

Online supplementary table 1: Descriptive statistics*

	Development dataset	Individuals with missing data	Validation dataset	Individuals with missing data
Sociodemographic and chronic disease variables				
Total (n=40,264)	24075 (66.7)	2782 (11.6)	12037 (33.3)	1370 (11.4)
Men	12422 (51.6)	0 (0)	6227 (51.7)	0 (0)
Mean age in years (SD)	62.4 (10.3)	0 (0)	65.7 (10.7)	0 (0)
45-50	3379 (14)		1658 (13.8)	
51-60	7969 (33.1)		3935 (32.7)	
61-70	7048 (29.3)		3462 (28.8)	
71-80	4496 (18.7)		2377 (19.7)	
81-85	1183 (4.9)		605 (5)	
Mood disorder**	3716 (15.4)	62 (0.2)	1801 (15)	26 (0.1)
Anxiety	1825 (7.6)	66 (0.2)	862 (7.2)	32 (0.1)
Osteoarthritis†	6233 (25.9)	295 (0.7)	3090 (25.7)	151 (0.4)
Rheumatoid arthritis	984 (4.1)	245 (0.6)	475 (3.9)	109 (0.3)
Other type of arthritis	3071 (12.8)	214 (0.5)	1496 (12.4)	116 (0.3)
Asthma	2874 (11.9)	69 (0.2)	1439 (12)	31 (0.1)
COPD	1357 (5.6)	95 (0.2)	672 (5.6)	36 (0.1)
Hypertension	8767 (36.4)	85 (0.2)	4344 (36.1)	42 (0.1)
Heart disease	2400 (10)	107 (0.3)	1240 (10.3)	45 (0.1)
Angina	1078 (4.5)	82 (0.2)	534 (4.4)	42 (0.1)
Myocardial infarction	1162 (4.8)	83 (0.2)	610 (5.1)	36 (0.1)
Peripheral vascular disease	1333 (5.5)	99 (0.2)	680 (5.6)	52 (0.1)
Stroke or TIA§	1016 (4.2)	121 (0.3)	491 (4.1)	74 (0.2)
Gastrointestinal ulcer	1789 (7.4)	78 (0.2)	906 (7.5)	38 (0.1)
Bowel disorder	2164 (9.0)	81 (0.2)	1069 (8.9)	38 (0.1)
Renal disease or failure	615 (2.6)	75 (0.2)	340 (2.8)	37 (0.1)
Hypothyroidism	2902 (12.1)	278 (0.7)	1515 (12.6)	134 (0.3)
Hyperthyroidism	548 (2.3)	280 (0.7)	266 (2.2)	131 (0.3)
Diabetes mellitus	3911 (16.2)	66 (0.2)	1958 (16.3)	23 (0.1)
Osteoporosis	2112 (8.8)	129 (0.3)	1091 (9.1)	68 (0.2)
Cataracts	6049 (25.1)	399 (1.0)	3061 (25.4)	175 (0.4)

Glaucoma	1070 (4.4)	83 (0.2)	575 (4.8)	40 (0.1)
Macular degeneration	939 (3.9)	127 (0.3)	446 (3.7)	60 (0.1)
Epilepsy	212 (0.9)	46 (0.1)	104 (0.9)	19 (0)
Parkinson's disease	93 (0.4)	41 (0.1)	35 (0.3)	15 (0)
Multiple sclerosis	152 (0.6)	36 (0.1)	73 (0.6)	25 (0.1)
Migraine	3210 (13.3)	62 (0.2)	1560 (13)	33 (0.1)
Allergies	9277 (38.5)	170 (0.4)	4505 (37.4)	63 (0.2)
Skin cancer: melanoma	404 (1.7)	0 (0)	208 (1.7)	0 (0)
Skin cancer: non-melanoma [¶]	800 (3.3)	0 (0)	413 (3.4)	0 (0)
Solid cancer [‡]	2295 (9.5)	0 (0)	1130 (9.4)	0 (0)
Haematological cancer [‡]	21 (0.1)	0 (0)	16 (0.1)	0 (0)
Ill-defined cancer	28 (0.1)	0 (0)	19 (0.2)	0 (0)
Current smoker [‡]	1754 (7.3)	144 (0.4)	906 (7.5)	64 (0.2)
Regular drinker [‡]	17496 (72.7)	17 (0)	8694 (72.2)	6 (0)
Healthcare utilisation and other measures of participant well-being				
Overnight hospitalisation	2113 (8.8)	14 (0)	1045 (8.7)	6 (0)
Emergency department visit	4968 (20.6)	38 (0.1)	2484 (20.6)	3 (0)
SWLS (SD)	28.3 (6.3)	332 (0.8)	28.3 (6.3)	170 (0.4)
OARS (SD)	27.8 (0.7)	184 (0.5)	27.8 (0.7)	110 (0.3)
Self-rated general health		19 (0)		14 (0)
Excellent	4977 (20.7)		2487 (20.7)	
Very good	9759 (40.5)		4986 (41.4)	
Good	6943 (28.8)		3434 (28.5)	
Fair	1966 (8.2)		928 (7.7)	
Poor	430 (1.8)		202 (1.7)	
Self-rated mental health		18 (0)		18 (0)
Excellent	7281 (30.2)		3604 (29.9)	
Very good	9835 (40.9)		4981 (41.4)	
Good	5774 (24.0)		2935 (24.4)	

Fair	1052 (4.4)		456 (3.8)	
Poor	133 (0.6)		61 (0.5)	

COPD: Chronic obstructive pulmonary disease; TIA: transient ischaemic attack; SWLS: Satisfaction with Life Scale; OARS: Older Americans Resources and Services measure of activities of daily living.

* n (% of total cohort), if not otherwise indicated.

** Mood disorder was defined as depression (including manic depression), bipolar disorder, mania, or dysthymia.

† All items related to osteoarthritis (e.g. osteoarthritis of the hand versus of the knee) were collapsed into osteoarthritis.

§ Stroke, cerebrovascular accident, and transient ischaemic attack (TIA) were collapsed into Stroke or TIA.

¶ Non-melanoma skin cancer and the not specified skin cancer were combined into skin cancer: non-melanoma.

* All cancers not included in the aforementioned categories, but excluding haematological cancer and ill-defined cancer, were combined into solid cancer.

‡ Mesothelial and soft tissue cancers were combined into haematological cancer.

** Current smokers were defined as having at least one cigarette daily for the last 30 days. All other participants were classified as non-current smokers.

‡ Regular alcohol consumption was defined as drinking alcohol at least once per month during the preceding 12 months. All other participants were classified as non-regular drinkers.

Online supplementary table 2: A comparison of participants with and without missing values requiring imputation

		Participants without missing values	Participants with missing values
Total n (%)		36,112 (89.7)	4,152 (10.3)
Mean CLSA analytical weight (SD)		1.0 (0.5)	0.9 (0.4)
Men n (%)		18,649 (51.6)	2,221 (53.5)
Mean age in years (SD)		62.4 (10.3)	65.7 (10.7)
Age in years n (%)	45-50	5,037 (13.9)	365 (8.8)
	51-60	11,904 (33.0)	1,062 (25.6)
	61-70	10,510 (29.1)	1,235 (29.7)
	71-80	6,873 (19.0)	1,095 (26.4)
	81-85	1,788 (5.0)	395 (9.5)

Online supplementary table 3: The five regression models, predicting overnight hospitalisation in the last 12 months

	Reference model	Model 1 – absolute count**	Model 2 – individual conditions	Model 3 – LASSO	Model 4 – AIC	Model 5 - BMA
Development dataset						
	Regression coefficient (95% CI)					
Intercept	-4.947 (-5.217, -4.678)	-4.829 (-5.101, -4.556)	-4.387 (-4.737, -4.038)	-3.685	-4.403 (-4.756, -4.049)	-2.721 (-2.824, -2.618)
Age	0.041 (0.037, 0.045)	0.024 (0.02, 0.029)	0.023 (0.018, 0.029)	0.018	0.024 (0.018, 0.03)	-
Sex	0.041 (-0.047, 0.128)	-0.009 (-0.097, 0.08)	0.154 (0.055, 0.252)	-	0.140 (0.04, 0.241)	0.077 (-0.271, 0.425)
Sum of diseases	-	0.186 (0.168, 0.203)	-	-	-	-
Sum*Age	-	-	-	-	-	-
Mood disorder	-	-	0.233 (0.114, 0.352)	0.060	0.239 (0.114, 0.363)	0.235 (0.14, 0.33)
Anxiety	-	-	0.204 (0.050, 0.358)	-	0.207 (0.047, 0.368)	-
Osteoarthritis	-	-	0.398 (0.298, 0.498)	0.268	0.407 (0.303, 0.511)	-
Rheumatoid Arthritis	-	-	0.077 (-0.121, 0.276)	-	-	0.354 (0.207, 0.501)
Other type of arthritis	-	-	0.084 (-0.042, 0.211)	-	0.092 (-0.204, 0.388)	-
Asthma	-	-	0.064 (-0.07, 0.199)	-	-	-
COPD	-	-	0.359 (0.196, 0.522)	0.187	0.385 (0.225, 0.545)	0.519 (0.401, 0.637)
Hypertension	-	-	0.277 (0.180, 0.373)	0.177	0.280 (0.184, 0.377)	-
Heart disease	-	-	0.380 (0.229, 0.531)	0.363	0.394 (0.243, 0.545)	-
Angina	-	-	0.123 (-0.07, 0.317)	0.049	0.131 (-0.255, 0.517)	0.512 (0.374, 0.65)
Myocardial infarction	-	-	0.249 (0.058, 0.441)	0.110	0.260 (0.068, 0.451)	0.495 (0.356, 0.633)
Peripheral vascular disease	-	-	0.267 (0.107, 0.427)	0.103	0.279 (0.120, 0.439)	-
Stroke or TIA	-	-	0.103 (-0.088, 0.295)	-	-	0.513 (0.378, 0.647)
Gastrointestinal ulcer	-	-	0.311 (0.166, 0.456)	0.109	0.320 (0.176, 0.464)	0.305 (0.191, 0.419)
Bowel disorder	-	-	0.151 (0.009, 0.294)	-	0.161 (0.019, 0.303)	0.238 (0.128, 0.349)
Kidney disease or failure	-	-	0.354 (0.134, 0.574)	0.082	0.357 (0.137, 0.577)	0.571 (0.408, 0.735)
Hypothyroidism	-	-	0.106 (-0.026, 0.239)	-	0.110 (-0.022, 0.243)	0.127 (-0.228, 0.482)
Hyperthyroidism	-	-	0.173 (-0.096, 0.443)	-	0.193 (-0.372, 0.758)	-
Diabetes	-	-	0.203 (0.09, 0.316)	0.076	0.205 (0.093, 0.318)	-
Osteoporosis	-	-	0.036 (-0.119, 0.19)	-	-	0.253 (0.134, 0.373)
Cataracts	-	-	0.070 (-0.049, 0.189)	-	-	-
Glaucoma	-	-	-0.084 (-0.29, 0.123)	-	-	-

Macular degeneration	-	-	-0.040 (-0.179, 0.258)	-	-	-
Epilepsy	-	-	0.498 (0.122, 0.873)	-	0.500 (0.125, 0.874)	-
Parkinson's disease	-	-	0.864 (0.337, 1.392)	-	0.877 (0.350, 1.404)	1.009 (0.624, 1.395)
Multiple sclerosis	-	-	0.607 (0.180, 1.034)	-	0.612 (0.185, 1.038)	-
Migraine	-	-	0.026 (-0.103, 0.155)	-	-	-
Allergies	-	-	0.071 (-0.024, 0.167)	-	0.086 (-0.007, 0.179)	-
Skin cancer: melanoma	-	-	-0.175 (-0.528, 0.177)	-	-	-
Skin cancer: non-melanoma	-	-	-0.077 (-0.328, 0.175)	-	-	-
Solid cancer	-	-	0.167 (0.022, 0.310)	-	0.170 (0.026, 0.313)	0.375 (0.273, 0.476)
Haematological and soft cancer	-	-	0.795 (-0.380, 1.97)	-	0.815 (-0.597, 2.226)	-
Ill-defined cancer	-	-	-0.133 (-1.444, 1.179)	-	-	-
Current smoker	-	-	0.115 (-0.048, 0.277)	-	0.116 (-0.201, 0.432)	-
Regular drinker	-	-	-0.274 (-0.370, -0.177)	-0.093	0.280 (-0.376, -0.183)	-
Validation dataset						
Nagelkerke's R ²	0.03	0.06	0.07	0.05	0.07	0.04
C-statistic (95% CI)	0.62 (0.60, 0.64)	0.67 (0.65, 0.68)	0.67 (0.67, 0.70)	0.67 (0.65, 0.68)	0.68 (0.66, 0.69)	0.63 (0.62, 0.65)
Pearson correlation [‡] (95% CI)	0.83 (0.74, 0.89)	0.90 (0.86, 0.93)	0.88 (0.83, 0.91)	0.88 (0.83, 0.91)	0.86 (0.81, 0.90)	0.86 (0.74, 0.90)

AIC: Akaike Information Criterion; BMA: Bayesian Model Averaging; COPD: Chronic obstructive pulmonary disease; TIA: transient ischaemic attack; 95% CI: 95% confidence interval

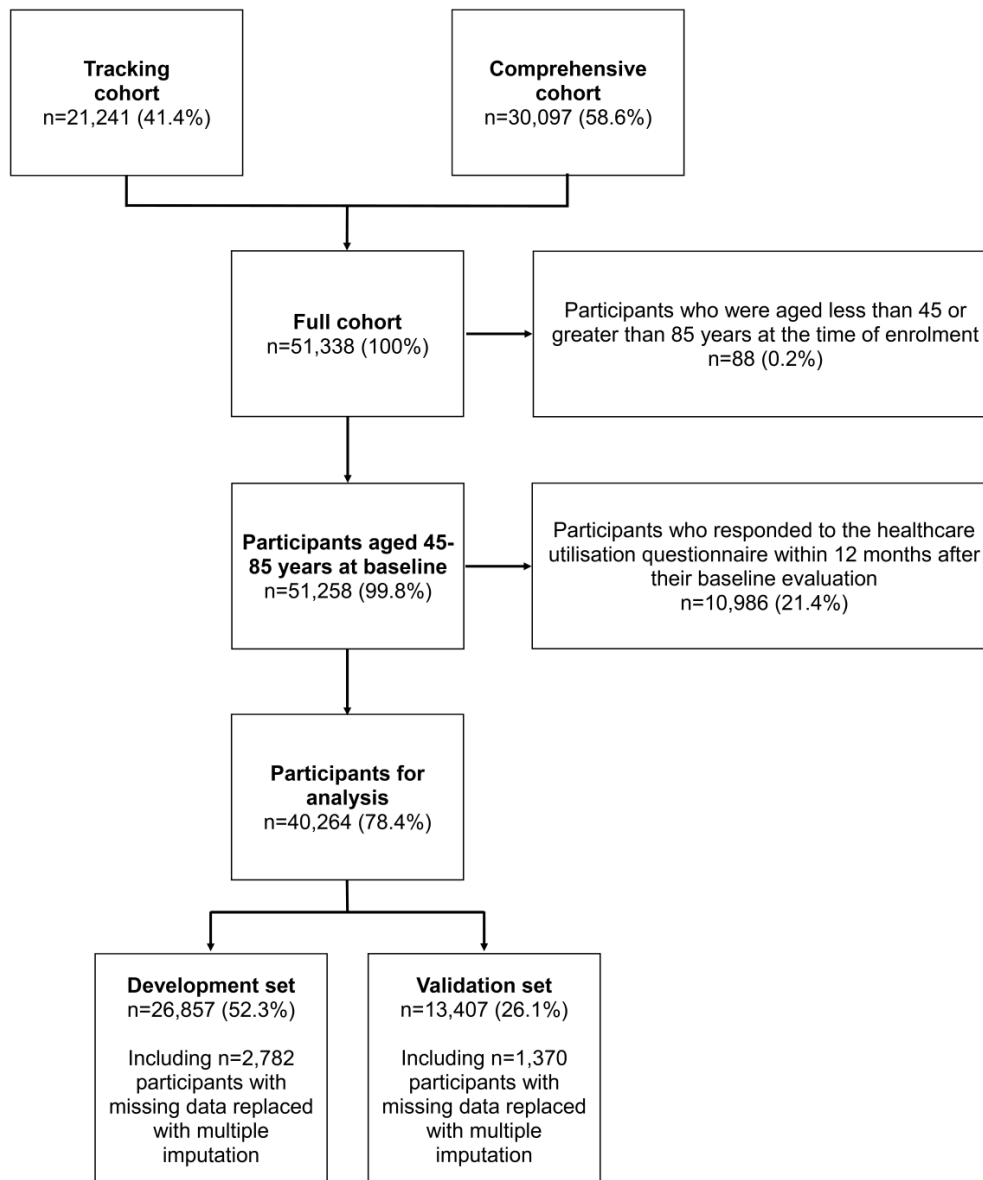
¶ In Index 3, some of the coefficients are shrunk to zero by LASSO regression. In addition, LASSO does not allow for the calculation of a 95% CI.

* Correlation between the predicted and the observed probability of the outcome, using 130 quantiles.

* There were many estimates for each measure due to multiple imputation. We report the median estimates.

** Model 1 is a regression model including the absolute count of chronic conditions as a single independent variable. Model 2 is a regression model where each chronic condition is treated as a dichotomous independent variable. Models 3-5 are regression models

that began with the same independent variables as those used for Model 2, but then involved variable selection using LASSO, stepwise selection based on the AIC, or BMA, respectively.

Online supplementary figure 1: Participant flow diagram

Online supplemental figure 2: Calibration plots for the 5 multimorbidity indices, without an interaction term