International Journal of Obesity (2011) 35, 300–308

0307-0565/11
www.nature.com/ijo

ORIGINAL ARTICLE

C-reactive protein levels and body mass index: elucidating direction of causation through reciprocal Mendelian randomization

NJ Timpson[1], BG Nordestgaard[2,3], RM Harbord[1,4], J Zacho[2,3], TM Frayling[5,6], A Tybjærg-Hansen[3,7] and G Davey Smith[1]

[1]*MRC Centre for Causal Analysis in Translational Epidemiology, Bristol University, Bristol, UK;* [2]*Faculty of Health Sciences, Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, University of Copenhagen, Copenhagen, Denmark;* [3]*Faculty of Health Sciences, The Copenhagen General Population Study, Herlev Hospital, Copenhagen University Hospital, University of Copenhagen, Copenhagen, Denmark;* [4]*Department of Social Medicine, Bristol University, Bristol, UK;* [5]*Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, UK;* [6]*Diabetes Genetics, Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, UK and* [7]*Faculty of Health Sciences, Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital, University of Copenhagen, Copenhagen, Denmark*

Context: The assignment of direction and causality within networks of observational associations is problematic outside randomized control trials, and the presence of a causal relationship between body mass index (BMI) and C-reactive protein (CRP) is disputed.
Objective: Using reciprocal Mendelian randomization, we aim to assess the direction of causality in relationships between BMI and CRP and to demonstrate this as a promising analytical technique.
Participants and methods: The study was based on a large, cross-sectional European study from Copenhagen, Denmark. Genetic associates of BMI (*FTO*(rs9939609)) and circulating CRP (*CRP*(rs3091244)) have been used to reexamine observational associations between them.
Results: Observational analyses showed a strong, positive association between circulating CRP and BMI (change in BMI for a doubling in logCRP of 1.03 $kg\,m^{-2}$ (95% confidence interval (95% CI): 1.00, 1.07), $P<0.0001$). Analysis using *CRP*(rs3091244) to re-estimate the causal effect of circulating CRP on BMI yielded null effects (change in BMI for a doubling in logCRP of $-0.24\,kg\,m^{-2}$ (95% CI: −0.58, 0.11), $P=0.2$). In contrast, analysis using *FTO*(rs9939609) to assess the causal effect of BMI on circulating CRP confirmed observational associations (ratio of geometric means of CRP per s.d. increase in BMI 1.41 (95% CI: 1.10, 1.80), $P=0.006$).
Conclusions: Taken together, these data suggest that the observed association between circulating CRP and measured BMI is likely to be driven by BMI, with CRP being a marker of elevated adiposity. More generally, the method of reciprocal randomization has general applicability in determining the direction of causation within inter-correlated networks of metabolic components.
International Journal of Obesity (2011) **35**, 300–308; doi:10.1038/ijo.2010.137; published online 17 August 2010

Keywords: Mendelian randomization; *FTO*; *CRP*; hsCRP; BMI

Introduction

The associations between inflammation and obesity related traits, including impaired insulin resistance, type II diabetes and coronary heart disease, have been investigated extensively in recent years. The acute-phase protein, C-reactive protein (CRP), has been a particular focus of these investigations which have and continue to report associations between circulating CRP levels, obesity and cardiovascular outcomes.[1–3] Some prospective studies have suggested that inflammatory markers in general, and CRP in particular, cause the development of elevated adiposity, obesity and diabetes.[4–6] Other evidence has suggested that obesity is a determinant of inflammatory marker status, including CRP

Correspondence: Dr NJ Timpson, MRC Centre for Causal Analysis in Translational Epidemiology, Bristol University, Oakfield House, Oakfield Grove, Bristol BS2 8BL, UK.
E-mail: n.j.timpson@bris.ac.uk
Received 13 February 2010; revised 22 April 2010; accepted 27 May 2010; published online 17 August 2010

level.[1,7,8] The direction of causation is difficult to determine outside the realm of experimental trial designs, in which exposures may be held constant, because of the highly intercorrelated nature of the factors involved (both exogenous confounders and in the form of bias), the existence of reverse causation and given the inevitable degree of measurement imprecision that may be encountered in such settings.[9–11]

Mendelian randomization, the utilization of genetic variants as proxies for particular phenotypic measures, offers a potential approach to assess the direction of association and, thus likely the causality in observational data. Germline genetic variants are generally neither associated with confounding factors nor can they be influenced by the outcome measure (that is, they are not susceptible to reverse causation).[12] Genetic variants related to an intermediate risk factor of interest (such as circulating CRP level)—particularly *cis* variants, which are likely to reflect effects on gene expression—should produce downstream effects on outcome phenotypes (such as body mass index (BMI)), only if the latter is influenced causally by the intermediate risk factor in question.

A study with measures of degree of adiposity through BMI, circulating CRP level and genetic variants related independently to both of these phenotypes allows a particularly clear assessment of the causal direction of association between BMI and CRP. In this paper, we exploit the properties of variation at the *CRP* and *FTO* gene loci to perform a bidirectional Mendelian randomization experiment, which we refer to as a reciprocal Mendelian randomization. We aimed to elucidate the driving agent behind observational associations between circulating CRP and BMI, and to illustrate the general potential of this technique.

Methods

The Copenhagen General Population Study is a cross-sectional study of the Danish general population initiated in 2003 and still recruiting;[13,14] the aim is to enroll 100 000 participants ascertained using the same methods as those used in the Copenhagen City Heart Study,[15] but with a focus on all multifactorial diseases. At the time of genotyping for this study, 37 027 individuals had been included (response rate 45%); however, complete data for BMI, CRP, covariates and genotypes (representing the smallest possible sample size for analyses within this sample) were available for 21 836 participants. All participants were White, of Danish descent and were selected on the basis of the National Danish Civil Registration System to be representative of the adult Copenhagen general population aged 20–80+ years. Details of data collection procedures have been reported previously.[16] For these analyses, examination data on height and weight and questionnaire data on age, sex, smoking and alcohol consumption, income level and educational level were utilized. Circulating CRP was assessed and genotyping carried out on extracted DNA.

Outcome variables

High-sensitivity measurement of circulating CRP was measured once by high-sensitivity laser nephelometry (Dade-Behring, Atterbury, Milton Keynes, UK). Upper and lower limits for circulating values for CRP were set at 30 and 0.174 mg l^{-1}, respectively. The lowest value was set at the limits for accurate high-sensitivity circulating CRP measurement, and the highest value was imposed to avoid the inclusion of those with acute elevation of CRP. Owing to the known highly skewed distribution of circulating CRP (confirmed in this cohort), this variable was log transformed before analyses to approximate a normal distribution. Where appropriate, results are back transformed by exponentiation and effects are expressed as ratios of geometric means.

Weight and height were measured once, and BMI was calculated as weight (kg) divided by height squared (m^2). In all regression analyses, bar that used for the generation of Figures 1 and 2, we used raw measures of BMI. To remove the dependence of BMI on sex, age and height, we generated the measure residual BMI. For this measure, BMI was regressed on sex, age, age squared, log(height) and an age–sex interaction. The residuals from this model give the difference between an individual's actual BMI and that expected for their sex, age and height. For analyses, those with a

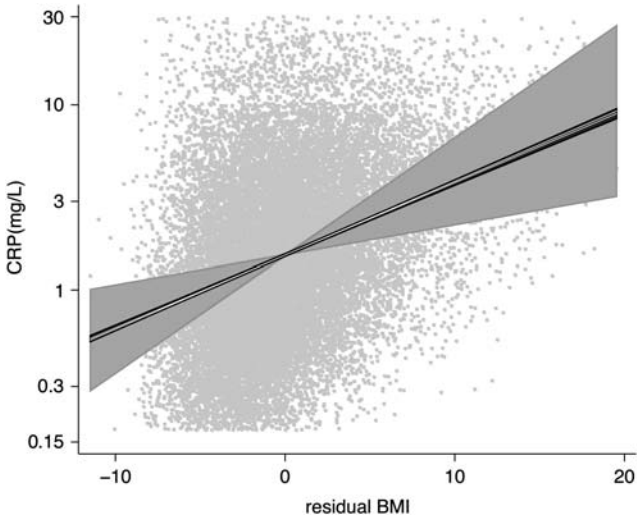

Figure 1 Comparison of linear relationships between circulating CRP and residual BMI observationally and when estimated using the *FTO* loci as an instrument for residual BMI. *X* and *Y* axes represent residual BMI and CRP, respectively. Light grey points represent a scatter plot of the correlation between circulating CRP and residual BMI. Gray areas represent 95% confidence regions around instrumental variables estimates. Black area represents 95% confidence regions around simple linear regression estimates. (The 50 individuals with extreme residual BMI over 20 kg m^{-2} are not shown on the plot but were included in the analyses that gave the fitted lines and confidence regions.)

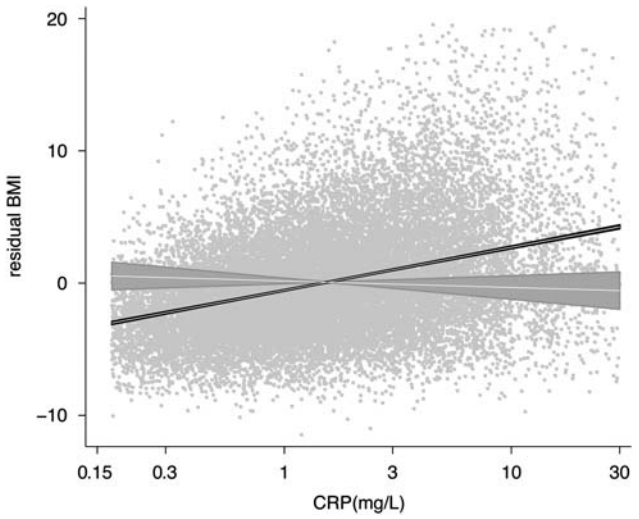

Figure 2 Comparison of linear relationships between residual BMI and circulating CRP observationally and when estimated using the *CRP* locus rs3091244 as an instrument for log-transformed CRP. *X* and *Y* axes represent CRP and residual BMI, respectively. Light grey points represent a scatter plot of the correlation between circulating CRP and residual BMI. Gray areas represent 95% confidence regions around instrumental variables estimates. Black area represents 95% confidence regions around simple linear regression estimates. (The 50 individuals with extreme residual BMI over 20 kg m^{-2} are not shown on the plot but were included in the analyses that gave the fitted lines and confidence regions.).

difference between predicted BMI (independent of variation attributable sex age and height) and observed BMI >20 kg m^{-2} were removed (because of their existence in the extreme tails of the BMI distribution).

Other covariates

Smoking and alcohol consumption were dichotomized and defined as 'ever' (ex-smoker or current smoker) versus 'never' smokers, and drinkers as those consuming >36 g alcohol per week. Other possible confounding factors related to social standing and educational attainment were recorded and incorporated into observational analyses. The two responses available for the assessment of these features were years of education completed and earned income at the date of examination. These responses formed the basis of the education and income variables used in further analyses which were coded as: education 0–9 years, 10–12 years, >13 years and annual income $<400\,000$ Kr, 400 000–600 000 Kr, $>600\,000$ Kr (100 000 Kr $\approx$ 13 000 Euro $\approx$ 17 000 USD).

Genotyping and selection of instruments

The ABI PRISM 7900HT Sequence Detection System (Applied Biosystems Inc., Foster City, CA, USA) was used to genotype the *FTO* locus rs9939609 and the *CRP* triallelic locus rs3091244 was scored using TaqMan (details available from authors). Genotyping was verified by DNA sequencing in >30 individuals with each genotype. As we performed reruns twice, $>99.9\%$ of all available participants were genotyped.

Using the same population studied here, we have previously shown that the *FTO* locus rs9939609 is associated with BMI, making this polymorphism suitable for a Mendelian randomization study similar to this study.[17] Similarly, we previously showed that the *CRP* triallelic promoter variant rs3091244 was associated with a 67% higher CRP level in the plasma for the rarest homozygote versus the most common homozygotes.[16]

In resequencing efforts at the *CRP* locus, Szalai *et al*.[18] assessed $\sim$1.2 kb of the *CRP* gene promoter, including rs3091244. Electrophoretic mobility shift assay confirmed that this single-nucleotide polymorphism was within an E-box regulatory factor-binding site (−394CACT*T*G-389) supporting hypotheses as to a functional role for this with respect to correlated variation in circulating levels of the CRP protein. We chose to restrict the primary analysis to this, the best apparently functional variant. Analyses used simple genotypic coding of this triallele and categorical analysis in the absence of an assumed genetic model; although in sensitivity analyses (not shown), other combinations were examined as described in the study by Zacho *et al*.[16]

Analyses

All data were gathered in a cross-sectional database summarizing individual characteristics at baseline collection. These data were transferred to Stata 10 (StataCorp LP, 2007; College station, TX, USA) for all analyses. Continuous effects were estimated using linear regression. Mean values for outcome variables by exposure group were estimated from linear regression models allowing for the incorporation of the covariates age and sex (descriptive analyses) and for sex, age, age squared, age–sex interaction, log(height), smoking, drinking, education and income (BMI–CRP associations).

Instrumental variable methods were used to obtain estimates of the directional effect of BMI on CRP and that of CRP on BMI.[19–21] The former was performed using *FTO*(rs9939609) as an instrument for BMI adjusting for sex, age, age squared, age–sex interaction, log(height), smoking, drinking, education and income. The latter was performed using the single-nucleotide polymorphism *CRP*(rs3091244) as an instrument for circulating CRP using the same covariates. The inclusion of baseline covariates not associated with instruments was undertaken to maximize the efficiency of instrumental variable regression models and to allow comparison of estimates from observational and instrumental variable analyses. We used the generalized method of moments with robust standard errors to fit the instrumental variable models in the main analyses, but checked results using limited information maximum likelihood and two-stage least squares. We compared the instrumental variable estimates with those from ordinary linear regression using the Durbin form of the Durbin–Wu–Hausman statistic.[22] We examined F-statistics

from first-stage regressions to evaluate the strength of the instruments. Values >10 are often considered to indicate the approximate validity of instrumental variable methods,[23,24] whereas values >30 are sufficient to ensure resulting estimates have under 5% bias and that tests for zero effect conducted at the 5% level have type I error rates no greater than 10%.[25]

Ethical approval

The studies were approved by Helve Hospital, by a Danish ethical committee (H-K 01-144/01), and were conducted according to the Declaration of Helsinki Principles. Written informed consent was obtained from participants. All participants were White and of Danish descent.

Results

There was an strong age-adjusted observational association between CRP and BMI among men and women as seen in Table 1. CRP levels were higher in women than in men across the BMI distribution, and across a large proportion of the BMI distribution, the association between CRP and BMI (as for residual BMI) was approximately linear (Supplementary Figure S1). In fully adjusted analyses, the change in logCRP per standard deviation increase in BMI can be summarized by a ratio of geometric means of 1.46 (95% confidence interval (95% CI): 1.45, 1.48) which approximates to a $0.71\,mg\,l^{-1}$ increase in circulating CRP. Sex-stratified analyses of this relationship showed a ratio of geometric means of 1.49 (95% CI: 1.47, 1.52), $\sim 0.76\,mg\,l^{-1}$ change in women and 1.42 (1.39, 1.45), $\sim 0.65\,mg\,l^{-1}$ change in men (P^{het} between men and women <0.0001). Conversely, if BMI was treated as the outcome, in fully adjusted analysis, there was a 1.06 (95% CI: 1.02, 1.09) $kg\,m^{-2}$ increase in BMI for a doubling in logCRP. In sex-stratified analyses, the same association was 1.26 (95% CI: 1.21, 1.31) $kg\,m^{-2}$ in women and 0.82 (0.77, 0.86) $kg\,m^{-2}$ men (P^{het} <0.0001).

In age- and sex-adjusted analyses, CRP levels were associated with sex, age, education, smoking, income and alcohol consumption (Table 2). BMI was associated with all of these, except smoking (Table 3). After stratification of these analyses by sex, largely consistent patterns were found. Out of these variables, the only relationships not shown to have strong evidence for association were those between the proportion with low income by quintile of CRP ($P=0.9$, males only) and the proportion with high income by quintile of BMI ($P=0.04$, males only) (Supplementary Tables S2–S5). All other features showed evidence for association ($P<0.005$). Furthermore, adjusting the associations between BMI and CRP for sex, age, age squared, age–sex interaction, log(height) and for confounders smoking, drinking, education and income had little effect on estimates (Table 1).

The *FTO*(rs9939609) genotype was associated with BMI in a manner expected, given previous work in this population[17]

Table 1 Age-adjusted means for CRP and BMI levels by decile of BMI

BMI/CRP by BMI decile	N	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	β	β′
BMI	23 073	20.20	22.20	23.39	24.40	25.35	26.34	27.43	28.72	30.55	35.26	1.134	1.137
		(20.12, 20.28)	(22.12, 22.28)	(23.31, 23.47)	(24.31, 24.48)	(25.27, 25.44)	(26.26, 26.43)	(27.35, 27.52)	(28.64, 28.81)	(30.47, 30.64)	(35.18, 35.34)	(1.129, 1.139)	(1.132, 1.142)
CRP	22 208	0.47	0.52	0.54	0.61	0.63	0.75	0.80	0.91	1.09	1.63		
		(0.44, 0.51)	(0.49, 0.56)	(0.51, 0.58)	(0.57, 0.65)	(0.59, 0.67)	(0.70, 0.80)	(0.75, 0.86)	(0.85, 0.98)	(1.02, 1.17)	(1.52, 1.75)		
BMIf	12 153	20.15	22.20	23.39	24.40	25.36	26.35	27.44	28.75	30.61	35.65	1.148	1.157
		(20.04, 20.26)	(22.09, 22.32)	(23.27, 23.51)	(24.29, 24.52)	(25.24, 25.49)	(26.22, 26.48)	(27.31, 27.58)	(28.62, 28.88)	(30.48, 30.74)	(35.53, 35.78)	(1.141, 1.155)	(1.150, 1.164)
CRPf	11 636	0.58	0.65	0.67	0.79	0.86	1.02	1.09	1.27	1.62	2.32		
		(0.53, 0.63)	(0.60, 0.71)	(0.61, 0.73)	(0.72, 0.87)	(0.78, 0.94)	(0.92, 1.12)	(0.98, 1.20)	(1.15, 1.40)	(1.47, 1.78)	(2.11, 2.55)		
BMIm	10 920	20.4	22.22	23.41	24.40	25.35	26.34	27.43	28.71	30.52	34.79	1.114	1.121
		(20.28, 20.52)	(22.11, 22.33)	(23.30, 23.52)	(24.29, 24.51)	(25.24, 25.46)	(26.24, 26.45)	(27.33, 27.53)	(28.60, 28.81)	(30.41, 30.62)	(34.68, 34.90)	(1.106, 1.121)	(1.113, 1.129)
CRPm	10 572	0.36	0.39	0.43	0.45	0.45	0.55	0.59	0.66	0.74	1.08		
		(0.32, 0.41)	(0.35, 0.43)	(0.39, 0.47)	(0.41, 0.49)	(0.401, 1.51)	(0.50, 0.60)	(0.54, 0.65)	(0.60, 0.73)	(0.67, 0.82)	(0.98, 1.20)		

Abbreviations: BMI, body mass index; CI, confidence interval; CRP, C-reactive protein. Means (95% CI) presented above are age adjusted and logCRP values are expressed as geometric means. 'f' and 'm' represent female- and male-specific values, respectively. β represents linear regression (95% CI)-derived β-coefficient expressed as a ratio of geometric means for the association between logCRP and decile of BMI. β′ indicates adjustment for sex, age, age squared, age–sex interaction, log(height), smoking, drinking, education and income.

Table 2 Age- and sex-adjusted relationships between confounders and quintiles of CRP

	CRP quintile					
Confounder	*1*	*2*	*3*	*4*	*5*	P-*value*
Mean age ($n=22\,304$)[a]	53.4 (53.0, 53.7)	57.0 (56.6, 57.4)	59.1 (58.7, 59.4)	59.9 (59.5, 60.3)	60.9 (60.5, 61.3)	<0.0001
% Male ($n=22\,304$)[b]	49.6 (48.4, 50.7)	48.7 (47.9, 49.5)	47.6 (46.9, 48.3)	46.3 (45.5, 47.1)	45.1 (44.0, 46.2)	<0.0001
% Ever smokers ($n=22\,304$)[c]	55.4 (54.2, 56.5)	59.3 (58.6, 60.2)	62.6 (62.0, 63.3)	65.5 (64.7, 66.2)	68.0 (66.9, 69.0)	<0.0001
% Ever drinkers ($n=22\,304$)[c]	74.6 (73.6, 75.6)	73.9 (73.2, 74.6)	72.1 (71.5, 72.7)	69.9 (69.2, 70.7)	66.9 (65.8, 68.0)	<0.0001
% Low income ($n=22\,017$)[c]	35.2 (34.0, 36.4)	45.9 (45.0, 46.8)	54.0 (53.2, 54.8)	59.3 (58.4, 60.2)	65.2 (64.0, 66.4)	<0.0001
% High income ($n=22\,017$)[c]	12.1 (11.3, 12.9)	8.8 (8.4, 9.3)	6.9 (6.5, 7.3)	5.7 (5.3, 6.1)	4.6 (4.2, 5.0)	<0.0001
% Low education ($n=22\,211$)[c]	23.7 (22.7, 24.8)	32.5 (31.6, 33.3)	39.7 (38.9, 40.4)	45.0 (44.1, 45.9)	5.1 (4.9, 5.2)	<0.0001
% High education ($n=22\,211$)[c]	13.5 (12.7, 14.4)	10.0 (9.5, 10.5)	8.1 (7.7, 8.5)	7.1 (6.7, 7.6)	6.2 (5.7, 6.7)	<0.0001

Abbreviations: CI, confidence interval; CRP, C-reactive protein. Smoking and drinking are binary variables and are coded as: smoking ever/never smoked, drinking >36g per week. High/low education and income are represented by upper and lower groups of the tripartite variables education = 0–9 years, 10–12 years, >13 years and income = <400 000 Kr, 400 000–600 000 Kr, >600 000 Kr. [a]Indicates sex-adjusted proportion (95% CI) of confounder by quintile of CRP. [b]Indicates age-adjusted proportion (95% CI) of confounder by quintile of CRP. [c]Indicates age- and sex-adjusted proportion (95% CI) of confounder by quintile of CRP.

Table 3 Age- and sex-adjusted relationships between confounders and quintiles of BMI

	BMI quintile					
Confounder	*1*	*2*	*3*	*4*	*5*	P-*value*
Mean age ($n=23\,073$)[a]	54.9 (54.5, 55.3)	56.7 (56.3, 57.1)	58.4 (58.0, 58.7)	59.8 (59.4, 60.2)	59.2 (58.8, 59.5)	<0.0001
% Male ($n=23\,073$)[b]	36.2 (35.1, 37.3)	41.6 (40.9, 42.4)	47.3 (46.64, 47.9)	53.0 (52.2, 53.8)	58.5 (57.4, 59.6)	<0.0001
% Ever smokers ($n=23\,073$)[c]	60.0 (58.9, 61.1)	61.5 (60.7, 62.3)	62.5 (61.9, 63.2)	63.0 (62.2, 63.8)	61.9 (60.8, 63.0)	0.3
% Ever drinkers ($n=23\,073$)[c]	73.3 (72.3, 74.3)	73.8 (73.1, 74.5)	73.3 (72.6, 73.9)	71.4 (70.7, 72.1)	66.4 (65.3, 67.5)	<0.0001
% Low income ($n=22\,780$)[c]	45.2 (43.9, 46.4)	47.8 (46.9, 48.8)	51.3 (50.5, 52.0)	55.2 (54.3, 56.1)	56.4 (55.2, 57.7)	<0.0001
% High income ($n=22\,780$)[c]	10.0 (9.3, 10.8)	8.8 (8.3, 9.3)	7.6 (7.2, 7.9)	6.4 (6.0, 6.8)	5.7 (5.2, 6.2)	<0.0001
% Low education ($n=22\,979$)[c]	23.0 (22.0, 24.0)	30.0 (29.2, 30.8)	37.8 (37.1, 38.5)	45.7 (44.8, 46.6)	49.5 (48.2, 50.8)	<0.0001
% High education ($n=22\,979$)[c]	13.7 (12.9, 14.6)	11.0 (10.5, 11.5)	8.7 (8.3, 9.1)	6.9 (6.5, 7.3)	6.1 (5.7, 6.7)	<0.0001

Abbreviations: BMI, body mass index; CI, confidence interval. Smoking and drinking are binary variables and are coded as: smoking ever/never smoked, drinking >36g per week. High/low education and income are represented by upper and lower groups of the tripartite variables education = 0–9 years, 10–12 years, >13 years and income = <400 000 Kr, 400 000–600 000 Kr, >600 000 Kr. [a]Indicates sex-adjusted proportion (95% CI) of confounder by quintile of BMI. [b]Indicates age-adjusted proportion (95% CI) of confounder by quintile of BMI. [c]Indicates age- and sex-adjusted proportion (95% CI) of confounder by quintile of BMI.

Table 4 Relationships between genotypic variation and BMI and circulating CRP

	FTO(rs9939609)				
	TT	*AT*	*AA*	*Per allele effect*	P-*value*
BMI	26.07 (25.98, 26.17)	26.37 (26.29, 26.45)	26.73 (26.59, 26.87)	0.32 (0.24, 0.40)	<0.0001
CRP	1.51 (1.48, 1.55)	1.55 (1.52, 1.58)	1.61 (1.56, 1.67)	1.03 (1.01, 1.05)	0.003

	CRP(rs3091244)							
	CC	*CT*	*TT*	*CA*	*AT*	*AA*	*Per allele effect*	P
BMI	26.32 (26.23, 26.41)	26.36 (26.27, 26.44)	26.24 (26.07, 26.42)	26.25 (26.02, 26.47)	26.29 (25.98, 26.61)	27.15 (26.02, 28.28)	−0.01 (−0.06, 0.04)	0.7
CRP	1.37 (1.34, 1.40)	1.61 (1.57, 1.64)	1.82 (1.74, 1.90)	1.71 (1.62, 1.81)	2.11 (1.95, 2.28)	2.56 (1.95, 3.37)	1.11 (1.10, 1.13)	<0.0001

Abbreviations: BMI, body mass index; CI, confidence interval; CRP, C-reactive protein. Means (95% CI) by genotypes with linear regression derived, per allele effect estimates (assuming additivity). CRP was log transformed for analyses; hence, geometric means are presented by genotype and ratios of geometric for effect estimates.

(Table 4), a finding that was similar in males and in females, as well as in different age groups (not shown). *FTO* variation was also associated with CRP levels, with genotypes associated with higher BMI being associated with higher CRP levels (ratio of geometric means from additive model 1.03 (95% CI: 1.01, 1.05), $P=0.003$). Variation at *CRP*(rs3091244) was strongly associated with CRP levels, again in the expected manner[16] (ratio of geometric means from additive model 1.11 (95% CI: 1.10, 1.13), $P<0.0001$) (Table 4). As for *FTO*(rs9939609), the effects were similar in both sexes and in different age groups (not shown). However, there was no association between the *CRP*(rs3091244) genotype and BMI.

Table 5 Observational and instrumental variable derived relationships between BMI and circulating CRP.

	Effect estimates				
Outcome/explanatory variable	*Observational*	*Instrumental variable*	P_{IV}	P_{diff}	F_{first}
CRP/BMI	1.46 (1.44, 1.48)	1.41 (1.10, 1.80)	0.006	0.8	31.1
BMI/CRP	1.03 (1.00, 1.07)	−0.24 (−0.58, 0.11)	0.2	<0.0001	57.3

Abbreviations: BMI, body mass index; CI, confidence interval; CRP, C-reactive protein. Observational analysis effects (95% CI) derived from linear regression adjusted for sex, age, age squared, age–sex interaction, log(height), smoking, drinking, education and income. CRP is log transformed for analyses above and effects on CRP are shown as ratios of geometric means for a s.d. increase in BMI. BMI effects are expressed as $kg\,m^{-2}$ for a doubling in logCRP. Instrumental variable derived estimates of the same effects include the same covariates. P_{IV} is the *P*-value from a test that the instrumental variable estimate is equal to the null. P_{diff} is the *P*-value from a test for difference between the observational and instrumental variable estimates. F_{first} is the first stage F-statistic from instrumental variable analysis.

There was no strong evidence of association between either *FTO*(rs9939609) or *CRP*(rs3091244) and confounding factors (Supplementary Table S1).

The joint associations of the *FTO* genotype and BMI, and the *FTO* genotype and CRP were used within an instrumental variable analysis to derive an estimate of the causal effect of BMI as an exposure on CRP as an outcome. In this instrumental variable analysis, the ratio of geometric means of circulating CRP per standard deviation (s.d.) increase in BMI was 1.41 (95% CI: 1.10, 1.80, $P=0.006$). In a test of equivalence between associations derived from observational analyses and those from instrumental variable analysis, there was no evidence for difference ($P^{diff}=0.8$). However, when *CRP* genotypes were used as an instrument for circulating CRP levels to estimate the causal effect of CRP level on BMI, observational associations were not corroborated. For a doubling in logCRP, the average change in BMI was estimated to be $-0.24\,kg\,m^{-2}$ (95% CI: −0.58, 0.11), $P=0.2$. In this case, a comparison of observational and instrumental variable estimates showed strong contrast (P^{diff} <0.0001). These results are summarized in Table 5. Instrumental variable analyses excluding baseline covariates show no substantive departures from reported effects (data not shown).

Figures 1 and 2 illustrate these findings graphically and note the contrast in effect estimates derived from instrumental variable analyses using *FTO* and *CRP* variations as instruments for BMI and CRP, respectively. These figures use the measure residual BMI to test for consistency in results when using a measure of BMI independent of sex, age and height.

These instrumental variable analyses were run separately for both sexes and in three different age groups, and generally consistent findings emerged (Supplementary Tables S6 and S7). In the case of age, wherein observational associations between BMI and circulating CRP diminish with age, instrumental variable-derived associations between BMI and CRP were similar across age strata.

Discussion

We used a reciprocal Mendelian randomization design to explore relationships between circulating CRP level and BMI. Not only was variation at *CRP* used to evaluate whether CRP has a causal effect on BMI, but simultaneously, variation at *FTO* was also used to evaluate whether BMI has a causal effect on CRP. This scheme is summarized in Figure 3 and exploits the availability of two independent instruments that yield nonconfounded estimates of causal effect. Overall, the use of variation at the *CRP* and *FTO* loci in reciprocal tests for a casual association between circulating CRP and BMI (in both basic genetic association and instrumental variable analyses) showed marked contrast and provided evidence that adiposity causally influences circulating CRP levels and not *vice versa*.

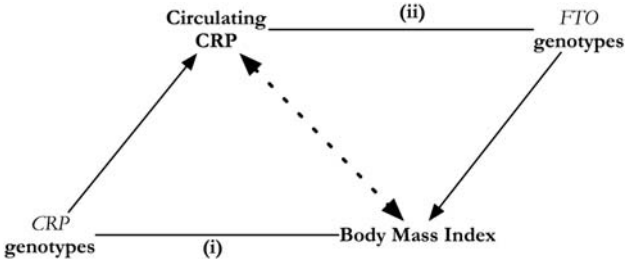

Figure 3 Graphical representation of the reciprocal Mendelian randomization framework used in main analyses. The dotted line represents the unknown direction of relationship between circulating CRP and BMI. Relationships (i) and (ii) denote the informative associations between *CRP* genotypes, *FTO* genotypes and circulating CRP and body mass index. Single-headed arrows represent the known (and assumed causal and largely nonconfounded) relationships between variation at the *CRP* and *FTO* loci and circulating CRP and body mass index, respectively.

Previous work has noted the consistent relationship between BMI and markers of chronic inflammation, including circulating CRP.[21,26–31] However, there have been questions regarding the causal nature of relationships between these measures and continued discussion of the appropriate interpretation of high-sensitivity CRP measures in a clinical setting.[32] Although the measurement of chronic elevation in circulating CRP has potential with respect to the refinement of prediction models for cardiovascular disease and end point events,[33] this is not dependent on causality (despite published assertions regarding this possibility).[5,34]

Work from studies such as the JUPITER trial[35] and from apparently independent assessment of CRP action[36,37] has continued to suggest that the role of CRP may be more than a marker of both BMI-derived and non-BMI-related cardiovascular risk. Difficulty clearly exists as to identifying a

causal role for CRP due in part to its confounded nature[38] and the potential for reverse causality.[21] In the absence of conventional randomization studies, reciprocal Mendelian randomization offers valuable insight. We consider the example presented—in which there is widespread but not universal consensus that the direction of causality runs from adiposity to CRP, not *vice versa*—as an illustration of a potentially highly valuable research strategy. Indeed in a number of other situations, it is already possible to identify genetic variants related to levels of associated pairs of factors in which the direction of causality requires elucidation and the increasing success of genome-wide association studies in identifying common variants related to intermediate phenotypes[39] will render this approach more generally applicable.

BMI as a causal agent

The available literature has confirmed the association between BMI and CRP and pointed toward BMI as a major contribution to observed variation in CRP levels in differing populations.[7,40–42] Along with this has also come recognition that the adipocyte itself is a key expressor of inflammatory molecules and that in situations of increased adiposity, levels of CRP expression are elevated.[1,8,40,42] This is not restricted to CRP; other inflammatory cytokines show elevated patterns of expression in situations of increased adiposity.[1,43] In addition to these biological lines of evidence suggesting that there is a direct link between increased levels of adiposity and changes in chronic inflammatory profiles, epidemiological evidence has shown similar patterns. Consistently, the adjustment of observational associations between circulating CRP and components of the metabolic syndrome for adiposity has led to an attenuation of these relationships. This suggests that although CRP is indeed marking the events leading to metabolic disturbance, it may not be driving them directly or independently.[1,44] However, the dilemma with observational data is that such statistical adjustments are highly dependent on the measurement characteristics of the confounders, and rely on the confounders being known and measurable.[9] Finally, available evidence about the impact weight loss has on circulating levels of CRP again favors the findings presented in this study. Weight loss interventions have been shown to decrease levels of circulating CRP and to improve the metabolic profile of those concerned.[45] Furthermore, investigation into the effects of weight reduction on the expression of inflammation-regulated loci has again shown that adiposity appears to be important in influencing levels of inflammatory proteins.[46]

Limitations to Mendelian randomization

There are several potential limitations to the application of Mendelian randomization, which have been discussed at length previously.[20,47] First, the biological consequences of variation at the *FTO* locus and the mechanism of the observed association of this with fat mass are still unclear. Several studies exist which point to a role for this locus in energy regulation and hypothalamically regulated patterns of appetite;[48–52] however, the possibility of complicating pleiotrophy in the action of *FTO* variation cannot be completely ruled out. In this case, utilizing multiple instruments—that is, multiple independent genetic variants that relate to an intermediate phenotype, such as BMI—can help strengthen causal inference, as pleiotropic effects are unlikely to influence the effects of each instrument in the same manner.[20] Second, the possibility of developmental plasticity altering the impact of chronic changes of metabolism delivered by heritable change (otherwise termed 'canalization'[53]) cannot be taken into account, although may influence the interpretation of findings.

Ultimately, such findings need to be interpreted against the background of other evidence on specific associations, but the reciprocal Mendelian randomization approach we advance in this study can provide powerful support to efforts at elucidating the direction of causal pathways. Through this, it offers to contribute to clinical and population-level efforts to improve health through modifying causal pathways leading to disease.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank Dorthe Uldall Andersen and Anne Bank for technical assistance with genotyping, staff and participants of the Copenhagen General Population for their important contribution and Chris Boustred for assistance in preparing this paper. *Funding*: NJT is funded by the MRC Centre for Causal Analyses in Translational Epidemiology grant number G0600705. GDS works within the MRC Centre for Causal Analyses in Translational Epidemiology, which is capacity funded by grant G0600705. RMH is supported in part by the MRC project grant G0601625. BGN is supported by the Copenhagen County Foundation and by Herlev Hospital, Copenhagen University Hospital. TMF is funded by the Peninsula College of Medicine and Dentistry, University of Exeter. The funders had no role in the study design, data collection, analysis, decision to publish or preparation of the paper.

Author contributions

NJT was involved in the concept and design of this study, undertook main analyses and writing of the paper. BGN is

the custodian of The Copenhagen General Population Study and was involved in the drafting of the paper. RMH was involved in statistical analysis and drafting of the paper. JZ was involved in data management and drafting of the paper. TMF was involved in initial design, concept and then drafting of the paper. ATH was involved in the running of The Copenhagen General Population Study and in drafting of the paper. GDS was involved in initial design, concept and then drafting of the paper. All authors had equal access to available data and take responsibility for data integrity.

References

1 Yudkin JS, Stehouwer CD, Emeis JJ, Coppack SW. C-reactive protein in healthy subjects: associations with obesity, insulin resistance, and endothelial dysfunction: a potential role for cytokines originating from adipose tissue? *Arterioscler Thromb Vasc Biol* 1999; **19**: 972–978.
2 Collaboration TERF. C-reactive protein concentration and risk of coronary disease, stroke, and mortality: an individual participant meta-analysis. *Lancet* 2010; **375**: 132–140.
3 Dandona P, Aljada A, Bandyopadhyay A. Inflammation: the link between insulin resistance, obesity and diabetes. *Trends Immunol* 2004; **25**: 4–7.
4 Dandona P, Weinstock R, Thusu K, Abdel-Rahman E, Aljada A, Wadden T. Tumor necrosis factor-alpha in sera of obese patients: fall with weight loss. *J Clin Endocrinol Metab* 1998; **83**: 2907–2910.
5 Bochud M, Fabienne M, Marques-Vidal P, Vollenweider P, Beckmann JS, Mooser V *et al*. Association between C-reactive protein and adiposity in women. *J Endocrinol Metab* 2009; AOP.
6 Schmidt MI, Duncan BB, Sharrett AR, Lindberg G, Savage PJ, Offenbacher S *et al*. Markers of inflammation and prediction of diabetes mellitus in adults (Atherosclerosis Risk in Communities study): a cohort study. *Lancet* 1999; **353**: 1649–1652.
7 Greenfield JR, Samaras K, Jenkins AB, Kelly PJ, Spector TD, Gallimore JR *et al*. Obesity is an important determinant of baseline serum C-reactive protein concentration in monozygotic twins, independent of genetic influences. *Circulation* 2004; **109**: 3022–3028.
8 Calabro P, Chang DW, Willerson JT, Yeh ETH. Release of C-reactive protein in response to inflammatory cytokines by human adipocytes: linking obesity to vascular inflammation. *J Am Coll Cardiol* 2005; **46**: 1112–1113.
9 Phillips AN, Smith GD. How independent are 'independent' effects? Relative risk estimation when correlated exposures are measured imprecisely. *J Clin Epidemiol* 1991; **44**: 1223–1231.
10 Davey Smith G. Reflections on the limitations to epidemiology. *J Clin Epidemiol* 2001; **54**: 325–331.
11 Davey Smith G, Ebrahim S. Mendelian randomisation: prospects, potentials and limitations. *Int J Epidemiol* 2004; **33**: 30–42.
12 Davey Smith G, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med* 2007; **4**: e352.
13 Nordestgaard BG, Benn M, Schnohr P, Tybjaerg-Hansen A. Nonfasting triglycerides and risk of myocardial infarction, ischemic heart disease, and death in men and women. *JAMA* 2007; **298**: 299–308.
14 Frikke-Schmidt R, Nordestgaard BG, Stene MCA, Sethi AA, Remaley AT, Schnohr P *et al*. Association of loss-of-function mutations in the ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. *JAMA* 2008; **299**: 2524–2532.
15 Schnohr P, Jensen JS, Scharling H, Nordestgaard BG. Coronary heart disease risk factors ranked by importance for the individual and community. A 21 year follow-up of 12 000 men and women from The Copenhagen City Heart Study. *Eur Heart J* 2002; **23**: 620–626.
16 Zacho J, Tybjaerg-Hansen A, Jensen JS, Grande P, Sillensen H, Nordestgaard BG. Genetically elevated C-reactive protein and ischaemic vascular disease. *N Engl J Med* 2008; **359**: 1897–1908.
17 Timpson NJ, Harbord R, Davey Smith G, Zacho J, Tybjaerg-Hansen A, Nordestgaard BG. Does greater adiposity increase blood pressure and hypertension risk? Mendelian randomisation using *FTO/MC4R* genotype. *Hypertension* 2009; **54**: 84–90.
18 Szalai AJ, Wu J, Lange EM, McCrory MA, Langefeld CD, Williams A *et al*. Single-nucleotide polymorphisms in the C-reactive protein (CRP) gene promoter that affect transcription factor binding, alter transcriptional activity, and associate with differences in baseline serum CRP level. *J Mol Med* 2005; **83**: 440–447.
19 Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000; **29**: 722–729.
20 Lawlor DA, Harbord RM, Sterne JAC, Timpson NJ, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008; **27**: 1133–1163.
21 Timpson NJ, Lawlor DA, Harbord R, Gaunt TR, Day IN, Palmer LJ *et al*. C-reactive protein and metabolic syndrome: Mendelian randomisation suggests associations are non-causal. *Lancet* 2005; **366**: 1954–1959.
22 Durbin J. Errors in variables. *Rev Int Stat Inst* 1954; **22**: 23–32.
23 Stock JH, Wright JH, Yogo M. A survey of weak instruments and weak identification in generalized method of moments. *J Bus Econ Stat* 2002; **20**: 518–529.
24 Staiger D, Stock JH. Instrumental variables regression with weak instruments. *Econometrica* 1997; **65**: 557–586.
25 Stock JH, Yogo M. Testing for weak instruments in linear IV regression. In: Andrews DWK, Stock JH (eds). *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge University Press: Cambridge, 2005. pp 80–108.
26 Frohlich M, Imhof A, Berg G, Hutchinson WL, Pepys MB, Boeing H *et al*. Association between C-reactive protein and features of the metabolic syndrome: a population based study. *Diabetes Care* 2000; **23**: 1835–1839.
27 Visser M, Bouter LM, McQuillan GM, Wener MH, Harris TB. Low-grade systemic inflammation in overweight children. *Pediatrics* 2001; **107**: E13.
28 Fantuzzi G. Adipose tissue, adipokines, and inflammation. *J Allergy Clin Immunol* 2005; **115**: 911–919.
29 Lowe GD, Rumley A, Wannamethee SG. Hemostatic abnormalities associated with obesity and the metabolic syndrome. *J Thromb Haemost* 2005; **3**: 1076–1078.
30 Ridker PM, Buring JE, Cook NR, Rifai N. C-reactive protein, the metabolic syndrome, and risk of incident cardiovascular events: an 8-year follow-up of 14 719 initially healthy American women. *Circulation* 2003; **107**: 391–397.
31 Rutter MK, Meigs JB, Sullivan LM, D'Agostino Sr RB, Wilson PW. C-reactive protein, the metabolic syndrome, and prediction of cardiovascular events in the Framingham Offspring Study. *Circulation* 2004; **110**: 380–385.
32 *Clin Chem* 2009; **55**. Special Issue—CRP.
33 Hingorani AD, Shah T, Casas JP, Humphries SE, Talmud PJ, Hingorani AD *et al*. C-reactive protein and coronary heart disease: predictive test or therapeutic target? *Clin Chem* 2009; **55**: 239–255.
34 Devaraj S, Singh U, Jialal I, Devaraj S, Singh U, Jialal I. The evolving role of C-reactive protein in atherothrombosis. *Clin Chem* 2009; **55**: 229–238.
35 Ridker PM, Danielson E, Fonseca FA, Genest J, Gotto Jr AM, Kastelein JJ *et al*. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N Engl J Med* 2008; **359**: 2195–2207.

36 Ridker PM, Rifai N, Pfeffer MA, Sacks F, Braunwald E. Long-term effects of pravastatin on plasma concentration of C-reactive protein. The Cholesterol and Recurrent Events (CARE) Investigators. *Circulation* 1999; **100**: 230–235.
37 Scirica BM, Cannon CP, Sabatine MS, Jarolim P, Sloane S, Rifai N *et al*. Concentrations of C-reactive protein and B-type natriuretic peptide 30 days after acute coronary syndromes independently predict hospitalization for heart failure and cardiovascular death. *Clin Chem* 2009; **55**: 265–273.
38 Ruckerl R, Peters A, Khuseyinova N, Andreani M, Koenig W, Meisinger C *et al*. Determinants of the acute-phase protein C-reactive protein in myocardial infarction survivors: the role of comorbidities and environmental factors. *Clin Chem* 2009; **55**: 322–335.
39 Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS *et al*. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009; **106**: 9362–9367.
40 Casas JP, Shah T, Hingorani AD, Danesh J, Pepys MB. C-reactive protein and coronary heart disease: a critical review. *J Inter Med* 2008; **264**: 295–314.
41 Park HS, Park JY, Yu R. Relationship of obesity and visceral adiposity with serum concentrations of CRP, TNF-alpha and IL-6. *Diabetes Res Clin Pract* 2005; **69**: 29–35.
42 Aronson D, Bartha P, Zinder O, Kerner A, Markiewicz W, Avizohar O *et al*. Obesity is the major determinant of elevated C-reactive protein in subjects with the metabolic syndrome. *Int J Obes Relat Metab Disord* 2004; **28**: 674–679.
43 Hotamisligil GS, Arner P, Caro JF, Atkinson RL, Spiegelman BM. Increased adipose tissue expression of tumor necrosis factor-alpha in human obesity and insulin resistance. *J Clin Invest* 1995; **95**: 2409–2415.
44 Kahn SE, Zinman B, Haffner SM, O'Neill MC, Kravitz BG, Yu D *et al*. Obesity is a major determinant of the association of C-reactive protein levels and the metabolic syndrome in type 2 diabetes. *Diabetes* 2006; **55**: 2357–2364.
45 Selvin E, Paynter NP, Erlinger TP. The effect of weight loss on C-reactive protein: a systematic review. *Arch Intern Med* 2007; **167**: 31–39.
46 Clement K, Viguerie N, Poitou C, Carette C, Pelloux V, Curat CA *et al*. Weight loss regulates inflammation-related genes in white adipose tissue of obese subjects. *FASEB J* 2004; **18**: 1657–1669.
47 Davey Smith G, Ebrahim S. 'Mendelian randomisation': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; **32**: 1–22.
48 Timpson NJ, Emmett P, Frayling TM, Rogers I, Hattersley AT, McCarthy MI *et al*. The FTO/obesity associated locus and dietary intake in children. *Am J Clin Nutr* 2008; **88**: 971–978.
49 Fischer J, Koch L, Emmerling C, Vierkotten J, Peters T, Bruning JC *et al*. Inactivation of the Fto gene protects from obesity. *Nature* 2009; **458**: 894–898.
50 Stratigopoulos G, Padilla SL, LeDuc CA, Watson E, Hattersley AT, McCarthy MI *et al*. Regulation of Fto/Ftm gene expression in mice and humans. *Am J Physiol Regul Integr Comp Physiol* 2008; **294**: R1185–R1196.
51 Cecil JE, Tavendale R, Watt P, Hetherington MM, Palmer CNA. An obesity-associated FTO gene variant and increased energy intake in children. *N Engl J Med* 2008; **359**: 2558–2566.
52 Gerken T, Girard CA, Tung YL, Webby CJ, Saudek V, Hewitson KS *et al*. The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* 2007; **318**: 1469–1472.
53 Waddington CH. Canalisation of development and the inheritance of acquired characters. *Nature* 1942; **150**: 563–565.

Supplementary Information accompanies the paper on International Journal of Obesity website (http://www.nature.com/ijo)