SUPPLEMENTAL MATERIAL

Supplemental Figure Legend

Supplemental Figure 1: Lexis Diagram illustrating the two-time scales involved in a design with person with diabetes together with matched controls

Supplemental figure 2: The hazard as a function of age

Supplemental Figure 3. Adjusted hazard ratios (95% CI) for patients with type 2 diabetes according to age at diagnosis, compared with matched controls in those without prior cardiovascular disease separately in men (blue) and women (red). All outcomes minus atrial fibrillation, non-acute myocardial infarction coronary heart disease. There was a significant age by sex interaction (p<0.0001) for all outcomes.

Supplemental Table 1. Definition of outcomes ICD-codes

Outcome	ICD 10 codes
AMI	'410' 'I21'
CHD	'410' '411' '412' '413' '414' 'I20' 'I21' 'I22' 'I23' 'I24' 'I25'
Stroke	'431' '432' '433' '434' '436' 'I61' 'I62' 'I63' 'I64'
CVD	'410' '431' '432' '433' '434' '436' 'I21' 'I61' 'I62' 'I63' 'I64'
Atrial fibrillation	'427D' 'I48'
Heart Failure	'428' 'I50'

ICD: International Classification of Diseases; AMI: acute myocardial infarction; CHD: coronary heart

disease; CVD: cardiovascular disease

Supplemental Table 2. Descriptive statistics at diagnosis for T2DM with less than 10 years duration and matched controls (entire cohort)

	Control	T2DM	p	SMD
n	1,575,108	318,038		
Male (%)	869300 (55·2)	175553 (55·2)	0·928	<0·001
Age (mean (sd))	63·74 (12·35)	63·87 (12·40)	<0·001	0·011
Education (%)			<0·001	0·221
College level	625910 (40·4)	129686 (41·6)		
Elementary school	540964 (34·9)	131141 (42·1)		
Upper secondary school	383752 (24·7)	50893 (16·3)		
Marital status (%)			<0·001	0·051
Married	884490 (56·2)	170827 (53·7)		
Separated	253722 (16·1)	53510 (16·8)		
Single	247584 (15·7)	51933 (16·3)		
Widowed	189237 (12·0)	41768 (13·1)		
Income (mean (sd))	2017 (4504)	1731 (2388)	<0·001	0·079
Country of origin (%)			<0·001	0·190
EU	40568 (2·6)	7685 (2·4)		
Nordic	75929 (4·8)	16827 (5·3)		
RoW	78387 (5·0)	31321 (9·8)		
Sweden	1380138 (87·6)	262197 (82·4)		
Previous CVD (%)	109880 (7·0)	41235 (13·0)	<0·001	0·201
Previous CHD (%)	122769 (7·8)	49265 (15·5)	<0·001	0·242
Previous AMI (%)	60748 (3·9)	26225 (8·2)	<0·001	0·185
Previous Stroke (%)	56002 (3·6)	17671 (5·6)	<0·001	0·096
Previous Renal (%)	1895 (0·1)	529 (0·2)	<0·001	0·012
Previous HF (%)	40409 (2·6)	17794 (5·6)	<0·001	0·154
Previous AF (%)	65942 (4·2)	22184 (7·0)	<0·001	0·122
Previous Amputation (%)	1242 (0·1)	611 (0·2)	<0·001	0·031
Previous Dementia (%)	11339 (0·7)	954 (0·3)	<0·001	0·059
Previous Cancer (%)	99414 (6·3)	22323 (7·0)	<0·001	0·028

T2DM: type 2 diabetes mellitus; SMD: Standardised mean difference; SD: standard deviation; EU: European Union; RoW: Rest of World; CVD: cardiovascular disease; CHD: coronary heart disease; AMI: acute myocardial infarction; HF: heart failure; AF: atrial fibrillation

Supplemental Table 3. Baseline Characteristics of Individuals with Type 2 Diabetes, According to Age at Diagnosis (entire cohort)

	Age at diagnosis of type 2 diabetes						
	40 or less	41 to 50	51 to 60	61 to 70	71 to 80	81 to 90	91 or older
n	7373	20816	40041	49866	28061	10010	538
Female (%)	3145	7660	15241	21303	16188	6179	353
	(42·7)	(36·8)	(38·1)	(42·7)	(57·7)	(61·7)	(65·6)
Age – mean (SD)	35·06	46·53	56·40	65·63	75·32	84·45	92·93
	(5·36)	(2·84)	(2·89)	(2·83)	(2·87)	(2·54)	(1·78)
HbA$_{1c}$ – mean (SD)	56·13	55·80	54·14	52·02	50·71	51·46	53·71
	(19·36)	(18·38)	(17·22)	(15·42)	(13·42)	(13·26)	(14·34)
Systolic BP – mean (SD)	127·13	131·53	136·37	139·38	142·11	142·86	144·23
	(15·12)	(15·83)	(16·58)	(17·06)	(18·06)	(19·38)	(19·84)
Diastolic BP – mean (SD)	79·45	81·56	81·54	79·46	76·88	75·13	74·31
	(10·50)	(10·19)	(9·62)	(9·43)	(9·65)	(9·99)	(9·81)
Triglycerides – mean (SD)	2·33	2·28	2·10	1·90	1·75	1·74	1·68
	(1·78)	(1·69)	(1·42)	(1·14)	(0·91)	(0·86)	(0·89)
Antihypertensives – n (%)	0·18	7910	22422	34407	21326	7908	425
	(39)	(38)	(56)	(69)	(76)	(79)	(79)
BMI – mean (SD)	33·62	32·02	30·80	30·11	28·93	27·34	25·96
	(7·50)	(6·35)	(5·46)	(5·19)	(4·82)	(4·42)	(4·10)
LDL cholesterol – mean (SD)	3·06	3·16	3·17	3·07	2·96	2·96	2·92
	(0·92)	(0·97)	(1·00)	(0·99)	(0·98)	(0·97)	(1·01)
HDL cholesterol – mean (SD)	1·06	1·14	1·22	1·28	1·33	1·33	1·38
	(0·32)	(0·35)	(0·37)	(0·39)	(0·40)	(0·40)	(0·56)
Total cholesterol – mean (SD)	5·14	5·29	5·31	5·19	5·08	5·09	5·13
	(1·18)	(1·17)	(1·18)	(1·13)	(1·12)	(1·13)	(1·22)
Statins – n (%)	0·13	5620	14815	22439	12066	3003	54
	(34)	(27)	(37)	(45)	(43)	(30)	(10)

Estimated GFR – mean (SD)[1]	107·31 (28·11)	97·72 (27·25)	90·10 (22·08)	82·21 (21·26)	71·88 (20·00)	63·14 (18·82)	57·40 (19·49)
Smokers – n (%)	1245 (22·0)	4273 (25·3)	7769 (23·5)	6595 (16·1)	1939 (8·3)	295 (3·6)	4 (1·0)
Physical activity[2] – n (%)							
No physical activity	485 (12·4)	1484 (12·4)	2617 (11·3)	3243 (10·7)	2345 (14·7)	1444 (26·2)	130 (45·6)
Less than once/week	542 (13·8)	1627 (13·6)	2906 (12·6)	3247 (10·7)	1800 (11·3)	818 (14·8)	36 (12·6)
1–2 times/week	896 (22·9)	2612 (21·8)	5061 (21·9)	5971 (19·7)	3135 (19·6)	1070 (19·4)	37 (13·0)
3–5 times/week	1000 (25·5)	3009 (25·2)	5848 (25·3)	7334 (24·3)	3465 (21·7)	919 (16·7)	41 (14·4)
Daily physical activity	996 (25·4)	3232 (27·0)	6714 (29·0)	10442 (34·5)	5212 (32·7)	1262 (22·9)	41 (14·4)

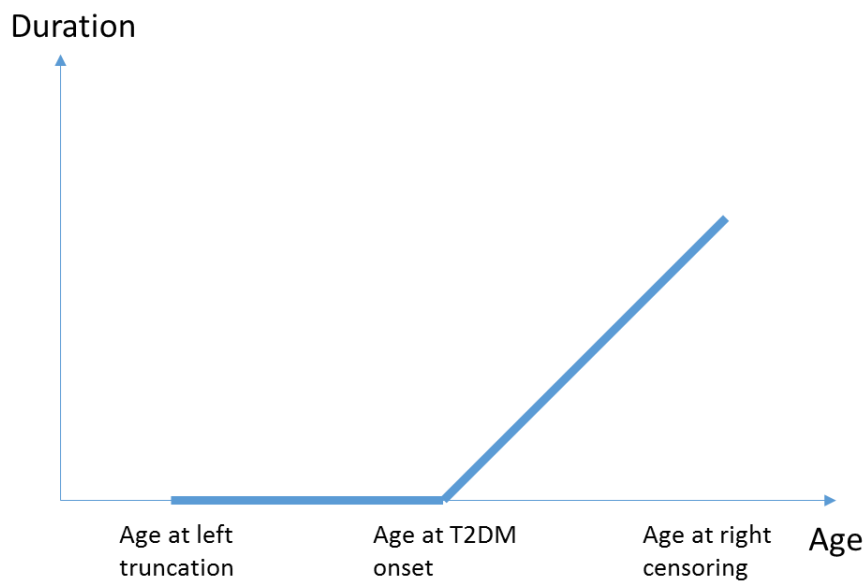SD: standard deviation; HbA1c: hemoglobin A1C; BP: blood pressure; BMI: body mass index; LDL: low-density lipoproteins; HDL: high-density lipoproteins; GFR: glomerular filtration rate

Supplemental Table 4. *Median survival time by age at diagnosis for T2DM and corresponding matched controls*

Age-group	Group	median	lower	upper
0 to 20	T2DM	75.76660	74.59491	78.35238
	Control	87.61522	85.72131	89.12252
21 to 30	T2DM	78.42790	76.56947	80.36083
	Control	85.19918	84.11214	86.67750
31 to 40	T2DM	78.48106	77.79671	79.37931
	Control	85.93635	85.39699	86.62891
41 to 50	T2DM	79.29763	79.02875	79.58031
	Control	85.54152	85.33949	85.75017
51 to 60	T2DM	81.98973	81.81793	82.14636
	Control	85.90349	85.81805	85.99863
61 to 70	T2DM	84.70226	84.57963	84.82341
	Control	86.73317	86.66872	86.80151
71 to 80	T2DM	87.68047	87.57278	87.79204
	Control	88.24013	88.18743	88.29489
81 to 90	T2DM	91.70922	91.56594	91.84725
	Control	91.48038	91.41056	91.55362
91 or older	T2DM	97.91170	97.52361	98.31348
	Control	97.49361	97.30036	97.71458

Supplemental statistical analyses: the statistical model

The statistical model we use differs slightly from the standard Cox proportional hazards model and the reason for this is the inclusion of the diabetes free controls in the design. We use age as the underlying time scale which allows the baseline hazard to capture the increase in hazard due to aging. To allow persons with diabetes to age at a different rate than the diabetes free controls we need to include duration of diabetes as a second time scale. The duration of diabetes is only defined for persons with diabetes and we need to find a way to handle that in the statistical model. There have been several attempts at this in prior papers including allocating the value of duration for a specific person with diabetes at the index date to the corresponding controls or stratifying on duration of diabetes for the persons with diabetes and letting the controls follow the corresponding person with diabetes. We have developed a more robust statistical method to model and therefore account for the effect of duration.

Duration

Supplemental Figure 1: Lexis Diagram illustrating the two-time scales involved in a design with person with diabetes together with matched controls

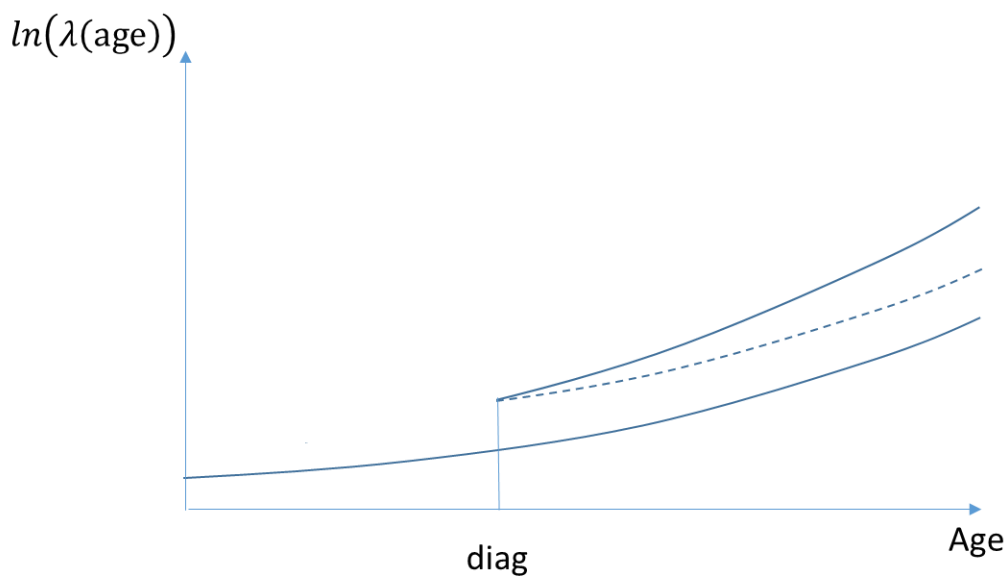T2DM: type 2 diabetes mellitus.

As an illustration let's look at the hazard function for a person using age as the underlying time scale. Let the hazard due to aging while diabetes free be denoted $\lambda_0(t)$ and that at some age t_d the patient is diagnosed with diabetes. At this point life becomes more hazardous which we model by adding α to the log hazard. For every

year beyond the age of diagnosis the patient now ages at a different rate compared to a diabetes free person which we represent by adding β to the log hazard for every year the patient lived with the disease. Mathematically the model looks like

$$\lambda(t) = \lambda_0(t)exp\big(\alpha I_{\{t>t_d\}} + (t-t_d)\beta I_{\{t>t_d\}}\big) =$$

$$\lambda_0(t)exp\big(\alpha I_{\{t>t_d\}} + D(t)\beta\big),$$

where $D(t)= (t-t_d)I_{\{t>t_d\}}$ is simply the duration at age t for patients diagnosed with diabetes and zero for those not diagnosed (including the controls) which can be illustrated as in figure 2 where the dotted line represents aging for a person with diabetes in absence of any effect of duration. In the design with controls match to diabetic patients based on age and sex we can view the controls as patients who have not yet been diagnosed with diabetes and therefore $t < t_d$ during the entire time we follow these patients and thus they get no contribution from the term $D(t)= (t-t_d)I_{\{t>t_d\}}$. It is practical to code for this in the data set by simply setting duration to zero.

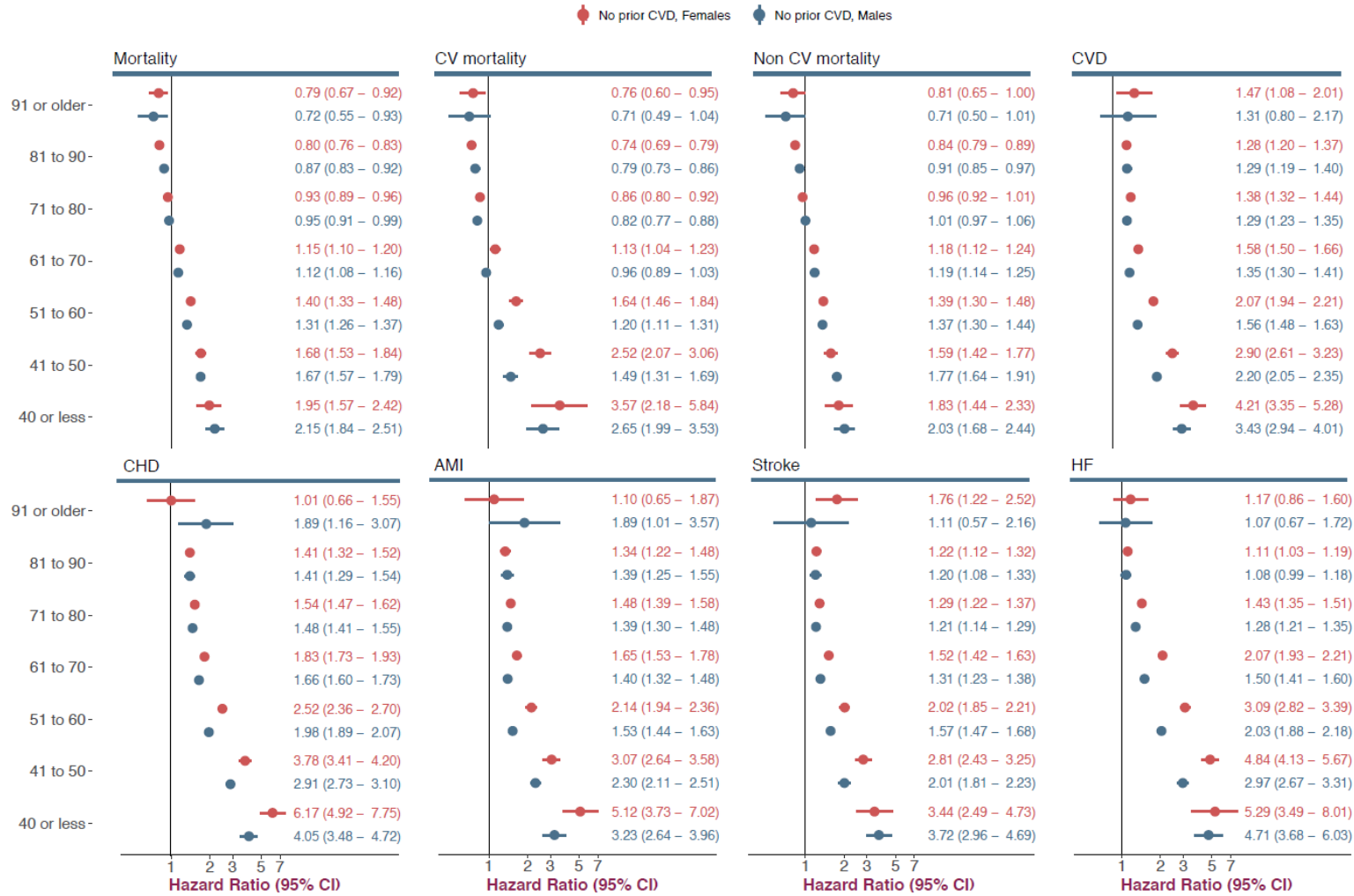**Supplemental figure 2: The hazard as a function of age**

diag: diagnosis.

Supplemental Table 5. Comorbidities by age at diagnosis DM

	Age Group													
	40 or less		41 to 50		51 to 60		61 to 70		71 to 80		81 to 90		91 or older	
	T2DM	Controls	T2DM	Controls	T2DM	Controls	T2DM	Controls	T2DM	Controls	T2DM	Controls	T2DM	Controls
n	15209	77520	40873	219128	71832	405423	69970	408513	34554	203600	8520	47702	320	1726
Previous Renal	12	33	68	134	103	294	88	358	28	169	1	27	0	1
(%)	(0·1)	(0·0)	(0·2)	(0·1)	(0·1)	(0·1)	(0·1)	(0·1)	(0·1)	(0·1)	(0·0)	(0·1)	(0·0)	(0·1)
Previous	2	5	27	40	97	84	99	159	59	148	26	109	2	5
Amputation (%)	(0·0)	(0·0)	(0·1)	(0·0)	(0·1)	(0·0)	(0·1)	(0·0)	(0·2)	(0·1)	(0·3)	(0·2)	(0·6)	(0·3)
Previous	3	2	6	25	37	212	89	1116	161	2825	85	1910	7	122
Dementia (%)	(0·0)	(0·0)	(0·0)	(0·0)	(0·1)	(0·1)	(0·1)	(0·3)	(0·5)	(1·4)	(1·0)	(4·0)	(2·2)	(7·1)
Previous Cancer	166	502	874	3419	3293	15456	5789	30636	3779	21023	1047	5561	40	199
(%)	(1·1)	(0·6)	(2·1)	(1·6)	(4·6)	(3·8)	(8·3)	(7·5)	(10·9)	(10·3)	(12·3)	(11·7)	(12·5)	(11·5)

T2DM: type 2 diabetes mellitus

Supplemental Figure 3.

Supplemental Figure 3. Adjusted hazard ratios (95% CI) for patients with type 2 diabetes according to age at diagnosis, compared with matched controls in those without prior cardiovascular disease separately in men (blue) and women (red). All outcomes minus atrial fibrillation, non-acute myocardial infarction coronary heart disease. There was a significant age by sex interaction (p<0.0001) for all outcomes.

CVD: cardiovascular disease; CV: cardiovascular; CHD: coronary heart disease; AMI: acute myocardial infarction; HF: heart failure; CI: confidence interval.