

Web extra table 1. Numbers of patients in CPRD and QResearch by geographical area

	<i>CPRD</i>	<i>Col %</i>	<i>QResearch</i>	<i>Col %</i>
East Midlands	109,428	3.3	500,970	7.4
East of England	397,008	12.1	528,379	7.8
London	563,353	17.2	1,711,956	25.3
North East	59,558	1.8	269,695	4.0
North West	474,457	14.5	830,047	12.3
South Central	411,571	12.6	696,070	10.3
South East	362,319	11.1	545,811	8.1
South West	397,735	12.2	700,041	10.4
West Midlands	348,614	10.7	589,548	8.7
Yorks & Humber	147,469	4.5	386,132	5.7
Total	3,271,512	100.00	6,758,649	100.00

Web table 2 Recording of ethnicity and family history of coronary heart disease (FH CHD) by geographical area

	CPRD linked data					QResearch					Ratio recording QResearch:CPRD	
	total patients	ethnicity recorded	FH CHD recorded		total patients	ethnicity recorded	FH CHD recorded				ethnicity	FH
	count	count	Row %	count	Row %	count	count	Row %	count	Row %		
East Midlands	109,428	23,636	21.6	7,912	7.2	500,970	284,194	56.7	60,278	12.0	2.6	1.7
East of England	397,008	127,427	32.1	17,597	4.4	528,379	319,742	60.5	54,252	10.3	1.9	2.3
London	563,353	308,285	54.7	21,034	3.7	1,711,956	1,095,835	64.0	162,282	9.5	1.2	2.5
North East	59,558	17,140	28.8	3,326	5.6	269,695	195,127	72.4	43,168	16.0	2.5	2.9
North West	474,457	196,987	41.5	25,915	5.5	830,047	460,640	55.5	112,718	13.6	1.3	2.5
South Central	411,571	140,448	34.1	16,989	4.1	696,070	380,908	54.7	73,051	10.5	1.6	2.5
South East	362,319	90,633	25.0	12,618	3.5	545,811	262,861	48.2	45,765	8.4	1.9	2.4
South West	397,735	137,806	34.6	17,829	4.5	700,041	375,155	53.6	75,091	10.7	1.5	2.4
West Midlands	348,614	148,012	42.5	20,375	5.8	589,548	347,479	58.9	63,412	10.8	1.4	1.8
Yorks & Humber	147,469	56,340	38.2	6,195	4.2	386,132	214,702	55.6	54,515	14.1	1.5	3.4
Total	3,271,512	1,246,714	38.1	149,790	4.6	6,758,649	3,936,643	58.2	744,532	11.0	1.5	2.4

Web extra table 3 comparison of baseline characteristics for the CPRD cohort with linked data used for the validation cohort in this study and the CPRD cohort without linked data (which was not used for the validation of the QPrediction scores).

	CPRD linked data	CPRD unlinked data
female	1682709 (51.4)	1166103 (51.3)
male	1588803 (48.6)	1108235 (48.7)
25-34 years	895167 (27.4)	665645 (29.3)
35-44 years	760830 (23.3)	503091 (22.1)
45-54 years	571937 (17.5)	384686 (16.9)
55-64 years	424453 (13.0)	293826 (12.9)
65-74 years	312352 (9.5)	216937 (9.5)
75+ years	306773 (9.4)	210153 (9.2)
Ethnicity recorded	1246714 (38.1)	645829 (28.4)
White or not recorded	3117325 (95.3)	2209396 (97.1)
Indian	32467 (1.0)	11751 (0.5)
Pakistani	12752 (0.4)	6358 (0.3)
Bangladeshi	4107 (0.1)	2682 (0.1)
Other Asian	22668 (0.7)	8854 (0.4)
Caribbean	11414 (0.3)	4812 (0.2)
Black African	27654 (0.8)	9751 (0.4)
Chinese	7090 (0.2)	3416 (0.2)
Other ethnic group	36035 (1.1)	17318 (0.8)
Smoking status recorded	3037626 (92.9)	2066777 (90.9)
Non smoker	1448554 (44.3)	1006511 (44.3)
Ex-smoker	475488 (14.5)	306460 (13.5)
Light smoker (1-9/day)	214330 (6.6)	127487 (5.6)
Moderate smoker (10-19/day)	362391 (11.1)	277693 (12.2)
Heavy smoker (20+/day)	229912 (7.0)	159718 (7.0)
Smoker amount not recorded	306951 (9.4)	188908 (8.3)
Alcohol status recorded	2617112 (80.0)	1759758 (77.4)
Non drinker	482513 (14.7)	361566 (15.9)
Trivial <1u/day	1186942 (36.3)	754470 (33.2)
Light 1-2u/day	701808 (21.5)	479407 (21.1)
Moderate 3-6/day	203091 (6.2)	135130 (5.9)
Heavy 7-9u/day	22465 (0.7)	15645 (0.7)
Very Heavy >/day	20293 (0.6)	13540 (0.6)
family history		
family history of CHD	149790 (4.6)	97401 (4.3)
family history of diabetes	229200 (7.0)	144713 (6.4)
family history osteoporosis	10942 (0.3)	6164 (0.3)
family history kidney disease	2839 (0.1)	1592 (0.1)

prescribed medication		
antidepressants	337350 (10.3)	232657 (10.2)
anticoagulants	29032 (0.9)	20338 (0.9)
antipsychotics	113398 (3.5)	76819 (3.4)
oral NSAIDs	592931 (18.1)	425739 (18.7)
tamoxifen	9343 (0.3)	6107 (0.3)
antiplatelets	190291 (5.8)	137793 (6.1)
oestrogen only HRT	119413 (3.7)	75448 (3.3)
corticosteroids	116949 (3.6)	70793 (3.1)
oral contraceptive pill	174288 (5.3)	126218 (5.5)
recorded diagnoses		
congestive cardiac failure	35543 (1.1)	25823 (1.1)
atrial fibrillation	40227 (1.2)	27032 (1.2)
coronary heart disease	138080 (4.2)	102493 (4.5)
cardiovascular disease	184597 (5.6)	134650 (5.9)
rheumatoid arthritis	24711 (0.8)	17427 (0.8)
chronic renal disease	8050 (0.2)	5774 (0.3)
type 1 diabetes	11162 (0.3)	7778 (0.3)
type 2 diabetes	94905 (2.9)	63240 (2.8)
venous thromboembolism	38162 (1.2)	23593 (1.0)
varicose veins	65991 (2.0)	44717 (2.0)
moderate/severe kidney failure	23607 (0.7)	15072 (0.7)
severe kidney failure	2728 (0.1)	1839 (0.1)
oesophageal varices	809 (0.0)	674 (0.0)
inflammatory bowel disease	19170 (0.6)	13095 (0.6)
SLE	1984 (0.1)	1273 (0.1)
peripheral vascular disease	30130 (0.9)	23066 (1.0)
dementia	23320 (0.7)	15858 (0.7)
Parkinson's disease	9222 (0.3)	5854 (0.3)
cancer	70774 (2.2)	45637 (2.0)
liver disease	6852 (0.2)	5041 (0.2)
malabsorption	16718 (0.5)	12007 (0.5)
endocrine diseases	17179 (0.5)	12479 (0.5)
COPD	44766 (1.4)	33190 (1.5)
chronic liver disease or pancreatitis	9572 (0.3)	6899 (0.3)
renal stones	19858 (0.6)	14935 (0.7)
care home resident	4873 (0.1)	2859 (0.1)
falls	90783 (2.8)	53221 (2.3)
prior fracture	70017 (2.1)	50346 (2.2)
asthma or COPD	312477 (9.6)	207765 (9.1)
treated hypertension	285293 (8.7)	190707 (8.4)
platelets < 150 or > 480	24333 (0.7)	12651 (0.6)
emergency admission or hip op	8749 (0.3)	6468 (0.3)
prior haemorrhage	177327 (5.4)	122024 (5.4)

Recorded values		
BMI recorded	2750153 (84.1)	1864134 (82.0)
SBP recorded	2949786 (90.2)	2010003 (88.4)
cholesterol/HDL ratio recorded	1193900 (36.5)	761573 (33.5)
platelets recorded	606260 (18.5)	302478 (13.3)
mean age (SD)	47.9 (17.0)	47.4 (17.2)
mean townsend score (SD)	-.5 (3.2)	.1 (3.7)
mean BMI (SD)	28.9 (6.9)	29.2 (7.1)
mean cholesterol ratio (SD)	4.1 (1.4)	4.1 (1.4)
mean SBP(SD)	130.7 (23.2)	130.1 (151.3)

Web extra table 4 Number of cases of upper gastrointestinal bleed and intracranial bleed on CPRD and QResearch (one third sample database). Incidence rates per 1000 pyrs have been standardised to the QResearch population in 5 year bands.

	<i>CPRD validation</i>		<i>QResearch validation</i>	
	cases on	age standardised Incidence rate per 1000pyrs	cases	Age standardised Incidence rate per 1000pyrs
no anticoagulants	13,314	1.41 (1.39 to 1.43)	6,447	1.33 (1.30 to 1.36)
anticoagulants	359	6.70 (4.06 to 9.34)	153	6.10 (3.20 to 8.98)
Intracranial bleed				
no anticoagulants	5,190	0.53 (0.51 to 0.54)	2,716	0.56 (0.54 to 0.58)
anticoagulants	233	2.45 (1.23 to 3.68)	104	2.87 (1.11 to 4.39)

Web extra table 5 Performance of QPrediction scores on the CPRD validation cohort, restricted to patients with complete data for relevant laboratory and clinical values

		<i>CPRD</i>	<i>CPRD</i>
		women	men
	statistic	mean (95%CI)	mean (95%CI)
QDiabetes-2013	ROC	0.849 (0.847 to 0.85)	0.814 (0.813 to 0.816)
(type 2 diabetes)	R ² (%)	49.8 (49.4 to 50.2)	44.4 (44 to 44.8)
	D statistic	2.04 (2.024 to 2.056)	1.828 (1.814 to 1.842)
QKidney-2010	ROC	0.847 (0.842 to 0.852)	0.839 (0.835 to 0.842)
(moderate or severe kidney failure)	R ² (%)	53.4 (52.8 to 54)	49.7 (49.3 to 50.1)
	D statistic	2.19 (2.165 to 2.215)	2.036 (2.018 to 2.054)
QKidney -2010	ROC	0.816 (0.798 to 0.834)	0.808 (0.795 to 0.822)
(severe kidney failure)	R ² (%)	47.3 (45.2 to 49.4)	46.1 (44.5 to 47.7)
	D statistic	1.938 (1.856 to 2.02)	1.895 (1.834 to 1.956)
QRISK2-2014	ROC	0.791 (0.787 to 0.796)	0.757 (0.753 to 0.761)
(cardiovascular disease)	R ² (%)	40.9 (40 to 41.8)	31.8 (30.9 to 32.7)
	D statistic	1.704 (1.673 to 1.735)	1.398 (1.371 to 1.425)
QStroke-2013	ROC	0.794 (0.79 to 0.797)	0.771 (0.768 to 0.774)
(ischaemic stroke or TIA)	R ² (%)	42.1 (41.4 to 42.8)	36.6 (35.9 to 37.3)
	D statistic	1.747 (1.723 to 1.771)	1.557 (1.535 to 1.579)
QThrombosis-2010	ROC	0.755 (0.75 to 0.76)	0.762 (0.756 to 0.767)
(venous thromboembolism)	R ² (%)	34.6 (33.8 to 35.4)	32.6 (31.7 to 33.5)
	D statistic	1.487 (1.462 to 1.512)	1.424 (1.395 to 1.453)
QBleed-20141	ROC statistic	0.773 (0.766 to 0.779)	0.751 (0.744 to 0.758)
(upper GI bleed)	R ² (%)	43.6 (42.1 to 45.1)	39.6 (38.1 to 41.1)
	D statistic	1.799 (1.744 to 1.854)	1.658 (1.605 to 1.711)
QBleed-2014	ROC statistic	0.812 (0.803 to 0.822)	0.791 (0.78 to 0.802)
(intracranial bleed)	R ² (%)	53 (51 to 55)	50.2 (48 to 52.4)
	D statistic	2.172 (2.086 to 2.258)	2.057 (1.967 to 2.147)
QFracture-2012	ROC	0.899 (0.896 to 0.901)	0.866 (0.86 to 0.872)
(fracture neck of femur)	R ² (%)	70.4 (69.9 to 70.9)	67.1 (66.2 to 68)
	D statistic	3.159 (3.124 to 3.194)	2.922 (2.861 to 2.983)
QFracture -2012	ROC	0.819 (0.816 to 0.821)	0.757 (0.751 to 0.763)
(osteoporotic fracture)	R ² (%)	53.9 (53.4 to 54.4)	46.1 (45 to 47.2)
	D statistic	2.213 (2.189 to 2.237)	1.893 (1.852 to 1.934)