American Journal of Epidemiology
Vol. 173, No. 12
DOI: 10.1093/aje/kwr026
Advance Access publication:
May 9, 2011

Practice of Epidemiology

Instrumental Variable Estimation of Causal Risk Ratios and Causal Odds Ratios in Mendelian Randomization Analyses

Tom M. Palmer*, Jonathan A. C. Sterne, Roger M. Harbord, Debbie A. Lawlor, Nuala A. Sheehan, Sha Meng, Raquel Granell, George Davey Smith, and Vanessa Didelez

* Correspondence to Dr. Tom M. Palmer, MRC Centre for Causal Analyses in Translational Epidemiology, Oakfield House, Oakfield Grove, Bristol BS8 2BN, United Kingdom (e-mail: tom.palmer@bristol.ac.uk).

Initially submitted January 14, 2010; accepted for publication January 20, 2011.

In this paper, the authors describe different instrumental variable (IV) estimators of causal risk ratios and odds ratios with particular attention to methods that can handle continuously measured exposures. The authors present this discussion in the context of a Mendelian randomization analysis of the effect of body mass index (BMI; weight (kg)/height $(m)^2$) on the risk of asthma at age 7 years (Avon Longitudinal Study of Parents and Children, 1991–1992). The authors show that the multiplicative structural mean model (MSMM) and the multiplicative generalized method of moments (MGMM) estimator produce identical estimates of the causal risk ratio. In the example, MSMM and MGMM estimates suggested an inverse relation between BMI and asthma but other IV estimates suggested a positive relation, although all estimates had wide confidence intervals. An interaction between the associations of BMI and fat mass and obesity-associated (*FTO*) genotype with asthma explained the different directions of the different estimates, and a simulation study supported the observation that MSMM/MGMM estimators are negatively correlated with the other estimators when such an interaction is present. The authors conclude that point estimates from various IV methods can differ in practical applications. Based on the theoretical properties of the estimators, structural mean models make weaker assumptions than other IV estimators and can therefore be expected to be consistent in a wider range of situations.

causal inference; causality; confounding factors (epidemiology); effect modifiers (epidemiology); generalized method of moments; instrumental variables; Mendelian randomization analysis; structural models

Abbreviations: ACE, average causal effect; BMI, body mass index; COR, causal odds ratio; CRR, causal risk ratio; *FTO*, fat mass and obesity-associated gene; GMM, generalized method of moments; IV, instrumental variable; LSMM, logistic structural mean model; MGMM, multiplicative generalized method of moments; MSMM, multiplicative structural mean model; SMM, structural mean model.

The instrumental variable (IV) approach to causal inference has the potential to control for unmeasured confounding and reverse causation, which can bias results from standard epidemiologic analyses (1). As is the case for all approaches to causal inference, IV methods depend on assumptions, some of which are untestable, that vary between the different estimators. Randomization of individuals' genotypes at conception, under Mendel's first and second laws of genetics (''Mendelian randomization''), motivates the use of genotypes as IVs (2–5). Such analyses permit testing for a causal effect of a phenotype (modifiable risk factor or exposure) on the outcome and, under additional modeling assumptions, estimation of causal effects.

A number of IV estimators are available for situations where the outcome variable is continuous (6). These have well-documented statistical properties and are available in statistical software. However, IV estimation poses particular statistical challenges when the outcome variable is binary and the phenotype continuous and investigators wish to estimate either a causal risk ratio (CRR) or a causal odds ratio

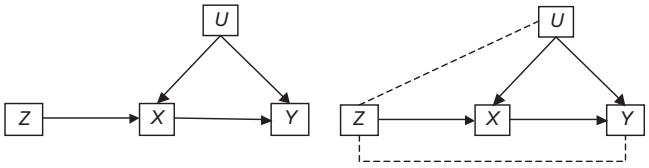

Figure 1. Directed acyclic graph (DAG) encoding the instrumental variable (IV) assumptions (left) and DAG encoding the IV assumptions with excluded associations shown by dotted lines (right). *U*, unmeasured confounders; *X*, phenotype; *Y*, outcome variable; *Z*, instrumental variable.

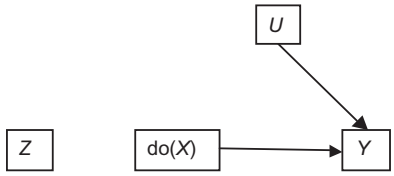

Figure 2. Directed acyclic graph representing the exclusion restriction—the instrumental variable assumptions under intervention in *X* (denoted do(*X*)).

(COR) (7). Various approaches have been proposed (8–11) and compared (12, 13) in the statistical, econometric, and epidemiologic literature, but these often consider only a binary phenotype. Given the increasing use of Mendelian randomization (14), epidemiologists need to be aware of the assumptions required for IV estimation of these causal parameters.

In this paper, we describe and compare IV estimators of the CRR and COR. We compare resulting estimates in the context of an example investigating the effect of body mass index (BMI; weight (kg)/height $(m)^2$) on the risk of asthma in children using fat mass and obesity-associated (*FTO*) genotypes as an instrument. We investigate reasons for differences between estimates using a simulation study based on the example.

THE IV ASSUMPTIONS

In this section, we describe aspects of causal inference relevant to IV analysis. We use the following notation: Y denotes the outcome variable, X the phenotype, Z the IV (genotype), U a set of unmeasured confounding variables, and p the probability of the outcome; we also define logit$(p) = \log(p/(1-p))$ and its inverse expit$(x) = \exp(x)/(1+\exp(x))$. The subscript i denotes an individual.

In the context of Mendelian randomization, the IV assumptions (7), which are common to all IV estimators, state that genotype should be

1) associated with the phenotype,
2) independent of the unmeasured confounding factors, and
3) independent of the outcome given the phenotype and unmeasured confounding factors.

The conditional independencies implied by these assumptions can be encoded in a directed acyclic graph (15) as shown on the left-hand side of Figure 1 (7). Assumption 1 is represented by the arrow between Z and X, while assumptions 2 and 3 are encoded by the absence of arrows. The dotted lines in the directed acyclic graph on the right-hand side of Figure 1 indicate associations (in either direction) or associations due to common causes that are excluded. If measured covariates are controlled for in the analysis, the IV assumptions are conditional on these.

To draw causal inferences, it is additionally necessary to make a "structural" assumption, which specifies how intervention on the phenotype operates on the system of variables (7, 16). In our context, this says that intervention does not affect genotype or the confounders and only affects the outcome through the changed value of the phenotype (7, 17). Using Pearl's do() operator to express the fact that X is *set* to a particular value as a result of intervention (18) and the symbol $\perp$ to denote conditional independence, the IV and structural assumptions imply (Figure 2) that

$$Z \perp Y | \text{do}(X). \quad (1)$$

The property in equation 1 is analogous to the "exclusion restriction" (17), which has also been described by Hernán and Robins (9) in terms of potential outcomes. The exclusion restriction implies the "conditional mean independence" assumption, that under intervention in X the mean of Y is independent of Z, which is a weaker form of the exclusion restriction and is sufficient for some estimation approaches.

For a continuous outcome, it is common to target an "average causal effect" (ACE)—the difference in the expected value of the outcome for a 1-unit difference in the phenotype:

$$\text{ACE}(x_0, x_0+1) = \text{E}(Y|\text{do}(X = x_0+1)) - \text{E}(Y|\text{do}(X = x_0)). \quad (2)$$

ACEs can also be estimated for binary outcomes, in which case they represent causal risk differences (1). For a binary outcome, the CRR for a 1-unit change in the phenotype is defined as the ratio of the probabilities of disease when X is set to x_0 and x_0+1:

$$\text{CRR}(x_0, x_0+1) = \frac{P(Y=1|\text{do}(X=x_0+1))}{P(Y=1|\text{do}(X=x_0))}. \quad (3)$$

Similarly, for a 1-unit change in the phenotype, the COR is defined as the ratio of the odds of disease when X is set to x_0 and x_0+1:

$$\text{COR}(x_0, x_0+1) = \frac{P(Y=1|\text{do}(X=x_0+1))\,P(Y=0|\text{do}(X=x_0))}{P(Y=0|\text{do}(X=x_0+1))\,P(Y=1|\text{do}(X=x_0))}. \quad (4)$$

The causal models discussed in this paper assume that their respective causal parameters are constant across values of x_0.

''Population causal effects'' are comparisons of the effect of setting X at different values for the whole population of interest (13). They are averaged over unobserved variables, particularly the confounders. For example, comparisons of the effect of treating all subjects compared with treating no subjects, obtained from a randomized trial with perfect compliance, are population causal effects. IV estimation of population effects typically relies on stronger assumptions than estimation of ''local causal effects,'' which are effects in specified population subgroups—for example, the effect of exposure on the exposed or treatment on the treated (9). The most common use of the term ''local'' is to refer to the ''complier causal effect,'' which compares treatment with control among persons who would always comply with assignment to treatment or control regardless of the actual assignment (17).

BINARY OUTCOME IV ESTIMATORS

We now describe IV estimators of the CRR and COR for a 3-level categorical instrument Z coded 0, 1, 2, denoting common homozygote, heterozygote, and rare homozygote genotypes, respectively; continuous phenotype X; and binary outcome Y. We denote log population CRR by θ and log population COR by ψ. We will make clear when methods estimate a corresponding local causal parameter. Estimates based on associations in the data will be described as ''associational,'' to distinguish them from estimates of causal parameters.

The Wald/ratio estimator

The ''Wald'' (19) or ''ratio'' estimator of the ACE for a 1-unit difference in X is defined as $\hat{\beta}_{YZ}/\hat{\beta}_{XZ}$, where $\hat{\beta}_{YZ}$ and $\hat{\beta}_{XZ}$ are the coefficients from linear regressions of Y on Z and X on Z, respectively. The estimator is consistent for the population ACE (that is, informally, the estimate will be close to the true value of its target parameter if the sample size is large) if the structural model for Y is linear in X and U and if U is not an effect modifier for the effect of X on Y. Under weaker assumptions (the additive structural mean model), this is consistent for a local ACE (9).

Several authors have suggested following the same principle to estimate the CRR (8) and the COR (20). Denoting the associational risk ratio (RR) and odds ratio (OR) between Y and Z by RR_{YZ} and OR_{YZ}, respectively, the estimators are

$$\hat{\theta} = \frac{\log(\mathrm{RR}_{YZ})}{\hat{\beta}_{XZ}} \tag{5}$$

and

$$\hat{\psi} = \frac{\log(\mathrm{OR}_{YZ})}{\hat{\beta}_{XZ}}. \tag{6}$$

Equation 5 is consistent for the population CRR if the structural model for Y is log-linear in X and U and if, additionally, X follows a linear model in Z and U and if U is not an effect modifier in any of these models (8, 13). Equation 6 is not consistent for the COR under any reasonable model, but it will approximate the CRR when the outcome is rare (13). The standard error of these estimators can be derived using Fieller's Theorem (20, 21) or a Taylor series expansion (22).

The 2-stage estimator

The first stage of the ''2-stage least squares'' estimator (23, 24) is a linear regression of X on Z, which generates predicted values $\hat{X}$. The second stage is a linear regression of Y on $\hat{X}$. This allows consistent estimation of the population ACE, assuming that, as with the ratio estimator, the structural model for Y is linear in X with no effect modification by U (13). For a single IV, the 2-stage least squares and ratio estimators are equivalent (17, 25).

For a binary outcome and a continuous phenotype, a 2-stage estimator of the population COR can be defined as

$$\text{Stage 1: obtain } \mathrm{X}_{\hat{i}} \text{ from } X_i = \alpha_0 + \alpha_1 Z_i + \varepsilon_i,\ \varepsilon_i \sim N\left(0, \sigma^2\right). \tag{7}$$

$$\text{Stage 2: estimate } \psi \text{ from } \mathrm{logit}(p_i) = \beta_0 + \psi \hat{X}_i,\ Y_i \sim \mathrm{Bern}(p_i). \tag{8}$$

The estimator $\hat{\psi}$ is the same as the ratio estimator from equation 6 and hence is also not generally consistent for the COR, but the bias may be reasonably small when the outcome is rare or when X is normally distributed and ψ is close to the null (13, 26–28). The 2-stage estimator $\hat{\theta}$ of the CRR is defined using a log-linear model at stage 2 and is the same as the ratio estimator of the population CRR from equation 5. Standard errors from the second-stage regression should be corrected to account for uncertainty in $\hat{X}$ (29)—for example, using the sandwich estimator available in standard statistical software (30).

The control function estimator

The ''control function'' estimator follows the same principle as the 2-stage estimator but additionally includes the estimated residuals from the first-stage regression in the second-stage regression (31–35). The rationale is that the first-stage residuals may be correlated with U, in which case they will help to control for the effect of U on Y. Therefore, the control function estimator targets a local causal effect conditional on U (the causal effect of X on Y within levels of U). The control function estimator of this local COR is defined as

$$\text{Stage 1: additionally estimate } \hat{\varepsilon}_i = X_i - \hat{X}_i. \tag{9}$$

$$\text{Stage 2: } \mathrm{logit}(p_i) = \beta_0 + \psi X_i + \beta_2 \hat{\varepsilon}_i. \tag{10}$$

An equivalent estimate is obtained with

$$\text{Stage 2: } \text{logit}(p_i) = \beta_0 + \psi\hat{X}_i + \beta'_2\hat{\varepsilon}_i, \text{ where } \beta'_2 = \psi + \beta_2. \quad (11)$$

For a linear model at stage 2, the control function estimator is equivalent to the 2-stage least squares estimator (25, 36, 37). The control function estimator of the CRR is obtained using a log-linear model in stage 2, which given a linear model at stage 1 is also the same as the ratio and 2-stage estimators of the population CRR. Because of the noncollapsibility of the odds ratio, this is not the case when a logistic regression is used at the second stage. In this latter case, when U is normally distributed and is not an effect modifier for the relation between X and Y, the control function estimator of the local COR is attenuated by the variance in U that is not explained by the first-stage residuals (28). The same holds on the probit scale if probit regression is used at the second stage (25, 38–40). As for the 2-stage estimator, standard errors from the second stage should be corrected to account for uncertainty in $\hat{X}$ (41).

Structural mean models

Structural mean models (SMMs) exploit IVs via G-estimation (42, 43). This involves finding the value of the causal parameter that fulfills the conditional mean independence assumption (9). In contrast to the estimators already described, G-estimation does not require specific distributional assumptions for the phenotype given Z and U. Under the conditional mean independence assumption, SMMs target a local causal effect of exposure on the exposed. In our case, this corresponds to the effect of setting X to a reference value, x_0, in all individuals. Under the stronger assumptions of no effect modification by U, which justifies the consistency of the ratio and 2-stage estimators, SMMs are also consistent for population causal effects.

Using an SMM with an identity link (sometimes referred to as the additive SMM) gives us the same estimator as the ratio and 2-stage estimators (9, 17, 44). We discuss the multiplicative structural mean model (MSMM) (9) and the logistic structural mean model (LSMM) (10, 11) below.

Multiplicative SMM. The MSMM assumes a log-linear structural model for the effect of X on Y with no effect modification by Z (9). The MSMM allows estimation of the local or population CRR by solving the estimating equation with respect to θ,

$$\sum_i Y_i \exp(-\theta X_i)(Z_i - \bar{Z}) = 0. \quad (12)$$

Details of the relation between the MSMM and the multiplicative generalized method of moments estimator (discussed below) are provided in Appendix 1.

Logistic SMM. Estimation of the COR using SMMs is more complex than estimation of the CRR (7), because G-estimation cannot be done within a single estimating equation (10, 45, 46). To overcome this problem, Vansteelandt and Goetghebeur (11) proposed an algorithm that fits an association model for the outcome, followed by the causal model.

The association model generates predicted probabilities $\hat{p}$ from a logistic regression of Y on X and Z, which are then used in an estimating equation for the target causal parameter. For the LSMM estimator to be consistent for the local COR, minimum requirements are that the association model contains an intercept, the unrestricted main effect of Z, and is estimated via maximum likelihood (11). As an example of such an algorithm, fit an association model to obtain estimated probabilities $\hat{p}_i$:

$$\text{logit}(p_i) = \beta_0 + \beta_1 X_i + \beta_2 Z_{1i} + \beta_3 Z_{2i}, \quad (13)$$

where, because of the second requirement, we replaced Z by 2 indicator variables Z_{1i} and Z_{2i} for the heterozygote and rare homozygote genotypes. The estimate of the COR is then obtained by solving the following estimating equation:

$$\sum_i \text{expit}(\text{logit}(\hat{p}_i) - \psi X_i)(Z_i - \bar{Z}). \quad (14)$$

The LSMM has been discussed in further detail by other authors (47–49).

In equations 12 and 14 and with SMMs in general, it is possible to use any function of the instrument to improve efficiency (11).

Generalized method of moments

The estimating equations for the MSMM and LSMM are ''moment conditions,'' since they set the sample version of an expectation to zero. Solution of these equations is feasible when the number of equations is equal to the number of parameters. Generalized method of moments (GMM) estimation allows for a greater number of equations (moment conditions) than parameters, known as overidentification (50). This occurs when multiple instruments are included in an analysis with a single phenotype. Here, we focus on the GMM estimator of the population CRR, using a multiplicative model (8).

The multiplicative GMM (MGMM) estimator assumes that the structural model for Y is log-linear in X and U (where U is not an effect modifier), which leads to the following estimating equations for the intercept α and the log CRR θ (8):

$$\sum_i (\exp(-\alpha - \theta X_i)Y_i - 1) = 0 \quad (15)$$

and

$$\sum_i (\exp(-\alpha - \theta X_i)Y_i - 1)Z_i = 0. \quad (16)$$

There are 2 equations because Z is accompanied by a vector of 1's to allow estimation of the intercept α. MGMM assumes that for any individual, a 1-unit change in X will have the same effect on the log risk of Y. MGMM is discussed in more detail elsewhere (26, 48, 51–54). Despite the different assumptions and target parameters of the MSMM and MGMM, the estimators are equal for a binary outcome (see

Table 1. Different Instrumental Variable Estimators and the Assumptions Required for Consistency

Estimator (Equation No.)	Target Parameter	Assumptions Required for Consistency
Ratio estimator (5)	Population CRR	Model for Y given do(X) and U is log-linear in X and U, without interaction; model for X given Z and U is linear without interaction, and X is approximately normally distributed (see reference 13).
Ratio estimator (6)	Population COR	Not generally consistent; approximately consistent for rare diseases under same assumptions as ratio estimator of the population CRR.
2-stage, logistic second stage (7, 8)	Population COR	Same as ratio estimator of population COR.
2-stage, log-linear second stage	Population CRR	Same as ratio estimator of population CRR.
Control function, logistic second stage (9, 10)	COR conditional on U	Generally not consistent, but converges to LSMM when X is normally distributed (see reference 47).
Control function, log-linear second stage	Population CRR	Same as 2-stage estimator with log-linear second stage.
MSMM (12)	CRR effect on exposed	Log-linear model for Y given do(X), X and Z, no effect modification by Z.
MSMM (12)	Population CRR	Log-linear model for Y given do(X) and U, no effect modification by U.
LSMM (13, 14)	COR effect on exposed	Logistic model for Y given do(X), X and Z, no effect modification by Z; association model for Y given X and Z has intercept, unrestricted main effect of Z and fitted by maximum likelihood.
MGMM (15, 16)	Population CRR	Same as MSMM estimator of the population CRR.

Abbreviations: COR, causal odds ratio; CRR, causal risk ratio; LSMM, logistic structural mean model; MGMM, multiplicative generalized method of moments; MSMM, multiplicative structural mean model.

Appendix 1 and Clarke and Windmeijer (49)). The MGMM estimator is implemented in Stata (Stata Corporation, College Station, Texas) with the command ivpois (55). We summarize the different IV estimators and the assumptions required for their consistency in Table 1.

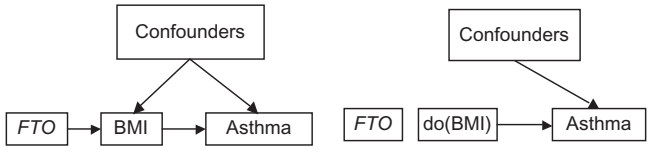

Figure 3. Directed acyclic graph (DAG) for Mendelian randomization analysis of fat mass and obesity-associated (*FTO*) genotype, body mass index (BMI; weight (kg)/height (m)^2), and asthma risk among children aged 7 years (left) and modified DAG under intervention in BMI, do(BMI) (right), Avon Longitudinal Study of Parents and Children, 1991–1992.

EXAMPLE: MENDELIAN RANDOMIZATION ANALYSIS OF THE CAUSAL EFFECT OF BMI ON ASTHMA RISK IN CHILDREN

We applied the estimators described above to a Mendelian randomization analysis using data from the Avon Longitudinal Study of Parents and Children (56) (www.bristol.ac.uk/alspac). In this population-based birth cohort study, investigators recruited 14,541 pregnant women with expected delivery dates between April 1991 and December 1992. A total of 13,988 infants survived to at least 1 year of age.

We targeted the causal effect of BMI, assessed at age 7 years, on the risk of physician diagnosis of asthma. A population effect considers the effect on asthma of setting BMI to x, compared with $x + 1$, for all children in the population. A local effect considers the effect, for a given child, of changing BMI from its observed level to the reference level (here we use the sample mean BMI of 16.1). Genotypes of the rs9939609 polymorphism in the *FTO* gene were used as an IV; this polymorphism is robustly associated with childhood and adult BMI and obesity (57). The 2 alleles of this *FTO* polymorphism are denoted A and T, where A is the risk allele associated with greater BMI, fat mass, and increased obesity. We assumed an additive genetic model for *FTO* genotypes. All analyses were carried out in 4,647 children with complete data on asthma, BMI, and *FTO*, of whom 649 (14%) had asthma. Analyses were performed using Stata, version 11.0 (58). The IV model is shown in the directed acyclic graph in Figure 3.

Assessment of IV assumptions

We investigated the extent to which *FTO* genotype was associated with BMI (IV assumption 1) using the first stage of the 2-stage estimator (equation 7). The mean increase in BMI per risk allele was 0.15 (95% confidence interval: 0.07, 0.23); this effect was small in relation to the standard deviation of BMI of 1.95. The R^2 and F statistics from this regression were 0.003 and 12.7, respectively. Although the F statistic was greater than the commonly used weak instrument threshold of 10 (59), the R^2 showed that *FTO* explained only 0.3% of the variation in BMI.

It is not strictly possible to test assumptions 2 and 3, as they involve unobservable variables (7, 9). We can find

Table 2. Distribution of Asthma and Possible Confounders by Fat Mass and Obesity-Associated (*FTO*) Genotype (rs9939609) in Children Aged 7 Years, Avon Longitudinal Study of Parents and Children, 1991–1992

	Total No.	TT		AT		AA		*P* Value From χ^2 Test
		No.	%	No.	%	No.	%	
No. and % of participants	4,647	1,699	37	2,220	48	728	16	0.95[a]
Asthma (yes)	4,647	234	13.8	302	13.6	113	15.5	0.41
Female sex	4,647	832	49	1,070	48	386	53	0.08
Low birth weight	4,594	75	4	80	4	36	5	0.21
Parental education (less than university degree)	4,593	893	54	1,214	56	390	55	0.44
Prenatal smoking	4,579	404	24	562	26	167	23	0.30
Postnatal smoking	4,407	270	17	390	19	115	17	0.23
Low parental social class	3,974	211	15	295	15	82	13	0.41

[a] Test for Hardy-Weinberg equilibrium.

support for assumption 2 by investigating whether *FTO* genotype is independent of measured covariates that might confound the association between BMI and asthma. Results from these analyses provided little evidence for such associations (Table 2). The plausibility of assumption 3 would ideally be justified by biologic knowledge of the functionality of the *FTO* gene, but research on this topic is not yet completed (60). Finally, there was some evidence that the association of BMI with asthma was stronger in girls than in boys ($P = 0.044$ for interaction in a logistic regression), which would violate the assumption of no effect modification underlying all estimators; from similar checks, we found little evidence of effect modification by other measured covariates.

Table 3. Instrumental Variable Estimates of the Causal Odds Ratio and Causal Risk Ratio for the Effect of Body Mass Index on Asthma Risk, Avon Longitudinal Study of Parents and Children, 1991–1992

	COR or CRR	95% CI
Standard logistic regression analysis		
Unadjusted odds ratio	1.06	1.02, 1.10
Adjusted[a] odds ratio	1.08	1.03, 1.13
Wald/ratio estimator[b]		
CRR	1.37	0.64, 2.96
COR	1.45	0.65, 3.43
2-stage estimator[c]		
CRR	1.37	0.68, 2.78
COR	1.45	0.64, 3.29
Control function[c]		
CRR	1.37	0.68, 2.76
COR	1.44	0.63, 3.28
Logistic structural mean model[d]		
COR	1.64	0.29, 9.31
Multiplicative structural mean model[d]		
CRR	0.81	0.44, 1.48
Multiplicative generalized method of moments[d]		
CRR	0.81	0.44, 1.48

Abbreviations: CI, confidence interval; COR, causal odds ratio; CRR, causal risk ratio.

[a] Adjusted for sex, birth weight, prenatal maternal smoking, postnatal maternal smoking, maternal education, and head-of-household social class.

[b] The 95% CI was based on the delta method standard error of the ratio of 2 means.

[c] The 95% CIs were based on robust standard errors.

[d] The 95% CIs were based on bootstrapped standard errors.

IV estimates of the causal effect of BMI on asthma risk

The associational estimates of the COR obtained from standard logistic regression models corresponded to 6% and 8% increases in the odds of asthma per 1-unit increase in BMI in analyses unadjusted and adjusted for possible confounders, respectively (Table 3). These estimates had narrow confidence intervals in comparison with the wide confidence intervals about all IV estimates, the latter being due to the small proportion of the variation in BMI explained by *FTO*.

Under IV assumptions, a test of the instrument-outcome association is a test for the presence of a causal effect of the phenotype on the outcome. The *FTO*-asthma odds ratio was 1.06 ($P = 0.372$), so there was no strong evidence against the null hypothesis of no causal effect of BMI on asthma. The ratio estimate of the COR was 1.45, identical to the 2-stage estimate as expected. The control function estimate of the COR was 1.44. These estimates had comparably wide confidence intervals. The ratio estimate of the CRR was 1.37, identical to the 2-stage and control function estimates and also with wide confidence intervals. The LSMM estimate of the COR was 1.64 with a very wide confidence interval.

The MSMM and MGMM estimates of the CRR were equal, as we expected and as we show in Appendix 1, but were in the opposite direction (CRR = 0.81) to the other IV estimates. The bootstrapped confidence intervals of these estimates spanned the null and overlapped considerably with those of the other estimates. We also estimated the MGMM confidence interval using a robust asymptotic standard error, which was slightly narrower (95% confidence interval: 0.63, 1.05).

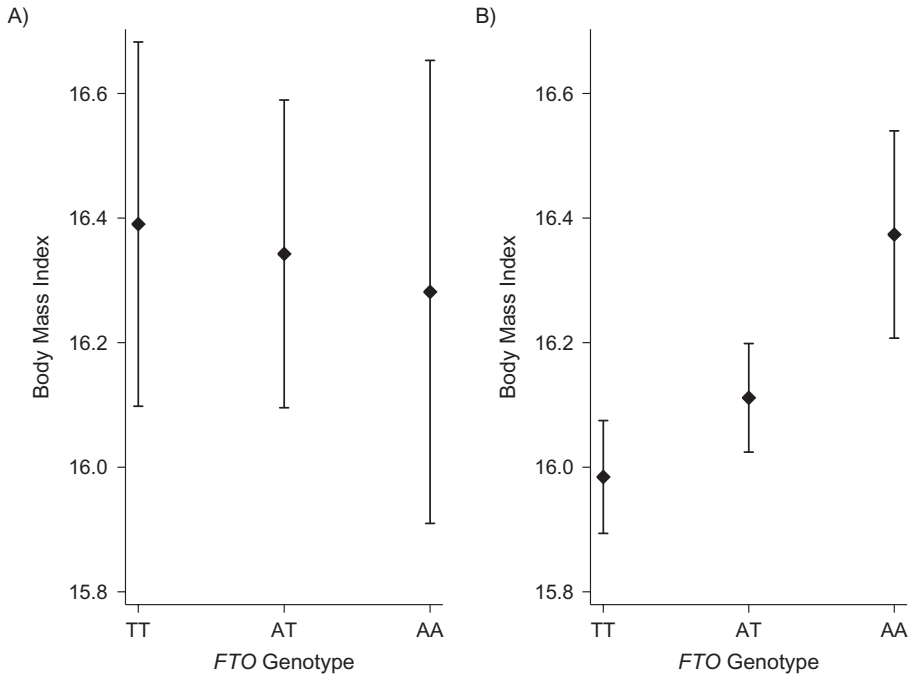

Figure 4. Mean body mass index (weight (kg)/height (m)^2), denoted by diamonds, according to fat mass and obesity-associated (*FTO*) genotype (rs9939609) for A) asthmatic and B) nonasthmatic children aged 7 years, Avon Longitudinal Study of Parents and Children, 1991–1992. Bars, 95% confidence interval.

The different direction of effect of the MSMM/MGMM estimates

For the ratio estimator, the *FTO*-BMI association (the denominator in equations 5 and 6) is estimated in the whole sample and is positive. Similarly, the 2-stage and control function estimators base the predictions of BMI from genotype on the whole sample (equation 7). In contrast, the MSMM/MGMM estimator is based on a single model for the joint relation between genotype, BMI, and asthma. Nonasthmatics (those for whom $y = 0$) only contribute to the mean genotype $\bar{Z}$ in the MGMM/MSMM estimating equations (equations 12, 15, and 16).

Therefore, we examined the associational relation between BMI, *FTO* genotype, and asthma status (Figure 4). Mean BMI increased from the TT genotype to the AA genotype in nonasthmatics but decreased in asthmatics, indicating an interaction between the associations of *FTO* and BMI with asthma status. We also fitted the logistic regression of asthma status on BMI, *FTO*, and their interaction, which gave some evidence of an associational interaction ($P = 0.038$). None of the structural assumptions of any IV estimators considered here imply the absence of such an associational interaction, which therefore does not imply that any of the IV assumptions is violated. This associational interaction could result from an interaction between *FTO* and unobserved confounders (as modeled in our simulation study described below) or could be a chance finding in this particular data set.

To investigate this issue further, we performed simulations in which we compared the MGMM and 2-stage estimators of the CRR, in data that were generated with and without an interaction between the causal effects of *FTO* and an unmeasured confounder U on BMI. This induces an associational interaction between *FTO* and BMI with asthma. Simulations were performed both under the null and with a small positive causal effect of BMI on asthma. Full details are given in Appendix 2. In scenarios 1 and 2, with the interaction, we found a negative correlation between the MGMM and 2-stage estimates (Table 4). In scenarios 3 and 4, without the interaction, there was a positive correlation between the estimates. In the scenarios with the interaction, a greater proportion of the MGMM and 2-stage estimates were on opposite sides of the true causal effect than in scenarios without the interaction. In scenario 2, with the causal effect and interaction, the 2-stage estimator is not consistent because the assumption of no interaction between the effects of the instrument and unmeasured confounders is violated. Corresponding confidence intervals from the 2-stage estimator had low coverage (67%), whereas the MGMM-based confidence interval had approximately correct coverage (92%). Both approaches had approximately correct coverage in the other scenarios.

DISCUSSION

IV methods applied in the context of Mendelian randomization can be used to estimate CRRs or CORs while avoiding bias due to uncontrolled confounding and/or reverse causality. We have described theoretical properties and

Table 4. Results of Simulations Comparing the Multiplicative Generalized Method of Moments and 2-Stage Estimators of the Causal Risk Ratio

	2-Stage Estimate for Log CRR (MCE)	MGMM Estimate for Log CRR (MCE)
Scenario 1: no causal effect with interaction		
Mean bias	−0.007 (0.0046)	0.009 (0.0094)
MSE	0.021 (0.0010)	0.088 (0.0042)
Coverage	0.952 (0.0068)	0.964 (0.0059)
Correlation between estimates	−0.23	
% of estimates on opposite sides of the CRR of 1	64.1	
Scenario 2: causal effect with interaction		
Mean bias	−0.206 (0.0042)	−0.146 (0.0100)
MSE	0.060 (0.0019)	0.120 (0.0055)
Coverage	0.674 (0.0148)	0.919 (0.0086)
Correlation between estimates	−0.12	
% of estimates on opposite sides of the CRR of 1.2	35.9	
Scenario 3: no causal effect with no interaction		
Mean bias	−0.005 (0.0049)	−0.001 (0.0053)
MSE	0.024 (0.0010)	0.029 (0.0018)
Coverage	0.942 (0.0074)	0.964 (0.0059)
Correlation between estimates	0.88	
% of estimates on opposite sides of the CRR of 1	7.3	
Scenario 4: causal effect with no interaction		
Mean bias	0.003 (0.0043)	0.003 (0.0049)
MSE	0.018 (0.0009)	0.024 (0.0014)
Coverage	0.954 (0.0066)	0.964 (0.0059)
Correlation between estimates	0.82	
% of estimates on opposite sides of the CRR of 1.2	15	

Abbreviations: CRR, causal risk ratio; MCE, Monte Carlo error; MGMM, multiplicative generalized method of moments; MSE, mean squared error.

assumptions of various such estimators. We found that there are essentially 2 classes of estimators: those that make distributional assumptions about the exposure/phenotype (ratio, 2-stage, control function) and those that avoid such assumptions (SMMs).

We demonstrated the equivalence of the MSMM and MGMM estimators of the CRR, which has been noted previously (48, 49). Additional IV estimators exist: for example, Robins and Rotnitzky (46) proposed an alternative estimator of the COR based on an SMM. Estimators using a probit link have convenient mathematical properties but do not easily lead to an estimate of the CRR or COR (39). GMM estimators of the CRR and COR using an additive moment condition exist, but the underlying models seem less plausible than those of the MGMM estimator (12, 48, 51, 61, 62).

We compared the IV estimators in an example data set and found that the MSMM and MGMM estimates of the CRR were below 1, whereas the other estimates of the CRR and COR were above 1. We explained this through an associational interaction between *FTO* and BMI with asthma in our data. This interaction may have arisen by chance, or it could have been induced by an interaction between the effects of *FTO* genotype and an unmeasured confounding variable on BMI. In simulations including such an interaction, we found a negative correlation between the MGMM and 2-stage estimates. Since an associational interaction is not excluded by the IV assumptions, it is not the case, as some authors have suggested (12), that different IV estimators will always estimate the same direction of effect. Despite the striking differences in both the magnitude and direction of the different IV estimates, their confidence intervals overlapped considerably.

Although we focused on methods that can handle continuous exposures/phenotypes, epidemiologists are often interested in using IV methods for binary exposure variables. A particular example is when the instrument is randomization to one of 2 treatments, and a binary X represents the treatment that is actually received (17). In this situation, the ratio, 2-stage, and control function estimators are not consistent for any causal risk ratio or odds ratio and thus should not be used (13). The MSMM/MGMM estimators do not make distributional assumptions about X and can be used to estimate causal effects such as the effect of treatment on the treated or exposure on the exposed. An alternative approach targets causal effects of X on Y within the latent (unobservable) class of compliers. Consistent estimation of complier causal effects on the risk difference or risk ratio scale is straightforward under the ''monotonicity'' assumption that there are no defiers (persons who only take treatment when randomized to the control group) (9, 26, 27); estimation of complier causal effects on the odds ratio scale is more problematic (49). The standard intention-to-treat estimate of the risk difference or risk ratio will point in the same direction as the effect of treatment on the compliers, providing that there are no defiers.

All IV estimators presented here rely on a version of a structural ''no effect modification'' assumption involving the unobserved confounders, rather than the ''no defiers'' assumption used to justify estimation of complier causal effects. Both sets of assumptions are impossible to test from

observable data and are therefore problematic to justify in practice. For continuous X, one might formulate a monotonicity assumption for the effect of genotype on phenotype—for example, that increasing a person's number of *FTO* risk alleles would not decrease his or her BMI. However, to our knowledge, the monotonicity assumption does not allow us to construct estimators of CRRs or CORs when the phenotype is continuous.

A limitation of our analysis is that we focused on the consistency of the various IV estimators. In practice, efficiency and finite sample-size coverage probabilities of confidence intervals are of major importance and deserve further investigation. In this context, it seems promising to consider the use of covariates to increase the efficiency of IV estimators (11). When covariates are included in IV analyses, the IV assumptions must hold, conditional on the covariates. Adjustment for covariates in IV analyses is common in econometrics (41) and might be relevant to Mendelian randomization analyses when a covariate is both a confounder and suspected of modifying the effect of phenotype on outcome.

In summary, investigators using IV methods to estimate causal risk ratios or odds ratios should be clear about the assumptions they are prepared to make. We identified situations in which different IV estimators are negatively correlated despite the validity of the IV assumptions; it follows that point estimates from different IV estimators can differ in practical applications. Based on the theoretical properties of the estimators, SMMs make weaker assumptions than ratio, 2-stage, and control function estimators and can therefore be expected to be consistent in a wider range of situations.

ACKNOWLEDGMENTS

Author affiliations: MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom (Tom M. Palmer, Debbie A. Lawlor, George Davey Smith); School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom (Jonathan A. C. Sterne, Roger M. Harbord, Raquel Granell); Departments of Health Sciences and Genetics, College of Medicine, Biological Sciences and Psychology, University of Leicester, Leicester, United Kingdom (Nuala A. Sheehan, Sha Meng); and School of Mathematics, Faculty of Science, University of Bristol, Bristol, United Kingdom (Vanessa Didelez).

This work was funded by the Medical Research Council (MRC) Collaborative Project (grant G0601625). Drs. Tom M. Palmer, Debbie A. Lawlor, and George Davey Smith are employed by the MRC Centre for Causal Analyses in Translational Epidemiology, which funds Dr. Palmer's salary (MRC grant G0600705). The MRC, the Wellcome Trust, and the University of Bristol provide core funding support for the Avon Longitudinal Study of Parents and Children (ALSPAC).

The authors thank Drs. Frank Windmeijer and Paul Clarke, Centre for Market and Public Organisation, University of Bristol, for very helpful comments. The authors also thank Dr. John Henderson, School of Social and Community Medicine, University of Bristol, for advice regarding the analysis of asthma data from ALSPAC. The authors are extremely grateful to the midwives for their help in recruiting the families and to the entire ALSPAC research team.

The views expressed in this paper are those of the authors and not necessarily those of any funding body or others whose support is acknowledged. The funders had no role in study design, data collection and analysis, the decision to publish, or preparation of the manuscript.

Conflict of interest: none declared.

REFERENCES

1. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000;29(4):722–729.
2. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1–22.
3. Thomas DC, Conti DV. Commentary: the concept of 'Mendelian randomization'. *Int J Epidemiol*. 2004;33(1):21–25.
4. Davey Smith G, Lawlor DA, Harbord R, et al. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med*. 2007;4(12):e352. (doi: 10.1371/journal.pmed.0040352).
5. Davey Smith G. Mendelian randomization for strengthening causal inference in observational studies: application to gene by environment interaction. *Perspect Psychol Sci.* 2010;5(5):527–545.
6. Lawlor DA, Harbord RM, Sterne JA, et al. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27(8):1133–1163.
7. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res.* 2007;16(4):309–330.
8. Mullahy J. Instrumental-variable estimation of count data models: applications to models of cigarette smoking behaviour. *Rev Econ Stat*. 1997;79(4):568–593.
9. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology.* 2006;17(4):360–372.
10. Robins JM, Rotnitzky A, Scharfstein DO. Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In: Halloran ME, Berry D, eds. *Statistical Models in Epidemiology: The Environment and Clinical Trials*. New York, NY: Springer Publishing Company; 1999:1–92.
11. Vansteelandt S, Goetghebeur E. Causal inference with generalized structural mean models. *J R Stat Soc Series B Stat Methodol*. 2003;65(4):817–835.
12. Rassen JA, Schneeweiss S, Glynn RJ, et al. Instrumental variable analysis for estimation of treatment effects with dichotomous outcomes. *Am J Epidemiol*. 2009;169(3):273–284.
13. Didelez V, Meng S, Sheehan NA. Assumptions of IV methods for observational epidemiology. *Stat Sci.* 2010;25(1):22–40.
14. Sheehan NA, Didelez V, Burton PR, et al. Mendelian randomisation and causal inference in observational epidemiology. *PloS Med*. 2008;5(8):e177. (doi: 10.1371/journal.pmed.0050177).
15. Cowell RG, Dawid AP, Lauritzen SL, et al. *Probabilistic Networks and Expert Systems*. New York, NY: Springer-Verlag New York, Inc; 1999.

16. Pearl J. *Causality: Models, Reasoning, and Inference*. Cambridge, United Kingdom: Cambridge University Press; 2000.
17. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc.* 1996; 91(434):444–455.
18. Pearl J. Causal diagrams for empirical research. *Biometrika*. 1995;82(4):669–688.
19. Wald A. The fitting of straight lines if both variables are subject to error. *Ann Math Stat*. 1940;11(3):284–300.
20. Thompson JR, Tobin MD, Minelli C. *GE1: On the Accuracy of Estimates of the Effect of Phenotype on Disease Derived From Mendelian Randomisation Studies*. (Technical report 2003/GE1). Leicester, United Kingdom: Department of Health Sciences, University of Leicester; 2003. (http://www2.le.ac.uk/departments/health-sciences/research/ships/gen-epi/reports/2003_GE1). (Accessed January 13, 2011).
21. Fieller EC. Some problems in interval estimation. *J R Stat Soc Series B Stat Methodol*. 1954;16(2):175–185.
22. Thomas DC, Lawlor DA, Thompson JR. Re: Estimation of bias in nongenetic observational studies using ''Mendelian triangulation'' by Bautista et al. [letter]. *Ann Epidemiol*. 2007;17(7):511–513.
23. Basmann RL. A generalized classical method of linear estimation of coefficients in a structural equation. *Econometrica*. 1957;25(1):77–83.
24. Theil H. *Repeated Least Squares Applied to Complete Equation Systems*. The Hague, the Netherlands: Central Planning Bureau; 1953.
25. Wooldridge JM. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, United Kingdom: MIT Press; 2002.
26. Angrist JD. Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice. *J Bus Econ Stat*. 2001;19(1):2–28.
27. Abadie A. Semiparametric instrumental variable estimation of treatment response models. *J Econom*. 2003;113(2):231–263.
28. Palmer TM, Thompson JR, Tobin MD, et al. Adjusting for bias and unmeasured confounding in Mendelian randomization studies with binary responses. *Int J Epidemiol*. 2008;37(5): 1161–1168.
29. Greene WH. *Econometric Analysis*. 6th ed. New York, NY: Prentice Hall; 2008.
30. White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*. 1980;48(4):817–838.
31. Garen J. The returns to schooling: a selectivity bias approach with a continuous choice variable. *Econometrica*. 1984;52(5): 1199–1218.
32. Nagelkerke N, Fidler V, Bernsen R, et al. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. *Stat Med*. 2000;19(14): 1849–1864.
33. Blundell RW, Powell JL. Endogeneity in nonparametric and semiparametric regression models. In: Dewatripont M, Hansen LP, Turnovsky SJ, eds. *Advances in Economics and Econometrics: Theory and Applications. 8th World Congress of the Econometric Society.* Cambridge, United Kingdom: Cambridge University Press; 2003:312–357.
34. Wooldridge JM. On two stage least squares estimation of the average treatment effect in a random coefficient model. *Econ Lett.* 1997;56(2):129–133.
35. Terza JV, Basu A, Rathouz PJ. Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *J Health Econ*. 2008;27(3):531–543.
36. Dhrymes PJ. *Econometrics: Statistical Foundations and Applications*. New York, NY: Harper and Row; 1970.
37. Nichols A. Weak instruments: an overview and new techniques [abstract]. Presented at the 5th North American Stata Users Group Meeting, Boston, Massachusetts, July 24–25, 2006. (http://www.stata.com/meeting/5nasug/wiv.pdf). (Accessed January 13, 2011).
38. Smith RJ, Blundell RW. An exogeneity test for a simultaneous equation tobit model with an application to labor supply. *Econometrica*. 1986;54(3):679–685.
39. Rivers D, Vuong QH. Limited information estimators and exogeneity tests for simultaneous probit models. *J Econometrics*. 1988;39(3):347–366.
40. Blundell RW, Smith RJ. Estimation in a class of simultaneous equation limited dependent variable models. *Rev Econ Stat*. 1989;56(1):37–57.
41. Cameron AC, Trivedi PK. *Microeconometrics: Methods and Applications*. New York, NY: Cambridge University Press; 2005.
42. Robins JM. The analysis of randomized and nonrandomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In: Sechrest L, Freeman H, Mulley A, eds. *Health Service Research Methodology: A Focus on AIDS*. Washington, DC: National Center for Health Services Research, US Public Health Service; 1989:113–159.
43. Robins JM. Correcting for non-compliance in randomized trials using structural nested mean models. *Commun Stat Theory Methods*. 1994;23(8):2379–2412.
44. Fischer-Lapp K, Goetghebeur E. Practical properties of some structural mean analyses of the effect of compliance in randomized trials. *Control Clin Trials*. 1999;20(6):531–546.
45. Robins JM. Causal inference from complex longitudinal data. In: Berkane M, ed. *Latent Variable Modeling and Applications to Causality.* (Lecture Notes in Statistics 120). New York, NY: Springer-Verlag New York, Inc; 1997:69–117.
46. Robins JM, Rotnitzky A. Estimation of treatment effects in randomised trials with non-compliance and dichotomous outcomes using structural mean models. *Biometrika*. 2004; 91(4):763–783.
47. Bowden J, Vansteelandt S. Mendelian randomization analysis of case-control data using structural mean models. *Stat Med*. 2011;30(6):678–694.
48. Clarke P, Windmeijer F. *Instrumental Variable Estimators for Binary Outcomes.* (Working Paper no. 09/209). Bristol, United Kingdom: Centre for Market and Public Organisation, University of Bristol; 2009. (http://www.bristol.ac.uk/cmpo/publications/papers/2009/wp209.pdf). (Accessed March 31, 2011).
49. Clarke PS, Windmeijer F. Identification of causal effects on binary outcomes using structural mean models. *Biostatistics*. 2010;11(4):756–770.
50. Hansen LP. Large sample properties of generalized method of moments estimators. *Econometrica*. 1982;50(4):1029–1054.
51. Windmeijer FA, Santos Silva JM. Endogeneity in count data models: an application to demand for health care. *J Appl Econom*. 1997;12(3):281–294.
52. Wooldridge JM. Quasi-likelihood methods for count data. In: Pesaran MH, Schimdt P, eds. *Handbook of Applied Econometrics. Volume II: Microeconomics*. Oxford, United Kingdom: Blackwell Publishing Ltd; 1997.
53. Windmeijer F. *ExpEnd, A Gauss Programme for Non-Linear GMM Estimation of Exponential Models With Endogenous Regressors for Cross Section and Panel Data*. (Working Paper CWP14/02). London, United Kingdom: Centre for Microdata Methods and Practice, Institute for Fiscal Studies; 2002.

54. Windmeijer F. *GMM for Panel Count Data Models*. (Working Paper CWP21/06). London, United Kingdom: Centre for Microdata Methods and Practice, Institute for Fiscal Studies; 2006.
55. Nichols A. *IVPOIS: Stata Module to Estimate an Instrumental Variables Poisson Regression via GMM*. Washington, DC: Urban Institute; 2007. (http://ideas.repec.org/c/boc/bocode/s456890.html). (Accessed January 13, 2011).
56. Golding J, Pembrey M, Jones R, et al. ALSPAC—The Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinatal Epidemiol*. 2001;15(1):74–87.
57. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*. 2007;316(5826):889–894.
58. Stata Corporation. *Stata Statistical Software, Release 11.1*. College Station, TX: Stata Corporation; 2010.
59. 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(4):518–529.
60. Han Z, Niu T, Chang J, et al. Crystal structure of the FTO protein reveals basis for its substrate specificity. *Nature*. 2010;464(7292):1205–1209.
61. Foster EM. Instrumental variables for logistic regression: an illustration. *Soc Sci Res*. 1997;26(4):487–504.
62. Johnston KM, Gustafson P, Levy AR, et al. Use of instrumental variables in the analysis of generalized linear models in the presence of unmeasured confounding with applications to epidemiological research. *Stat Med*. 2008;27(9):1539–1556.

APPENDIX 1

We aim to show that the sample moment conditions of the multiplicative structural mean model (MSMM) and multiplicative generalized method of moments (MGMM) estimators are equivalent for a binary outcome. Firstly, we examine the sample moment condition of the MGMM estimator,

$$\sum_i (\exp(-\alpha - \theta X_i)Y_i - 1)Z_i = 0.$$

Since y_i is either 0 or 1, this becomes

$$\begin{aligned}&\sum_{i:y=1} (\exp(-\alpha)\exp(-\theta X_i)Z_i) - n\bar{Z} = 0\\ &\Leftrightarrow \sum_{i:y=1} \exp(-\theta X_i)Z_i = n\bar{Z}\exp(\alpha). \qquad \text{(A1)}\end{aligned}$$

Secondly, we examine the sample moment condition for the MSMM estimator with a standardized instrument,

$$\sum_i Y_i \exp(-\theta X_i)(Z_i - \bar{Z}) = 0.$$

Again, using the fact that y_i is either 0 or 1, we have

$$\begin{aligned}&\sum_{i:y=1} \exp(-\theta X_i)(Z_i - \bar{Z}) = 0\\ &\Leftrightarrow \sum_{i:y=1} \exp(-\theta X_i)Z_i = \bar{Z}\sum_{i:y=1} \exp(-\theta X_i). \qquad \text{(A2)}\end{aligned}$$

We then note that the left-hand sides of expressions A1 and A2 are equal. Hence, to complete the proof, we need to show that the right-hand sides of these expressions are also equal. We continue by equating the right-hand side of these expressions,

$$\begin{aligned}&n\bar{Z}\exp(\alpha) = \bar{Z}\sum_{i:y=1} \exp(-\theta X_i)\\ &\Leftrightarrow \exp(\alpha) = \frac{1}{n}\sum_{i:y=1} \exp(-\theta X_i). \qquad \text{(A3)}\end{aligned}$$

In order for the proof to be complete, we need to show that expression A3 is true. We argue that this is the case, because we can derive this expression for $\exp(\alpha)$ using the MGMM moment condition alone. We show this by noting that for the MGMM estimator, the instrument is actually a vector with the first term a constant 1 to allow estimation of the intercept α. Hence, for $z_i = 1$, the MGMM moment condition becomes

$$\begin{aligned}&\sum_{i:y=1} (\exp(-\alpha - \theta X_i) - 1) = 0\\ &\Leftrightarrow \exp(-\alpha)\sum_{i:y=1} \exp(-\theta X_i) - n = 0\\ &\Leftrightarrow \exp(-\alpha) = \frac{1}{n}\sum_{i:y=1} \exp(-\theta X_i). \qquad \text{(A4)}\end{aligned}$$

Expressions A3 and A4 are identical, which completes the proof. Hence, we conclude that for a binary outcome, the MSMM and MGMM moment conditions are equivalent, and therefore these estimators give the same estimates of the causal risk ratio.

APPENDIX 2

The following details the simulations used to compare the 2-stage and multiplicative generalized method of moments (MGMM) estimators of the causal risk ratio. In the following, the 3-category instrument, continuous phenotype, binary outcome, and continuous unmeasured confounder are denoted *Z*, *X*, *Y*, and *U*, respectively. We performed simulations for 4 scenarios. In scenarios 1 and 2, we simulated a structural interaction between the effects of fat mass and obesity-associated (*FTO*) genotype and the unmeasured confounder *U* on body mass index, which induces an associational interaction between the effects of *FTO* and body mass index on asthma. Scenarios 3 and 4 were simulated without the interaction. Scenarios 1 and 3 were

generated with no causal effect, and scenarios 2 and 4 were simulated with a causal risk ratio of 1.2.

Following the frequencies in the data, Z was simulated with 3 categories, 0, 1, and 2, with frequencies 37%, 47%, and 16%, respectively. Where the subscript i denotes an observation, the other variables were simulated as follows:

$u_i \sim \text{Uniform}(-1,1)$.

Scenarios 1 and 2: $x_i \sim \text{Normal}(15.5 + 0.4u_i + 0.25z_{1i} + 0.4z_{2i} - 0.9u_i.z_1 - 1.5u_i.z_2, 1.9^2)$.

Scenarios 3 and 4: $x_i \sim \text{Normal}(15.5 + 0.4u_i + 0.25z_{1i} + 0.4z_{2i}, 1.9^2)$.

X was then standardized using $s_i = (x_i - 15.5)/1.9$.

Scenarios 1 and 3: $p_i = \exp(-2.5 + u_i)$.

Scenarios 2 and 4: $p_i = \exp(-2.5 + u_i + 0.35s_i)$.

$y_i \sim \text{Bernoulli}(p_i)$.

For each scenario, we simulated data sets with 20,000 observations and fitted the 2-stage and MGMM estimators of the causal risk ratio to each data set. The simulations were repeated for 1,000 replications. In scenarios 1 and 2, the ratio of odds ratios for the interaction between the associations of X and Z with Y was 0.93, as compared with 1.00 (no interaction) in scenarios 3 and 4.