Supplementary Appendix: Authors: Zöller B, Pirouzifard M, Holmquist B, Sundquist J, Halling A, Sundquist K

Table of content pages 1-32

Zöller B, *et al. BMJMED* 2023; 2:e000070. doi: 10.1136/bmjmed-2021-000070

Supplementary Table S1. Diseases with ICD codes included in the multimorbidity score modified and adapted to Swedish ICD-10 codes after Barnett K et al (2012)*. Infectious diseases, i.e. viral hepatitis, were excluded and six common non-communicable diseases (NCDs) were added: dermatitis or eczema, gout, arthrosis, osteoporosis, obesity, and pancreas diseases. One point for each disease. The theoretical range of the score is 0-45 points.

	Included disorders in the multimorbidity score	ICD-10**
1	Hypertension	I10-I15
2	Affective disorders	F30-F39
3	Painful back condition	M50-54
4	Asthma	J45-J46
5	CHD	I20-I25
6	Ulcer disease	K221, K25-K28
7	Diabetes	E10-E14
8	Thyroid disorders	E00-E07
9	Connective tissue disease	M05, M06, M08, M09, M30-M36, D86
10	Hearing loss (or impaired)	H90-H91
11	Chronic obstructive pulmonary disease (COPD)	J41-J44
12	Anxiety	F40-F48
13	Irritable bowel syndrome (IBS)	K58
14	Cancer	C00-C99
15	Alcohol use disorders	F10
16	Psychoactive substance misuse	F11-F14, F16, F18, F19
17	Constipation	K590
18	Cerebrovascular disease	I60-I69
19	Renal disease	N00-N19, Q61
20	Diverticular disease of intestine	K57
21	Atrial fibrillation	I48

22	Atherosclerosis	I70-I74, I77
23	Heart failure	I50, I110, I130, I132
24	Prostate disease	N40-N42
25	Glaucoma	H40, H42
26	Epilepsy	G40, G41
27	Dementia	F00-F03
28	Schizophrenia disorders	F20, F21
29	Psoriasis	L40
30	Dermatitis or Eczema	L20-L30
31	Inflammatory bowel disease (IBD)	K50, K51
32	Migraine	G43, G440, G441
33	Blindness & low vision	H53-H54
34	Chronic sinusitis	J32
35	Learning disability	F81
36	Anorexia or bulimia	F50
37	Bronchiectasis	J47
38	Parkinson's disease	G20, G21, G22
39	Multiple sclerosis	G35
40	Liver disease	K70-K77
41	Gout	M10
42	Arthrosis	M15-M19
43	Osteoporosis	M80-M82
44	Obesity	E65, E66
45	Pancreas diseases	K85, K86
*Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012; 380: 37-43. **ICD=International classification of disease.		

Supplementary Table S2. Distribution of multimorbidity scores in relatives (twins, siblings, half-siblings, cousins). One point for every disease. Theoretical range is 0-45. Median score =0, range 0-20.

	Twins		Siblings		Half-Siblings		Cousins	
Score	n	%	n	%	n	%		
0	14652	61.00	938702	60.71	134405	54.58	621041	59.46
1	5838	24.30	365498	23.64	60336	24.50	246478	23.60
2	2153	8.96	142931	9.24	27262	11.07	99744	9.55
3	807	3.36	56830	3.68	12752	5.18	42464	4.07
4	339	1.41	23956	1.55	6117	2.48	19037	1.82
5	128	0.53	10142	0.66	2819	1.14	8410	0.81
6	49	0.20	4294	0.28	1318	0.54	3840	0.37
7	34	0.14	1992	0.13	646	0.26	1790	0.17
8	11	0.05	916	0.06	286	0.12	852	0.08
9	4	0.02	444	0.03	155	0.06	440	0.04
10	2	0.01	193	0.01	68	0.03	211	0.02
11	0	0.00	114	0.01	39	0.02	112	0.01
12	1	0.00	55	0.00	23	0.01	74	0.01
13	0	0.00	24	0.00	9	0.00	27	0.00
14	0	0.00	5	0.00	3	0.00	12	0.00
15	0	0.00	5	0.00	3	0.00	7	0.00
16	1	0.00	4	0.00	2	0.00	4	0.00
17	0	0.00	1	0.00	1	0.00	1	0.00
18	0	0.00	1	0.00	0	0.00	0	0.00
19	0	0.00	1	0.00	0	0.00	0	0.00
20	0	0.00	0	0.00	0	0.00	2	0.00

*One point for every disease in patient. Included diseases in score: Hypertension, Heart failure, coronary heart disease (CHD), Diabetes, Obesity, Atrial fibrillation, Gout, Atherosclerosis, Renal disease, Affective disorders, Anxiety, Psychoactive substance misuse, Alcohol use disorders, Anorexia or bulimia, Schizophrenia disorders, Inflammatory bowel disease (IBD), Liver disease, Pancreas

diseases, Ulcer disease, Epilepsy, Blindness & low vision, Cerebrovascular disease, Cancer, Hearing loss (or impaired), Connective tissue disease, Osteoporosis, Thyroid disorders, Psoriasis, Prostate disease, Arthrosis, Painful back condition, Diverticular disease of intestine, Chronic sinusitis, Bronchiectasis, Parkinson disease, Glaucoma, Learning disability, Irritable bowel syndrome (IBS), Asthma, Dermatitis and Eczema, Constipation, Chronic obstructive pulmonary disease (COPD), Migraine, Multiple sclerosis and Dementia.

Supplementary Table S3. Distribution of participants according to county.

County	Number of individuals	Percentage of total
1	434700	16.13
2	33223	1.23
3	89055	3.31
4	76406	2.84
5	134927	5.01
6	101967	3.78
7	57382	2.13
8	80311	2.98
9	20357	0.76
10	51700	1.92
11	79884	2.96
12	244757	9.08
13	79637	2.96
14	260384	9.66
15	117344	4.36

16	82428	3.06
17	88541	3.29
18	86066	3.19
19	78912	2.93
20	91167	3.38
21	95137	3.53
22	87241	3.24
23	42654	1.58
24	90520	3.36
25	89742	3.33

Supplementary Table S4. Descriptive findings for twins, siblings, half-siblings, and cousins stratified according to multimorbidity scores for number of individuals, sex, education, age at the end of study, and birth date.

	Multimorbidity scores								
	All	**0**	**1**	**2**	**3**	**4**	**≥ 5**	**≤ 1**	**≥ 2**
Twins n (%)	24020	14652 (61.00)	5838 (24.30)	2153 (8.96)	807 (3.36)	339 (1.41)	231 (0.96)	20490 (85.30)	3530 (14.70)
Sex, Female, n(%)	12003 (49.97)	7029 (47.97)	2951 (50.55)	1188 (55.18)	489 (60.59)	205 (60.47)	141 (61.04)	9980 (48.71)	2023 (57.31)
Education, >11 years; n(%)	4940 (20.57)	3236 (22.09)	1101 (18.86)	369 (17.14)	142 (17.60)	59 (17.40)	33 (14.29)	4337 (21.17)	603 (17.08)
Age at end of study, Median (IQR) Range (min-max)	22 (16-34) (0 -64)	22 (16-33) (0 -64)	22 (15-33) (1 -62)	23 (16-35) (9 -60)	25 (18-38) (10-64)	28 (21-42) (1 -59)	36 (24-47) (6 -62)	22 (16-33) (0 -64)	24 (17-37) (1 -64)
Year of birth, Median (IQR) Range (min-max)	1992 (1981-1999) (1951-2005)	1992 (1981-1999) (1951-2005)	1993 (1981-2000) (1953-2005)	1992 (1980-1999) (1953-2005)	1990 (1976-1997) (1951-2005)	1986 (1973-1994) (1953-2005)	1979 (1968-1991) (1953-2004)	1993 (1981-1999) (1951-2005)	1991 (1977-1998) (1951-2005)
Sibling, (n (%)	1546108	938702 (60.71)	365498 (23.64)	142931 (9.24)	56830 (3.68)	23956 (1.55)	18191 (1.18)	1304200 (84.35)	241908 (15.65)
Sex, Female, n(%)	755981 (48.90)	423077 (45.55)	186466 (51.61)	78875 (55.89)	33263 (59.34)	14423 (61.16)	11704 (64.34)	616009 (47.23)	139972 (57.86)
Education, >11 years; n(%)	470199 (30.76)	298311 (32.12)	108534 (30.04)	39295 (27.84)	14608 (26.06)	5641 (23.92)	3810 (21.32)	411972 (31.59)	64308 (26.58)

Age at end of study, Median (IQR) Range (min-max)	31 (21-42) (0 -67)	30 (21-41) (0 -66)	31 (21-43) (0 -67)	32 (22-44) (0 -66)	35 (24-45) (1 -66)	37 (26-47) (6 -65)	41 (30-49) (6 -65)	30 (21-42) (0 -67)	34 (23-45) (0 -66)
Year of birth, Median (IQR) Range (min-max)	1984 (1973-1993) (1948-2005)	1985 (1973-1994) (1949-2005)	1984 (1972-1994) (1948-2005)	1982 (1971-1993) (1949-2005)	1980 (1969-1991) (1949-2005)	1978 (1968-1989) (1950-2005)	1974 (1966-1985) (1950-2005)	1984 (1973-1994) (1948-2005)	1981 (1970-1992) (1949-2005)
Half sibling	246244	134405 (54.58)	60336 (24.50)	27262 (11.07)	12752 (5.18)	6117 (2.48)	5372 (2.18)	194741 (79.08)	51503 (20.92)
Sex, Female, n(%)	120574 (48.97)	59922 (44.58)	30760 (50.98)	15175 (55.66)	7546 (59.18)	3740 (61.14)	3431 (63.87)	90682 (46.57)	29892 (58.04)
Education, >11 years; n(%)	54819 (22.26)	32349 (24.07)	13214 (21.90)	5431 (19.92)	2213 (17.35)	954 (15.60)	658 (12.25)	45563 (23.40)	9256 (17.97)
Age at end of study, Median (IQR) Range (min-max)	31 (23-41) (0 -67)	30 (22-40) (0 -67)	31 (22-41) (0 -67)	32 (24-43) (1 -66)	34 (25-45) (5 -66)	36 (27-47) (4 -65)	40 (30-50) (7 -66)	30 (22-40) (0 -67)	34 (25-45) (1 -66)
Year of birth, median (IQR) Range (min-max)	1984 (1973-1992) (1948-2005)	1985 (1975-1993) (1948-2005)	1984 (1973-1992) (1948-2005)	1983 (1972-1991) (1949-2005)	1981 (1970-1990) (1949-2005)	1979 (1968-1988) (1950-2005)	1975 (1965-1984) (1949-2005)	1985 (1974-1993) (1948-2005)	1981 (1970-1990) (1949-2005)
Cousins	1044546	621041 (59.46)	246478 (23.60)	99744 (9.55)	42464 (4.07)	19037 (1.82)	15782 (1.51)	867519 (83.05)	177027 (16.95)
Sex, Female, n(%)	505755 (48.42)	278051 (44.77)	126265 (51.23)	55215 (55.36)	24958 (58.77)	11427 (60.03)	9839 (62.34)	404316 (46.61)	101439 (57.30)
Education, >11 years; n(%)	320112 (30.65)	202015 (32.53)	73937 (30.00)	26999 (27.07)	10217 (24.06)	4088 (21.47)	2856 (18.10)	275952 (31.81)	44160 (24.95)
Age at end of study, Median (IQR) Range (min-max)	33 (24-44) (0 -68)	32 (23-42) (0 -66)	34 (24-44) (0 -67)	35 (25-46) (0 -67)	37 (27-48) (1 -66)	40 (29-49) (1 -67)	44 (33-52) (3 -68)	32 (23-43) (0 -67)	37 (26-48) (0 -68)

Year of birth, median (IQR) Range (min-max)	1982 (1971-1991) (1947-2005)	1983 (1972-1992) (1949-2005)	1981 (1970-1991) (1948-2005)	1980 (1969-1990) (1948-2005)	1978 (1967-1988) (1949-2005)	1975 (1965-1986) (1948-2005)	1971 (1963-1982) (1947-2005)	1982 (1972-1991) (1948-2005)	1978 (1967-1989) (1947-2005)
IQR=Interquartile Range.									

Supplementary Table S5. Stratified ORs according to sex and age among twins.

	OR (95%CI)	
	Crude	Adjusted*
Twins Stratified for Sex		
Male -Male	3.65 (2.92-4.56)	3.44 (2.74-4.32)
Female – Female	3.82 (3.17-4.61)	3.60 (2.97-4.37)
Male – Female	2.01 (1.63-2.49)	1.92 (1.55-2.39)
Female -Male	2.01 (1.63-2.49)	1.91 (1.53-2.37)
Twins stratified for Age at end of study		
Age <= 20	3.27 (2.69-3.98)	3.22 (2.65-3.92)
20 < Age <= 30	3.33(2.66-4.15)	3.14 (2.50-3.93)
Age > 30	2.58 (2.12-3.15)	2.29 (1.87-2.80)

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S6. Stratified ORs according to sex and age among siblings.

	OR (95%CI)	
	Crude	Adjusted*
Siblings Stratified for Sex		
Male -Male	2.11 (2.04-2.18)	1.91 (1.84-1.98)
Female – Female	2.00 (1.94-2.05)	1.89 (1.84-1.95)
Male – Female	1.85 (1.81-1.89)	1.74 (1.71-1.78)
Female -Male	1.85 (1.81-1.89)	1.73 (1.69-1.77)
Siblings stratified for Age at end of study		
Age <= 24	2.33 (2.26-2.39)	2.29 (2.23-2.36)
24 < Age <= 39	1.83 (1.79-1.88)	1.73 (1.69-1.78)
Age > 39	1.63 (1.60-1.67)	1.52 (1.48-1.55)

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S7. Stratified ORs according to sex and age among half-siblings.

	OR (95%CI)	
	Crude	Adjusted*
Half-siblings Stratified for Sex		
Male -Male	1.36 (1.31-1.41)	1.27 (1.22-1.32)
Female – Female	1.36 (1.32-1.41)	1.31 (1.27-1.35)
Male – Female	1.30 (1.27-1.33)	1.23 (1.20-1.26)
Female -Male	1.30 (1.27-1.33)	1.24 (1.21-1.27)
Half-siblings stratified for Age at end of study		
Age <= 24	1.37 (1.33-1.41)	1.36 (1.32-1.40)
24 < Age <= 39	1.30 (1.27-1.33)	1.25 (1.22-1.28)
Age > 39	1.25 (1.21-1.28)	1.20 (1.16-1.23)

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S8. Stratified ORs according to sex and age among cousins.

	OR (95%CI)	
	Crude	Adjusted*
Cousins Stratified for Sex		
Male -Male	1.20 (1.18-1.23)	1.13 (1.11-1.15)
Female – Female	1.19 (1.18-1.21)	1.15 (1.13-1.17)
Male – Female	1.18 (1.16-1.19)	1.12 (1.11-1.14)
Female -Male	1.18 (1.16-1.19)	1.12 (1.11-1.14)
Cousins stratified for Age at end of study		
Age <= 24	1.22 (1.20-1.24)	1.21 (1.19-1.24)
24 < Age <= 39	1.15 (1.14-1.17)	1.12 (1.10-1.13)
Age > 39	1.14 (1.12-1.15)	1.10 (1.08-1.11)

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplement Table S9. Result of principal component analysis (PCA) followed by factor analysis. Unrotated Eigenvalues for the principal components factors are show. Number of observations = 1570128. Retained factors were 14 factors beyond 1.

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	8.00168	4.69514	0.1778	0.1778
Factor2	3.30654	0.71200	0.0735	0.2513
Factor3	2.59454	0.35234	0.0577	0.3090
Factor4	2.24220	0.41290	0.0498	0.3588
Factor5	1.82930	0.25639	0.0407	0.3994
Factor6	1.57292	0.09361	0.0350	0.4344
Factor7	1.47931	0.05200	0.0329	0.4673
Factor8	1.42731	0.11487	0.0317	0.4990
Factor9	1.31243	0.11358	0.0292	0.5281
Factor10	1.19885	0.00682	0.0266	0.5548
Factor11	1.19204	0.06940	0.0265	0.5813
Factor12	1.12264	0.09269	0.0249	0.6062
Factor13	1.02995	0.01665	0.0229	0.6291
Factor14	1.01330	0.07554	0.0225	0.6516

Eigenvalue=An eigenvalue is the variance of the factor.
Difference= difference between the current and following eigenvalue.
Proportion= proportion of variance accounted for by the factor.
Cumulative= cumulative proportion of variance accounted for by this factor plus all of the previous ones.

Supplementary Table S10. Distribution of the factor 1-9 (F1-F9) scores in respective relatives.

	Twins		Siblings		Half-Siblings		Cousins	
Score	n	%	n	%	n	%	n	%
F1: Hypertension, Heart failure, CHD, Diabetes, Obesity, Atrial fibrillation, Gout, Atherosclerosis and Renal disease.								
0	22692	94.47	1437411	92.97	223790	90.88	960831	91.99
1	1161	4.83	91401	5.91	18472	7.50	67741	6.49
2	129	0.54	12829	0.83	2865	1.16	11413	1.09
3	32	0.13	3327	0.22	806	0.33	3293	0.32
4	3	0.01	818	0.05	223	0.09	920	0.09
5	2	0.01	239	0.02	63	0.03	229	0.02
6	1	0.00	65	0.00	24	0.01	89	0.01
7	0	0.00	17	0.00	1	0.00	25	0.00
8	0	0.00	1	0.00	0	0.00	5	0.00
F2: Affective disorders, Anxiety, Psychoactive substance misuse, Alcohol use disorders, Anorexia or bulimia, and Schizophrenia disorders.								
0	21461	89.35	1364992	88.29	202473	82.22	909925	87.11
1	1581	6.58	113198	7.32	25236	10.25	82539	7.90
2	671	2.79	48679	3.15	12461	5.06	36894	3.53
3	235	0.98	14503	0.94	4405	1.79	11277	1.08
4	65	0.27	4303	0.28	1534	0.62	3542	0.34
5	7	0.03	421	0.03	131	0.05	357	0.03
6	0	0.00	12	0.00	4	0.00	12	0.00
F3: IBD, Liver disease, Pancreas diseases and Ulcer disease.								
0	23712	98.72	1518543	98.22	240913	97.84	1023304	97.97
1	295	1.23	25962	1.68	4983	2.02	19903	1.91
2	12	0.05	1466	0.09	316	0.13	1249	0.12
3	1	0.00	127	0.01	29	0.01	84	0.01

4	0	0.00	10	0.00	3	0.00	6	0.00
F4: Epilepsy, Blindness & low vision, Cerebrovascular disease, Cancer and Hearing loss (or impaired).								
0	22679	94.42	1463005	94.63	231525	94.02	982870	94.10
1	1236	5.15	77738	5.03	13622	5.53	57396	5.49
2	96	0.40	4833	0.31	994	0.40	3796	0.36
3	9	0.04	487	0.03	99	0.04	450	0.04
4	0	0.00	43	0.00	3	0.00	32	0.00
5	0	0.00	2	0.00	1	0.00	2	0.00
F5: Connective tissue disease, Osteoporosis, Thyroid disorders and Psoriasis.								
0	23212	96.64	1480406	95.75	234321	95.16	996672	95.42
1	774	3.22	62476	4.04	11326	4.60	45455	4.35
2	32	0.13	3098	0.20	571	0.23	2302	0.22
3	2	0.01	127	0.01	26	0.01	116	0.01
4	0	0.00	1	0.00	0	0.00	1	0.00
F6: Prostate disease, Arthrosis, Painful back condition, Diverticular disease of intestine, and Chronic sinusitis.								
0	22686	94.45	1426623	92.27	223215	90.65	952047	91.14
1	1261	5.25	111705	7.22	21355	8.67	85577	8.19
2	70	0.29	7423	0.48	1586	0.64	6560	0.63
3	3	0.01	335	0.02	85	0.03	351	0.03
4	0	0.00	22	0.00	3	0.00	10	0.00
5	0	0.00	0	0.00	0	0.00	1	0.00
F7: Bronchiectasis, Parkinson disease, Glaucoma, Learning disability and IBS.								
0	23696	98.65	1521190	98.39	241666	98.14	1027143	98.33
1	322	1.34	24777	1.60	4556	1.85	17300	1.66
2	2	0.01	141	0.01	22	0.01	103	0.01
F8: Asthma, Dermatitis and Eczema, Constipation, COPD and Migraine.								
0	19438	80.92	1308665	84.64	205953	83.64	895829	85.76

1	3821	15.91	203282	13.15	34170	13.88	128054	12.26
2	686	2.86	31087	2.01	5495	2.23	18772	1.80
3	74	0.31	2923	0.19	594	0.24	1789	0.17
4	1	0.00	149	0.01	31	0.01	101	0.01
5	0	0.00	2	0.00	1	0.00	1	0.00
F9: Multiple sclerosis and Dementia.								
0	23985	99.85	1542704	99.78	245630	99.75	1042116	99.77
1	35	0.15	3404	0.22	613	0.25	2425	0.23
2	0	0.00	0	0.00	1	0.00	5	0.00

.

Supplementary Table S11. Odds ratio for F1 multimorbidity (≥ 2 score) according to F1 multimorbidity score in proband relatives. Odds Ratios (ORs) with 95 % confidence interval (CI) for F1 multimorbidity scores among relatives. Reference are relatives with probands with no F1 diseases (score<1). ORs for the included covariates (sex, year birth, education, county (region)) in the adjusted models are presented. F1=Hypertension, Heart failure, CHD, Diabetes, Obesity, Atrial fibrillation, Gout, Atherosclerosis and Renal disease.

	Twins		**Siblings**		**Half-siblings**		**Cousins**	
	OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Proband F1 score	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Score 0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Score 1	5.75 (3.93-8.42)	4.09 (2.75-6.11)	2.44(2.33-2.55)	1.85(1.76-1.94)	1.72(1.63-1.82)	1.46(1.38-1.54)	1.39(1.36-1.43)	1.18(1.14-1.21)
Score 2	9.02 (3.05-26.73)	3.11 (0.99-9.81)	6.14(5.56-6.78)	2.75(2.48-3.05)	2.75(2.38-3.17)	1.64(1.42-1.89)	2.15(2.01-2.30)	1.30(1.22-1.39)
Score 3	26.43 (7.29-95.83)	5.77 (1.14-29.30)	8.11(7.02-9.37)	2.98(2.56-3.47)	2.93(2.36-3.62)	1.45(1.16-1.80)	2.45(2.21-2.71)	1.34(1.22-1.48)
Score 4	---	---	8.73(6.70-11.39)	2.81(2.12-3.71)	3.40(2.39-4.84)	1.53(1.06-2.20)	2.74(2.30-3.25)	1.41(1.19-1.68)
Score ≥5	---	---	11.16(7.59-16.42)	3.33(2.22-4.98)	3.62(2.16-6.07)	1.45(0.84-2.53)	3.53(2.74-4.55)	1.73(1.34-2.24)

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S12. Odds ratio for F2 multimorbidity (≥ 2 score) according to F2 multimorbidity score in proband relatives. Odds Ratios (ORs) with 95 % confidence interval (CI) for F2 multimorbidity scores among relatives. Reference are relatives with probands with no F2 diseases (score<1). ORs for the included covariates (sex, year birth, education, county (region)) in the adjusted models are presented. F2= Affective disorders, Anxiety, Psychoactive substance misuse, Alcohol use disorders, Anorexia or bulimia, and Schizophrenia disorders.

	Twins		**Siblings**		**Half-siblings**		**Cousins**	
	OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Proband F2 score	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Score 0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Score 1	3.66 (3.09-4.45)	3.40 (2.83-4.10)	2.15(2.10-2.20)	2.13(2.08-2.18)	1.29(1.25-1.32)	1.27(1.24-1.30)	1.22(1.21-1.24)	1.22(1.20-1.23)
Score 2	5.98 (3.09-7.96)	5.59 (3.93-7.10)	3.20(3.08-3.33)	3.16(3.04-3.29)	1.51(1.45-1.58)	1.48(1.42-1.55)	1.35(1.31-1.38)	1.33(1.30-1.37)
Score 3	10.31 (7.19-14.79)	9.10 (6.29-13.20)	4.36(4.13-4.60)	4.29(4.06-4.52)	1.71(1.62-1.82)	1.67(1.58-1.77)	1.53(1.47-1.59)	1.52(1.46-1.58)
Score 4	---	---	5.13(4.71-5.59)	5.01(4.59-5.45)	1.85(1.70-2.01)	1.81(1.66-1.97)	1.65(1.55-1.75)	1.62(1.53-1.72)
Score ≥5	---	---	4.56(3.50-5.93)	4.64(3.56-6.04)	1.52(1.11-2.07)	1.50(1.09-2.05)	1.52(1.26-1.83)	1.52(1.26-1.84)

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S13. Odds ratio for F3 multimorbidity (≥ 2 score) according to F3 multimorbidity score in proband relatives. Odds Ratios (ORs) with 95 % confidence interval (CI) for F3 multimorbidity scores among relatives. Reference are relatives with probands with no F3 diseases (score<1). ORs for the included covariates (sex, year birth, education, county (region)) in the adjusted models are presented. F3= IBD, Liver disease, Pancreas diseases and Ulcer disease.

	Twins		**Siblings**		**Half-siblings**		**Cousins**	
	OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Proband F3 score	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Score 0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Score 1	14.71 (3.24-66.68)	8.97 (1.96-41.14)	3.19 (2.56-3.99)	2.45 (1.96-3.07)	1.46 (1.05-2.02)	1.13 (0.81-1.56)	1.28 (1.11-1.48)	1.10 (0.95-1.27)
Score 2	---	---	9.02 (4.22-19.29)	6.32 (2.95-13.53)	4.00 (1.28-12.47)	2.67 (0.86-8.29)	1.24 (0.55-2.80)	1.03 (0.45-2.31)
Score 3	---	---	24.39 (5.59-106.48)	17.10 (3.76-77.70)	---	---	1.42 (0.20-10.15)	1.04 (0.14-7.48)
Score 4	---	---	---	---	---	---	---	---
Score ≥5	---	---	---	---	---	---	---	---

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S14. Odds ratio for F4 multimorbidity (≥ 2 score) according to F4 multimorbidity score in proband relatives. Odds Ratios (ORs) with 95 % confidence interval (CI) for F4 multimorbidity scores among relatives. Reference are relatives with probands with no F4 diseases (score<1). ORs for the included covariates (sex, year birth, education, county (region)) in the adjusted models are presented. F4= Epilepsy, Blindness & low vision, Cerebrovascular disease, Cancer and Hearing loss (or impaired).

	Twins		**Siblings**		**Half-siblings**		**Cousins**	
	OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Proband F4 score	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Score 0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Score 1	2.73 (1.52-4.90)	2.17 (1.18-4.02)	1.72 (1.56-1.90)	1.37 (1.24-1.52)	1.21 (1.06-1.37)	1.06 (0.93-1.21)	1.16 (1.10-1.22)	1.03 (0.97-1.08)
Score 2	11.16 (2.63-47.24)	6.37 (1.25-32.54)	3.58 (2.49-5.15)	2.64 (1.83-3.80)	1.02 (0.51-2.03)	0.84 (0.42-1.68)	1.23 (0.94-1.61)	1.05 (0.80-1.37)
Score 3	---	---	2.49 (0.93-6.65)	1.82 (0.67-4.91)	---	---	0.85 (0.43-1.70)	0.77 (0.39-1.53)
Score 4	---	---	---	---	---	---	---	---
Score ≥5	---	---	---	---	---	---	---	---

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S15. Odds ratio for F5 multimorbidity (≥ 2 score) according to F5 multimorbidity score in proband relatives. Odds Ratios (ORs) with 95 % confidence interval (CI) for F5 multimorbidity scores among relatives. Reference are relatives with probands with no F4 diseases (score<1). ORs for the included covariates (sex, year birth, education, county (region)) in the adjusted models are presented. F5= Connective tissue disease, Osteoporosis, Thyroid disorders and Psoriasis.

	Twins		**Siblings**		**Half-siblings**		**Cousins**	
	OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Proband F5 score	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Score 0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Score 1	11.86 (5.47-25.73)	7.85 (3.50-18.40)	2.58 (2.29-2.90)	2.07 (1.84-2.33)	1.66 (1.41-1.96)	1.41 (1.20-1.67)	1.32 (1.23-1.42)	1.13 (1.05-1.22)
Score 2	67.21 (8.29-545.26)	40.95 (3.22-520.09)	8.11 (5.40-12.18)	5.76 (3.80-8.74)	4.93 (2.69-9.02)	3.76 (2.03-6.97)	2.13 (1.54-2.95)	1.67 (1.21-2.30)
Score 3	---	---	---	---	17.13 (6.33-46.34)	10.20 (3.60-28.87)	2.59 (0.97-6.93)	1.83 (0.68-4.96)
Score 4	---	---	---	---	---	---	---	---
Score ≥5	---	---	---	---	---	---	---	---

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S16. Odds ratio for F6 multimorbidity (≥ 2 score) according to F6 multimorbidity score in proband relatives. Odds Ratios (ORs) with 95 % confidence interval (CI) for F6 multimorbidity scores among relatives. Reference are relatives with probands with no F6 diseases (score<1). ORs for the included covariates (sex, year birth, education, county (region)) in the adjusted models are presented. F6= Prostate disease, Arthrosis, Painful back condition, Diverticular disease of intestine and Chronic sinusitis.

	Twins		**Siblings**		**Half-siblings**		**Cousins**	
	OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Proband F6 score	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Score 0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Score 1	3.95 (2.11-7.40)	1.45 (0.75-2.81)	3.10 (2.93-3.29)	1.59 (1.50-1.69)	1.91 (1.77-2.06)	1.27 (1.18-1.37)	1.70 (1.64-1.75)	1.14 (1.10-1.18)
Score 2	31.65 (8.83-113.39)	4.00 (0.98-16.18)	7.94 (6.63-9.51)	2.48 (2.07-2.99)	3.60 (2.81-4.60)	1.65 (1.28-2.11)	2.34 (2.05-2.67)	1.20 (1.06-1.37)
Score 3	---	---	14.10 (8.66-22.95)	3.50 (2.14-5.75)	6.39 (3.36-12.15)	2.07 (1.04-4.15)	2.62 (1.80-3.83)	1.16 (0.79-1.70)
Score 4	---	---	23.46 (5.48-100.39)	5.24 (1.12-24.55)	---	---	---	---
Score ≥5	---	---	---	---	---	---	---	---

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S17. Odds ratio for F7 multimorbidity (≥ 2 score) according to F7 multimorbidity score in proband relatives. Odds Ratios (ORs) with 95 % confidence interval (CI) for F7 multimorbidity scores among relatives. Reference are relatives with probands with no F7 diseases (score<1). ORs for the included covariates (sex, year birth, education, county (region)) in the adjusted models are presented. F7= Bronchiectasis, Parkinson disease, Glaucoma, Learning disability and IBS.

	Twins		**Siblings**		**Half-siblings**		**Cousins**	
	OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Proband F7 score	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Score 0	---	---	1.00	1.00	1.00	1.00	1.00	1.00
Score 1	---	---	3.21 (1.50-6.86)	3.06 (1.43-6.57)	1.98 (0.63-6.05)	1.92 (0.63-5.85)	1.15 (0.65-2.05)	1.13 (0.63-2.00)
Score 2	---	---	---	---	---	---	---	---
Score 3	---	---	---	---	---	---	---	---
Score 4	---	---	---	---	---	---	---	---
Score ≥5	---	---	---	---	---	---	---	---

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S18. Odds ratio for F8 multimorbidity (≥ 2 score) according to F8 multimorbidity score in proband relatives. Odds Ratios (ORs) with 95 % confidence interval (CI) for F8 multimorbidity scores among relatives. Reference are relatives with probands with no F8 diseases (score<1). ORs for the included covariates (sex, year birth, education, county (region)) in the adjusted models are presented. F8= Asthma, Dermatitis and Eczema, Constipation, COPD and Migraine.

	Twins		**Siblings**		**Half-siblings**		**Cousins**	
	OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Proband F8 score	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Score 0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Score 1	3.50 (2.95-4.16)	2.84 (2.39-3.39)	2.52 (2.46-2.60)	2.09 (2.04-2.15)	1.43 (1.38-1.49)	1.36 (1.31-1.41)	1.34 (1.32-1.37)	1.20 (1.18-1.22)
Score 2	15.83 (11.98-20.93)	11.62 (8.72-15.48)	7.02 (6.67-7.40)	4.91 (4.65-5.17)	2.14 (1.94-2.37)	1.95 (1.76-2.16)	1.80 (1.71-1.90)	1.40 (1.33-1.48)
Score 3	31.42 (18.69-52.81)	21.18 (12.53-35.78)	13.07 (11.82-14.46)	8.43 (7.61-9.34)	2.78 (2.27-3.40)	2.51 (2.06-3.07)	2.23 (1.98-2.52)	1.65 (1.46-1.86)
Score 4	---	---	18.04 (12.34-26.36)	12.84 (8.77-18.79)	2.47 (1.13-5.40)	2.77 (1.34-5.72)	2.17 (1.24-3.78)	1.77 (1.02-3.08)
Score ≥5	---	---	---	---	---	---	---	---

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S19. Odds ratio for F9 multimorbidity (≥ 2 score) according to F9 multimorbidity score in proband relatives. Odds Ratios (ORs) with 95 % confidence interval (CI) for F9 multimorbidity scores among relatives. Reference are relatives with probands with no F9 diseases (score<1). ORs for the included covariates (sex, year birth, education, county (region)) in the adjusted models are presented. F9= Multiple sclerosis and Dementia.

	Twins		**Siblings**		**Half-siblings**		**Cousins**	
	OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Proband F9 score	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Score 0	---	---	---	---	---	---	---	---
Score 1	---	---	---	---	---	---	---	---
Score 2	---	---	---	---	---	---	---	---
Score 3	---	---	---	---	---	---	---	---
Score 4	---	---	---	---	---	---	---	---
Score ≥5	---	---	---	---	---	---	---	---

Abbreviations: OR = odds ratio. The ORs were derived from double entry. Model 1 is a crude model (univariate). Model 2 is an adjusted model (multivariate), with adjustments for sex, year of birth, county, and educational attainment. In the logistic regression models, the "variance covariance (vce) cluster" method in STATA was used, which calculated robust standard errors using families as clusters.

Supplementary Table S20. Number of observation and cases in Table 2, i.e. Cases, No./person at risk No in Table 2.[£]

	Twins (n=24020)	Siblings (n=1546108)	Half-siblings (n=984976)	Cousins (n=6623156)
Proband score	Cases, No./person at risk No.	Cases, No./person at risk No.	Cases, No./person at risk No.	Cases, No./person at risk No.
Score 0	1502/14652	121783/938702	102229/543007	623597/3982972
Score 1	980/5838	62125/365498	50313/240617	261183/1555796
Score 2	543/2153	30869/142931	24698/107084	110063/621032
Score 3	269/807	14383/56830	12284/50049	48363/258695
Score 4	126/339	6871/23956	6223/23976	22516/113660
Score ≥5	110/231	5877/18191	5605/20243	18666/91001

[£]Based on double entry.

Supplementary Table S21. Number of observation and cases in Table 3, i.e. Cases, No./person at risk No in Table 3.[£]

	Twins (n=24020)	**Siblings (n=1546108)**	**Half-siblings (n=984976)**	**Cousins (n=6623156)**
Proband score	Cases, No./person at risk No.	Cases, No./person at risk No.	Cases, No./person at risk No.	Cases, No./person at risk No.
Score ≤1	2482/20490	183908/1304200	152542/783624	884780/5538768
Score ≥2	1048/3530	58000/241908	48810/201352	199608/1084388

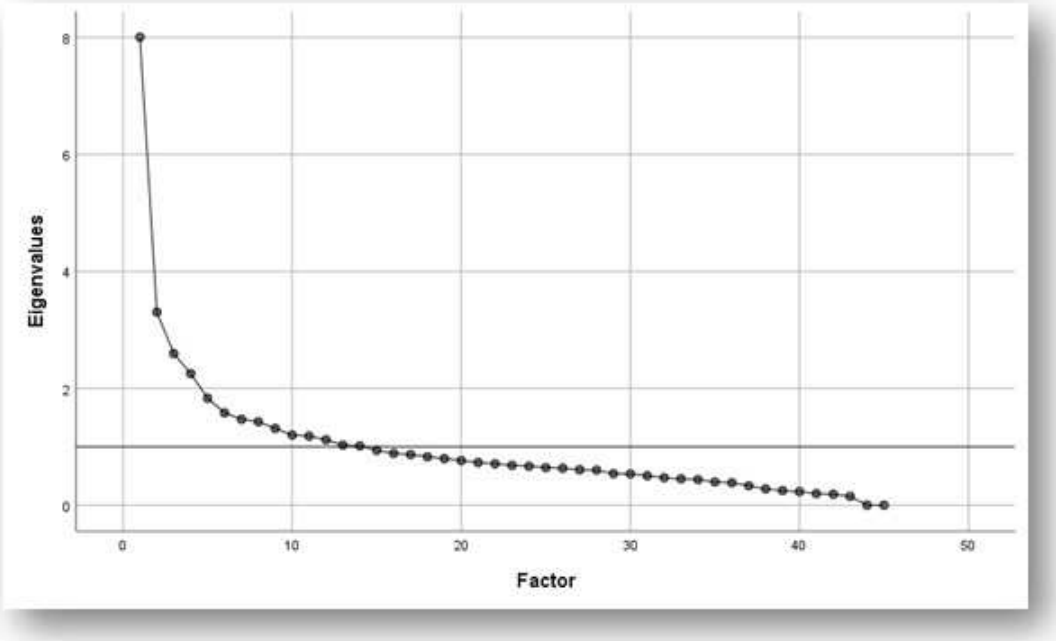
[£]Based on double entry.

Supplementary Table S22. Number of observation and cases in Table 5, i.e. Cases, No./person at risk No in Table 5.[£]

		Twins (n=24020)	Siblings (n=1546108)	Half-siblings (n=984976)	Cousins (n=6623156)
	Proband score	Cases, No./person at risk No.	Cases, No./person at risk No.	Cases, No./person at risk No.	Cases, No./person at risk No.
F1	**Score ≤1**	157/23853	16218/1528812	13911/970521	88399/6532075
	Score ≥2	10/167	1078/17296	544/14455	2682/91081
F2	**Score ≤1**	792/23042	59662/1478190	65695/911433	295130/6308020
	Score ≥2	186/978	8256/67918	7848/73543	20006/315136
F3	**Score ≤1**	NA	1587/1544505	1275/983695	7986/6615158
	Score ≥2	NA	16/1603	6/1281	12/7998
F4	**Score ≤1**	101/23915	5301/1540743	4171/980789	24829/6598217
	Score ≥2	4/105	64/5365	16/4187	110/24939
F5	**Score ≤1**	32/23986	3178/1542882	2358/982588	15374/6607706
	Score ≥2	2/34	48/3226	30/2388	76/15450
F6	**Score ≤1**	67/23947	7516/1538328	6285/978551	39268/6583364
	Score ≥2	6/73	264/7780	140/6425	524/39792
F7	**Score ≤1**	NA	NA	NA	NA
	Score ≥2	NA	NA	NA	NA
F8	**Score ≤1**	579/23259	30275/1511947	22215/961669	122665/6496331
	Score ≥2	182/761	3886/34161	1092/23307	4160/126825
F9	**Score ≤1**	NA	NA	NA	NA
	Score ≥2	NA	NA	NA	NA

[£]Based on double entry.

Supplement Figure S1. Plot of Eigenvalues according to number of factors in principal component analysis (PCA). The curve is flattening after 9 factors.



Statistical section

Principal component analysis (PCA) is a technique to describe the directions, i.e. linear combinations of the variables (i.e. diagnoses), that contain the majority of the co-variation of the diagnoses as is described by a correlation matrix.[28,29] These directions (the so-called eigenvectors) together with the eigenvalues (describing the variances in those directions) can fully describe the correlation structure seen among the 45 included diagnoses. A reduced number of eigenvectors (with large corresponding eigenvalues) can to a large extent reproduce the pattern seen in the correlation matrix. A factor analysis model is an attempt to explain the pattern seen in the correlation matrix as some variations occurring in the diagnoses from some common factors (the communality) and some other variations (the specific variances) that are unique to each individual variable (diagnosis) and not shared with other diagnoses as opposed to the communality.

In estimating such a model description, one possibility is to start with the eigenvectors from a PCA decomposition and rotate these in an oblique manner (non-orthogonal rotation) in order to assign as much as possible of the correlation pattern into the communality part, describing the influence from several – possibly correlated – common factors. Hence an initial principal component decomposition followed by a factor analysis with the principal factor method and an oblique promax rotation was used on the correlation matrix of tetra-choric correlations between the 45 diagnoses in patients to identify disease clusters. The disease clusters obtained by the factor analysis at the individual level were used for determination of familial risks among those smaller disease clusters. Thus, using PCA and factor analysis, we aimed to determine if disease clustering in multimorbidity is disease-specific or not.