

Additional file 2: Characteristics of Fenland study participants included and excluded from the analytic sample

Variable ⁱ		Included n = 11,396 (91.65%)	Excluded n = 1,038 (8.35%)	Statistical tests ⁱⁱ
	Male	5,321 (46.69)	422 (40.66)	
	Female	6,075 (53.31)	616 (59.34)	
Age (years)	Median (IQR)	48.9 (42.7, 54.8)	48.0 (42.2, 53.5)	z = -3.69, p=0.0002
Alcohol (grams/day)	Median (IQR)	5.47 (1.27, 10.72)	5.14 (0.79, 10.88)	z = -0.94, p=0.35
	≤16	4,570 (40.10)	362 (38.39)	
	>16 to ≤18	2,521 (22.12)	204 (21.63)	
	>18	4,305 (37.78)	377 (39.98)	
Physical activity (kj ^{iv} /kg ^v /day)	No	10,045 (88.14)	742 (83.75)	
	Yes	1,351 (11.86)	144 (16.25)	
	No	8,677 (76.14)	798 (76.88)	
	Yes	2,719 (23.86)	240 (23.12)	
	Median (IQR)	51.00 (37.84, 66.75)	48.25 (34.45, 64.61)	z = -3.70, p=0.0002
	No	1,959 (17.19)	217 (20.91)	
Yes	9,437 (82.81)	821 (79.09)		
	No	10,116 (88.77)	896 (89.60)	
	Yes	1,280 (11.23)	104 (10.40)	
	<3x/week	704 (6.18)	79 (7.79)	
	3-5x/week	3,688 (32.36)	360 (35.50)	
	>5x/week	7,004 (61.46)	575 (56.71)	

ⁱ Results shown as number (column percentage). Median (inter-quartile range) shown for: age, alcohol, physical activity

ⁱⁱ Testing for significant differences between included and excluded populations using Mann-Whitney test with z-scores for continuous variables, and Pearson Chi squared test for categorical variables. Significance at 1% level

ⁱⁱⁱ History of diabetes in first degree relative

^{iv} kj = kilojoules

^v kg = kilograms