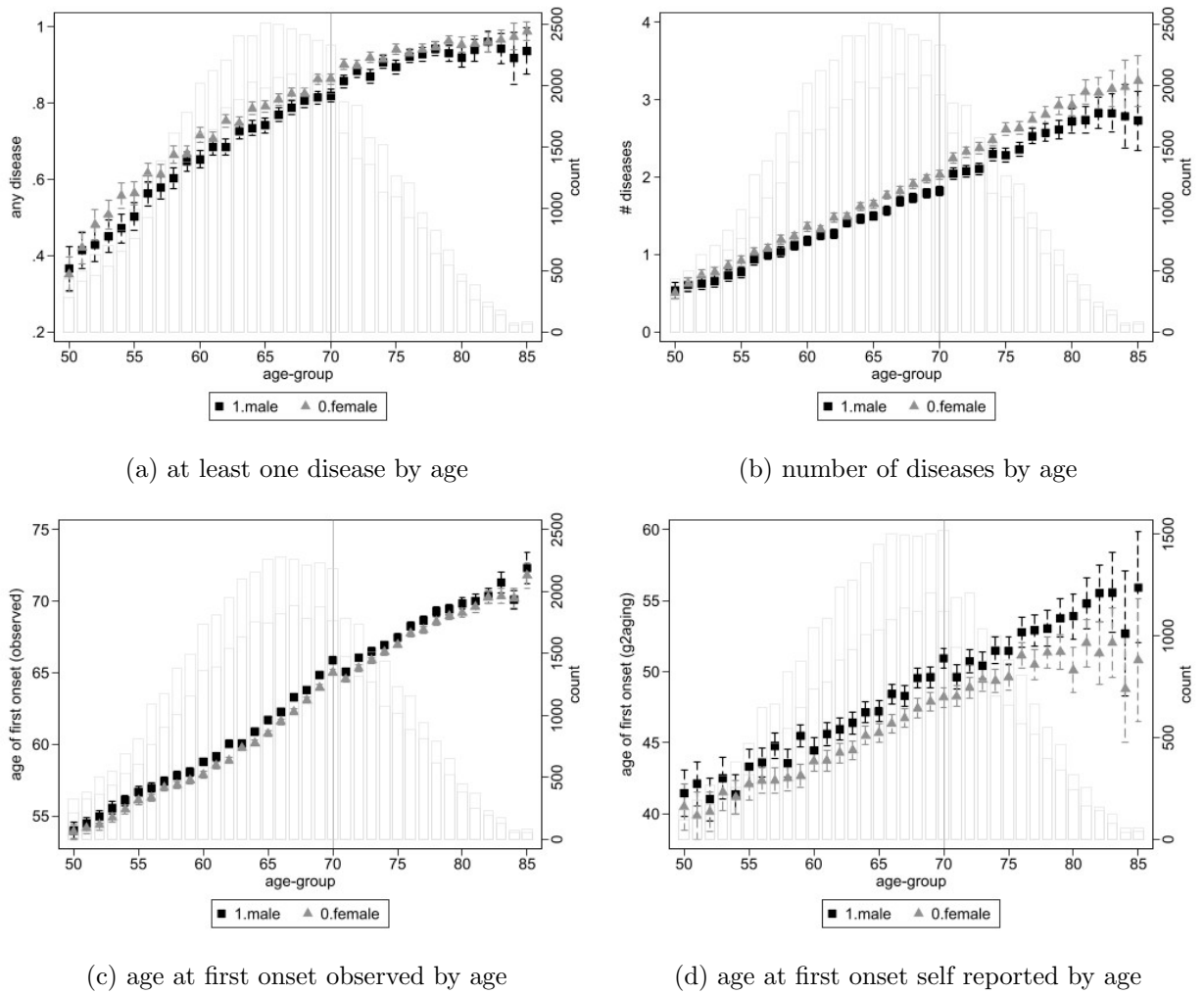
```
. *tab countryID wave,m
. sum agemin
```

Variable	Obs	Mean	Std. dev.	Min	Max
agemin	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, 18:07:39
```

Figures and Tables

Figure 1: Outcome variable by age group



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

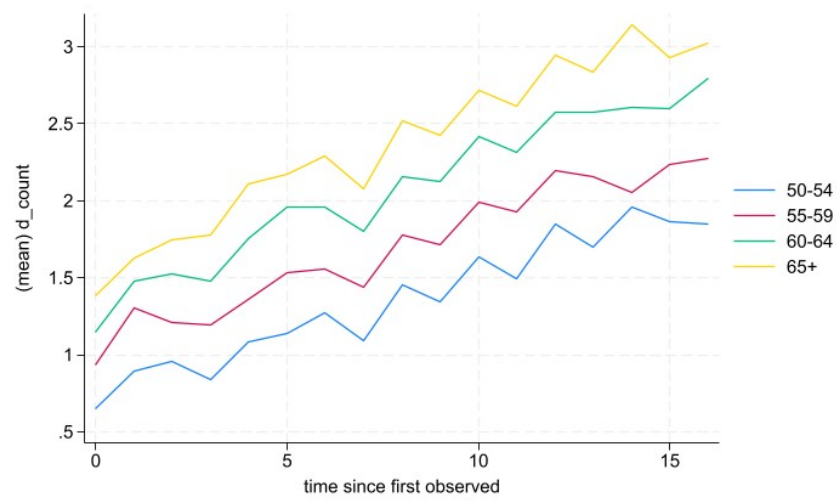
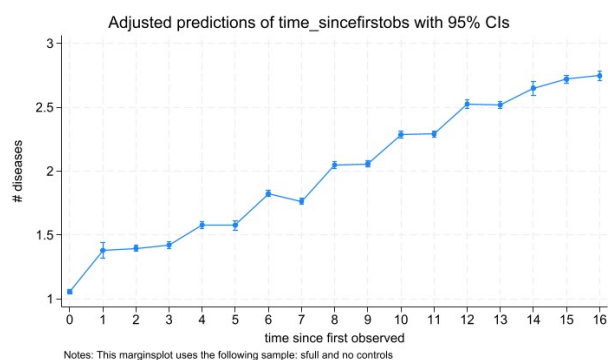
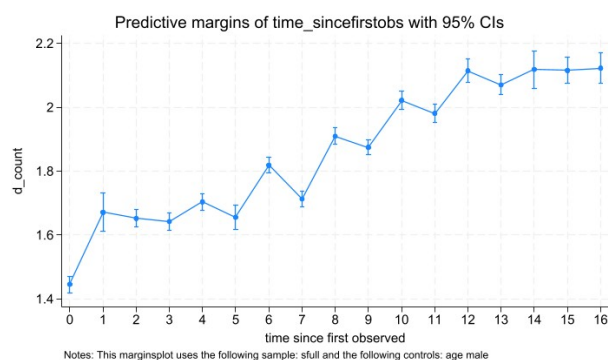


Figure 3: Predictions of disease count adjusted by time since baseline without (left) and with (right - age, gender) control variables

Panel A: Full Sample

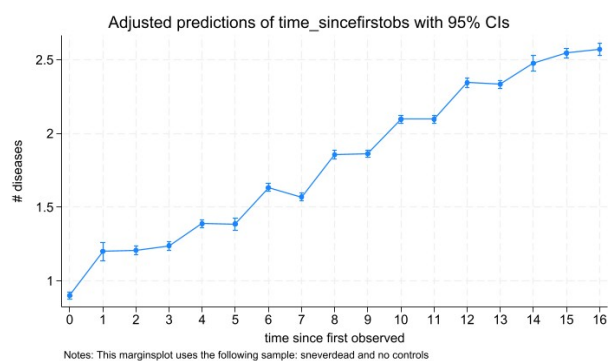


(a) without controls
number of diseases by time of follow-up (years)

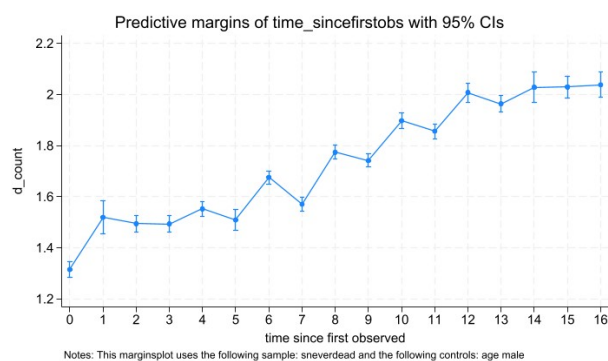


(b) with controls
number of diseases by time of follow-up (years) - adjusted by age and sex

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

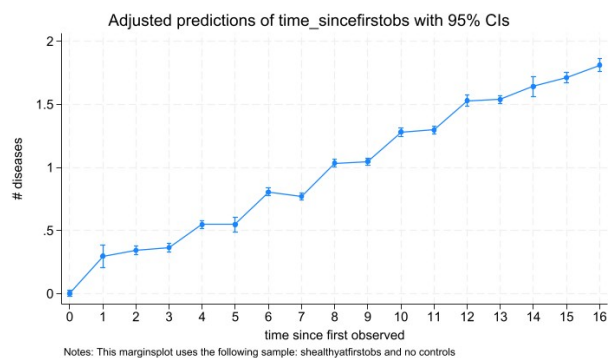


(c) without controls

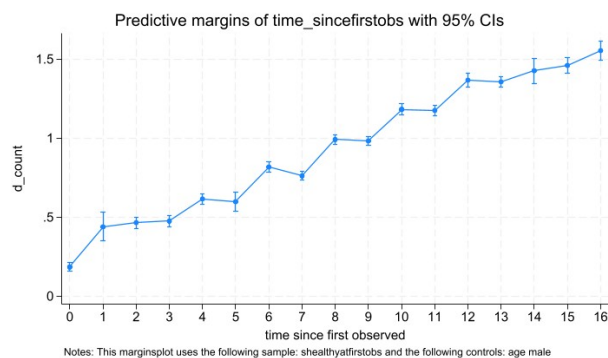


(d) with controls

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



(e) without controls

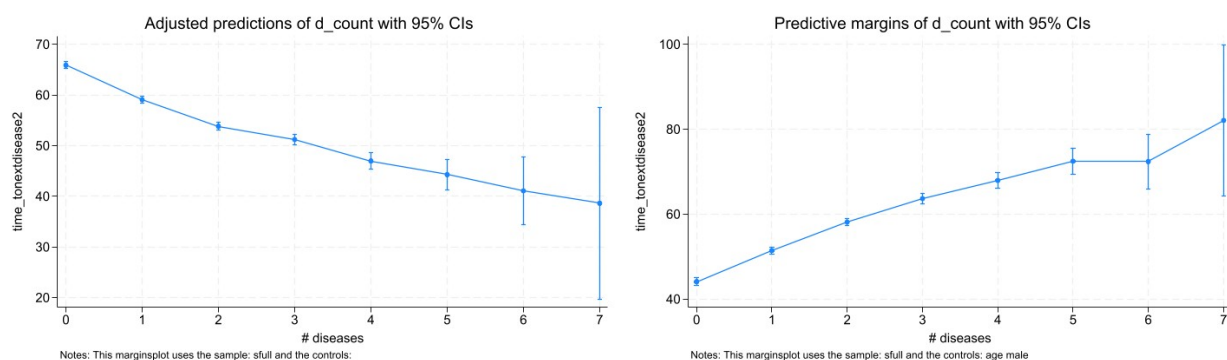


(f) with controls

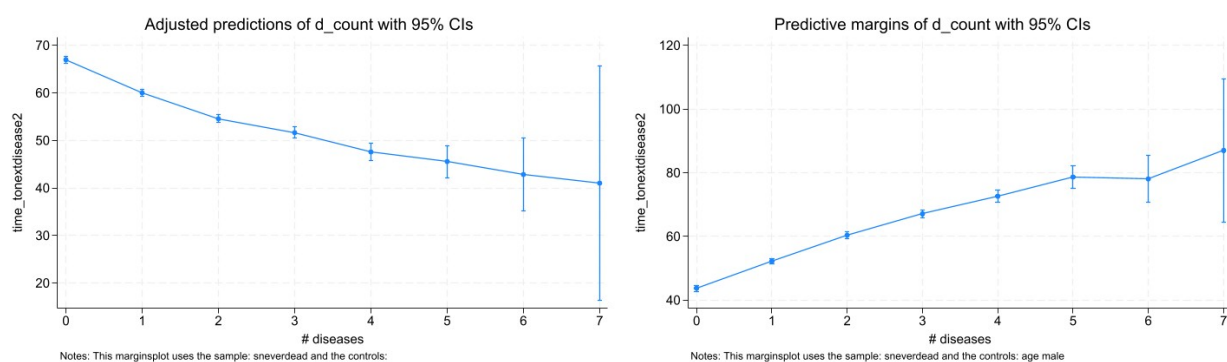
Notes: time_sincefirstobs refers to years (year at interview rather than survey year/wave)

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

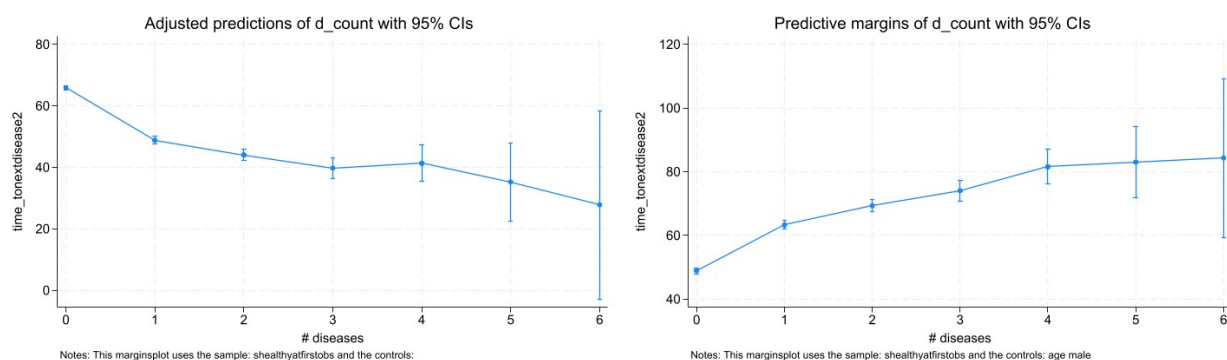
Panel A: Full Sample



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



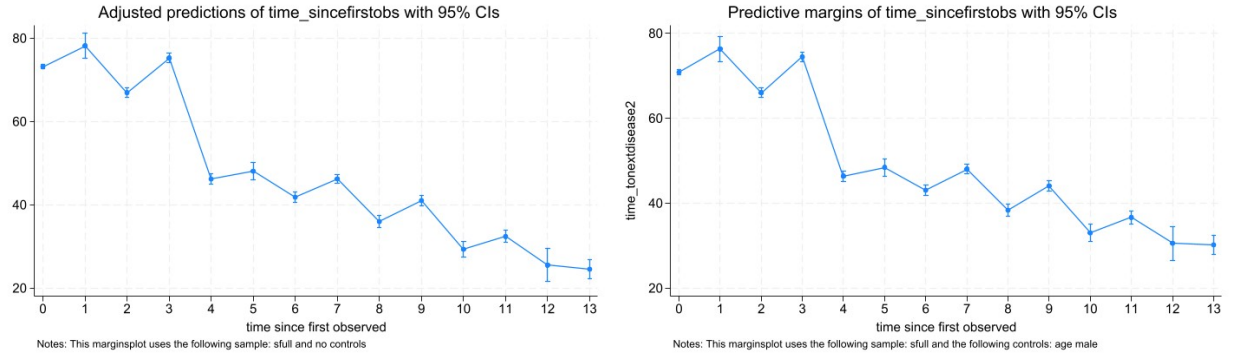
Panel C: Full sample, but only people who are healthy-at-baseline



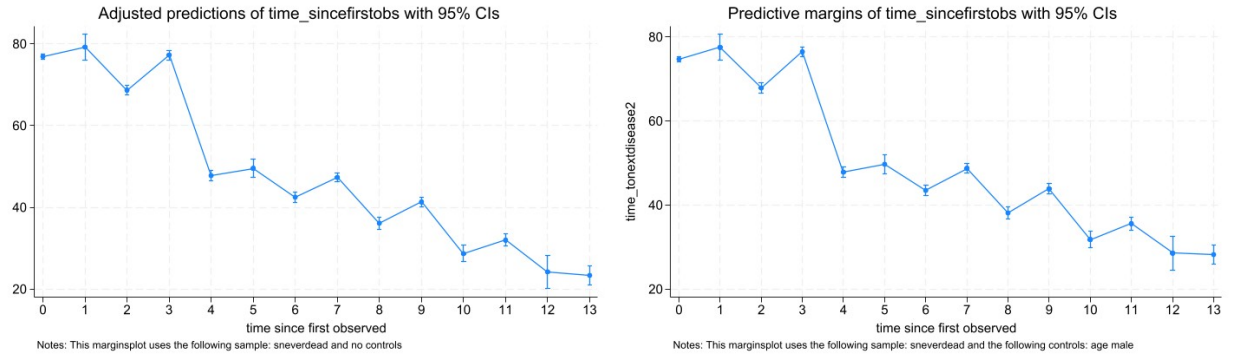
Notes: Only observations with positive changes c to $c+1$ diseases are used, which may cause the plots to be similar

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

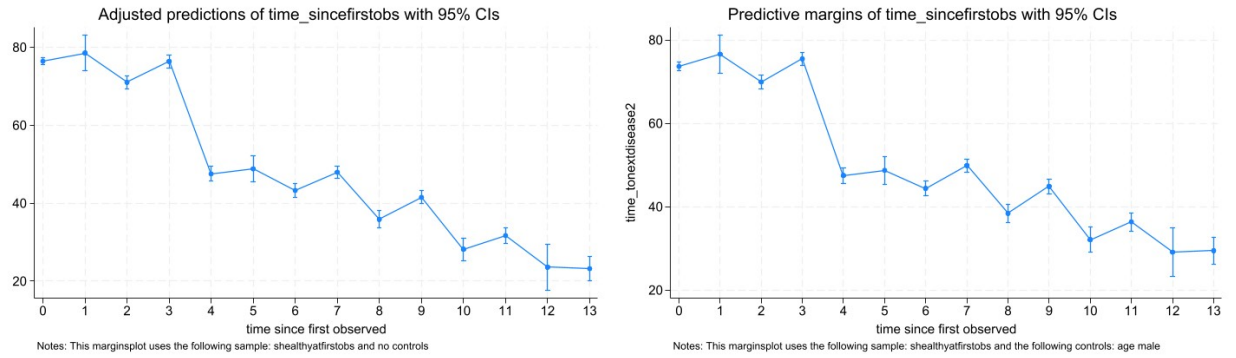
Panel A: Full Sample



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	10,255 0.57	13,855 0.57	16,084 0.60	16,391 0.59	16,936 0.59	17,504 0.60	15,421 0.61
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	10,238 0.57	13,570 0.65	14,945 0.73	14,706 0.79	14,447 0.83	13,756 0.86	10,388 0.88
# diseases	10,234 0.93 (1.06)	13,554 1.20 (1.21)	14,936 1.44 (1.29)	14,695 1.71 (1.40)	14,435 1.96 (1.47)	13,722 2.16 (1.52)	10,373 2.32 (1.55)
age of first onset (observed)	9,062 62.21 (6.61)	12,273 61.94 (6.55)	14,185 61.99 (6.44)	14,410 62.04 (6.46)	14,941 62.03 (6.40)	15,367 62.03 (6.40)	13,494 62.21 (6.34)
age of first onset (g2aging)	5,633 47.49 (12.87)	7,684 47.47 (12.70)	9,219 46.94 (13.18)	9,203 47.21 (13.16)	9,646 47.28 (13.06)	9,941 47.23 (13.07)	8,885 47.10 (13.27)
years first onset to death (observed)	1,847 7.62 (4.47)	2,430 7.30 (4.35)	3,028 6.40 (4.08)	2,750 6.71 (4.21)	3,112 6.52 (4.11)	3,631 6.31 (4.11)	3,830 6.30 (4.10)
months (observed) 1 to 2+ diseases	6,839 38.30 (47.19)	9,283 36.22 (45.06)	10,647 32.91 (42.78)	10,763 33.91 (43.34)	11,217 33.02 (42.81)	11,515 32.01 (42.71)	10,108 31.23 (43.22)
months (observed) 2 to 3+ diseases	4,352 36.98 (43.42)	5,935 35.43 (41.47)	6,746 32.38 (39.15)	6,851 33.12 (39.67)	7,171 32.54 (39.29)	7,362 31.94 (39.22)	6,456 31.42 (39.73)
months (observed) 3 to 4+ diseases	2,333 35.15 (40.92)	3,190 34.34 (39.92)	3,567 31.80 (37.47)	3,639 32.22 (37.84)	3,838 32.09 (37.96)	3,932 31.49 (37.96)	3,436 30.79 (38.00)
months (observed) 4 to 5+ diseases	987 30.46 (36.18)	1,342 30.86 (35.44)	1,486 29.83 (33.91)	1,530 29.83 (34.06)	1,623 29.18 (33.76)	1,649 29.00 (34.15)	1,460 29.04 (34.59)
months (observed) 5 to 6+ diseases	324 28.98 (34.94)	426 26.84 (32.85)	461 26.96 (31.26)	486 26.79 (31.65)	504 27.19 (31.54)	516 26.33 (31.50)	445 26.33 (31.85)
months (observed) 6 to 7+ diseases	65 28.58 (32.47)	92 28.75 (30.44)	92 27.82 (28.65)	103 28.41 (28.13)	106 28.49 (28.35)	110 26.47 (27.37)	94 29.13 (29.68)
months (observed) 7 to 8+ diseases	7 37.57 (24.65)	9 41.22 (22.90)	12 30.92 (27.00)	10 30.90 (26.02)	11 33.73 (26.41)	12 30.92 (27.00)	9 27.44 (23.74)
time_tonextdisease2	7,200 79.11 (45.53)	8,601 78.37 (32.00)	8,066 49.37 (27.13)	6,323 44.00 (22.01)	4,453 36.41 (15.39)	2,101 31.42 (1.89)	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_hlbp	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_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)
radiaghlbp: 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)
radiaglung: 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

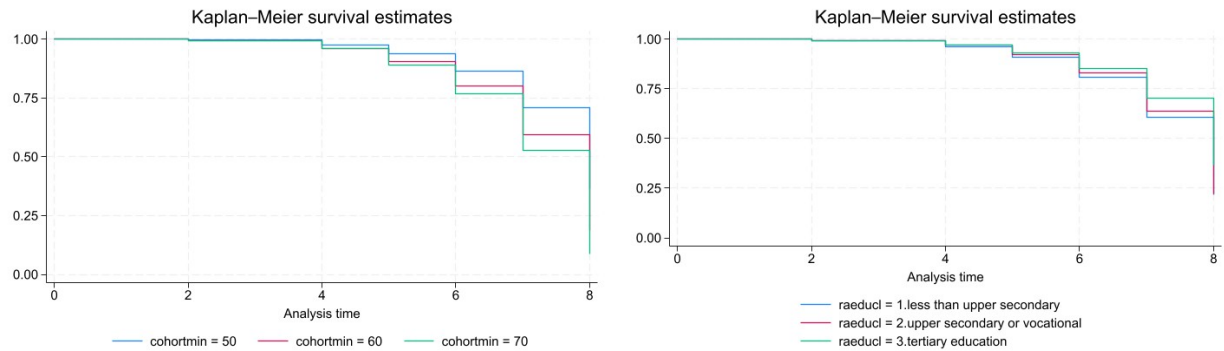
	2004 wave	2006/07 wave	2011/12 wave	2013 wave	2015 wave	2017 wave	2019/20 wave
	4,490 (10.3%)	5,949 (13.6%)	6,520 (14.9%)	6,811 (15.6%)	6,875 (15.7%)	7,043 (16.1%)	6,034 (13.8%)
age at interview	4,490 57.70 (6.27)	5,891 59.02 (6.49)	6,225 62.22 (6.60)	6,334 64.24 (6.56)	6,173 66.02 (6.54)	6,001 67.83 (6.59)	4,657 70.34 (6.46)
r interview year	4,490 2004.33	5,891 2006.92	6,225 2011.01	6,334 2013.00	6,173 2015.00	6,001 2017.00	4,657 2019.48
rubyyear: r birth year	4,490 1946.26	5,949 1947.31	6,520 1948.00	6,811 1947.87	6,875 1947.97	7,043 1948.11	6,034 1948.06
rubyyear: r death year	618 2013.17	784 2013.57	893 2013.81	856 2013.69	961 2013.72	1,158 2014.11	1,198 2014.19
dead	4,490 0.00	5,949 0.01	6,520 0.05	6,811 0.07	6,875 0.10	7,043 0.15	6,034 0.23
already has disease at baseline	4,434 0.00	5,906 0.00	6,489 0.00	6,794 0.00	6,863 0.00	6,982 0.00	6,017 0.00
male	4,490 0.48	5,949 0.47	6,520 0.47	6,811 0.47	6,875 0.47	7,043 0.48	6,034 0.49
any disease	4,417 0.00	5,842 0.19	6,183 0.36	6,321 0.51	6,170 0.62	5,949 0.69	4,651 0.74
# diseases	4,413 0.00 (0.00)	5,834 0.24 (0.56)	6,182 0.52 (0.83)	6,320 0.82 (1.01)	6,168 1.07 (1.14)	5,932 1.30 (1.24)	4,646 1.49 (1.29)
age of first onset (observed)	3,241 64.93 (6.83)	4,324 64.48 (6.73)	4,590 64.40 (6.72)	4,813 64.48 (6.70)	4,868 64.44 (6.63)	4,845 64.44 (6.65)	4,090 64.62 (6.64)
age of first onset (g2aging)	156 52.85 (12.33)	195 52.96 (11.54)	199 52.25 (13.08)	217 52.89 (12.60)	208 52.40 (12.79)	211 52.39 (12.86)	179 52.03 (13.06)
years first onset to death (observed)	337 5.69 (3.68)	443 5.31 (3.48)	455 4.83 (3.40)	462 5.00 (3.45)	504 4.75 (3.32)	543 4.71 (3.31)	554 4.79 (3.34)
months (observed) 1 to 2+ diseases	1,965 38.49 (39.48)	2,577 35.20 (37.37)	2,660 34.00 (36.22)	2,794 33.80 (36.14)	2,840 32.89 (35.35)	2,834 32.60 (35.62)	2,386 34.17 (37.36)
months (observed) 2 to 3+ diseases	963 30.73 (33.93)	1,238 29.64 (32.20)	1,239 28.78 (31.79)	1,321 28.75 (31.66)	1,344 27.63 (30.72)	1,348 27.73 (30.82)	1,139 29.09 (32.21)
months (observed) 3 to 4+ diseases	364 23.22 (27.73)	461 22.10 (26.35)	439 22.05 (25.81)	482 22.40 (26.48)	493 22.18 (26.00)	496 22.25 (26.28)	416 22.51 (26.66)
months (observed) 4 to 5+ diseases	132 23.20 (27.81)	157 21.53 (27.59)	154 21.16 (27.39)	166 21.99 (27.40)	168 20.84 (26.68)	169 21.17 (27.36)	139 21.04 (28.04)
months (observed) 5 to 6+ diseases	35 19.26 (22.62)	39 20.03 (22.55)	35 18.11 (21.44)	41 19.83 (22.29)	41 19.83 (22.29)	41 18.46 (21.72)	30 19.97 (23.65)
months (observed) 6 to 7+ diseases	10 10.60 (13.99)	10 13.90 (15.06)	8 9.12 (12.70)	10 13.90 (15.06)	10 13.90 (15.06)	11 12.64 (14.89)	6 19.67 (15.56)
months (observed) 7 to 8+ diseases	0 . (.)	0 . (.)	0 . (.)	0 . (.)	0 . (.)	0 . (.)	0 . (.)
time_tonextdisease2	3,222 82.05 (46.18)	3,902 79.17 (32.89)	3,514 51.05 (28.02)	2,866 44.91 (22.37)	2,011 36.54 (15.41)	961 31.41 (1.94)	0 . (.)
r educ							
1.less than upper secondary	1,378 (38.7%)	2,206 (37.1%)	2,261 (34.7%)	2,363 (34.7%)	2,380 (34.6%)	2,478 (35.2%)	2,049 (34.0%)
2.upper secondary or vocational	1,519 (33.8%)	2,177 (36.6%)	2,568 (39.4%)	2,608 (38.3%)	2,706 (39.4%)	2,772 (39.4%)	2,430 (40.3%)
3.tertiary education	1,233 (27.5%)	1,566 (26.3%)	1,691 (25.9%)	1,840 (27.0%)	1,789 (26.0%)	1,793 (25.5%)	1,555 (25.8%)
cohort							
50-59	2,515 (59.4%)	3,022 (53.4%)	2,224 (36.3%)	1,575 (25.1%)	916 (14.9%)	429 (7.2%)	137 (2.9%)
60-69	1,615 (38.1%)	2,265 (40.0%)	2,901 (47.3%)	3,242 (51.7%)	3,380 (55.1%)	3,233 (54.1%)	2,094 (45.1%)
70-79	107 (2.5%)	377 (6.7%)	1,004 (16.4%)	1,458 (23.2%)	1,739 (28.3%)	2,045 (34.2%)	1,967 (42.3%)
80+	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)	100 (1.6%)	267 (4.5%)	447 (9.6%)

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