

## References

- Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;**32**:1–22.
- Burgess S, Timpson NJ, Ebrahim S, Davey Smith G. Mendelian randomization: where are we now and where are we going? *Int J Epidemiol* 2015;**44**:379–88.
- Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr* 2016;**103**:965–78.
- Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 2016;**40**:304–14.
- Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;**44**:512–25.
- Lango Allen H, Estrada K, Lettre G *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010;**467**:832–38.
- Davies NM, von Hinke Kessler Scholder S, Farbmacher H, Burgess S, Windmeijer F, Davey Smith G. The many weak instruments problem and Mendelian randomization. *Stat Med* 2015;**34**:454–68.
- Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013;**3**:658–65.
- Wood AR, Esko T, Yang J *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 2014;**46**:1173–86.
- Kemp JP, Sayers A, Davey Smith G, Tobias JH, Evans DM. Using Mendelian randomization to investigate a possible causal relationship between adiposity and increased bone mineral density at different skeletal sites in children. *Int J Epidemiol* 2016;**45**:1560–72.
- Locke AE, Kahali B, Berndt SI *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;**518**:197–206.
- Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol* 2013;**178**:1177–84.
- Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol* 2015;**30**:543–52.

## Authors' response to Hartwig and Davies

International Journal of Epidemiology, 2016, 1678–1679

doi: 10.1093/ije/dyw241

Advance Access Publication Date: 20 September 2016



From John P Kemp,<sup>1,2</sup> Adrian Sayers,<sup>3</sup> George Davey Smith,<sup>2</sup> Jonathan H Tobias<sup>3</sup> and David M Evans<sup>1,2,\*</sup>

<sup>1</sup>University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, QLD, Australia, <sup>2</sup>MRC Integrative Epidemiology Unit and <sup>3</sup>School of Clinical Sciences, University of Bristol, Bristol, UK

\*Corresponding author. MRC Integrative Epidemiology Unit at the University of Bristol, School of Social and Community Medicine, Oakfield Grove, Bristol BS8 2BN, UK. E-mail: Dave.Evans@bristol.ac.uk

In their letter, Hartwig and Davies<sup>1</sup> raise an important issue that was not discussed in the original Mendelian randomization Egger regression (MR Egger) paper by Bowden *et al.*<sup>2</sup> Hartwig and Davies point out that, similar to other varieties of MR, MR-Egger is also susceptible to weak instrument bias. In the case of single-sample MR-Egger, this means that estimates of the causal effect may be biased towards the observational association when weak instruments are used in the analysis (as is the case with traditional single-sample MR). Hartwig and Davies also point out a potential solution to this problem, the utilization of more precise externally derived estimates of the relevant single nucleotide polymorphism (SNP)-exposure association (i.e. from larger publicly available genome-wide association meta-analyses).

We agree with Hartwig and Davies' conclusions that weak instrument bias is a problem in MR-Egger as it is in traditional MR analyses, and that it is critically important that users of the technique are aware of this possibility. Hartwig and Davies also refer to a recent study of ours<sup>3</sup> in the *International Journal of Epidemiology*, where we used several different types of MR analyses (including MR-Egger) to examine a possible causal effect of adiposity on bone mineral density (BMD). As Hartwig and Davies fairly acknowledge in their letter, not only did we discuss the possibility of weak instrument bias in our study, we also performed preliminary simulations to investigate its effect on MR-Egger (the results of which broadly agree with their assertions). As Hartwig and Davies note, MR-Egger was only one small component of our paper and none of our key results (which we believe

to be robust) rely on the results of these analyses in isolation, and indeed many of the other analyses reported in our paper do not suffer from potential bias due to weak instruments.

Hartwig and Davies did, however, suggest that we could have used estimates from an external source to obtain less biased results in our MR-Egger analyses. Whereas we agree that this would be good practice in most situations, we do not feel that it would have been appropriate in our study, for two reasons. First, the focus of our article was not on a possible causal relationship between body mass index and BMD (which is well-known and widely accepted), but rather on a possible causal relationship between adiposity [as operationalized as fat mass calculated from total body dual-energy X-ray absorptiometry (DXA)] and BMD. There are no publicly available genome-wide association studies of total body fat mass as measured by total body DXA, and therefore no external estimates that we could have applied in our analyses (i.e. as far as we are aware, we are currently the largest such study). We could have used external estimates for analyses involving body mass index, but this

would have been of limited utility since body mass index is a far from perfect measure of adiposity. Second, our study involved 9-year-old children from the Avon Longitudinal Study of Parents and Children. It is unclear the extent to which effect sizes of adiposity-associated variants in adults reflect effect sizes of adiposity-associated variants in children (as Hartwig and Davies recognize), and we therefore feel it would have been inappropriate to use adult-derived external estimates in our study of children.

## References

1. Hartwig F, Davies N. Why internal weights should be avoided (not only) in MR Egger regression. *Int J Epidemiol* 2016;**45**:1676–78.
2. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;**44**:512–25.
3. Kemp JP, Sayers A, Davey Smith G, Tobias JH, Evans DM. Using Mendelian randomization to investigate a possible causal relationship between adiposity and increased bone mineral density at different skeletal sites in children. *Int J Epidemiol* 2016;**45**: 1560–72.

## Response to Hartwig and Davies

From Jack Bowden,<sup>1,2\*</sup> Stephen Burgess<sup>3</sup> and George Davey Smith<sup>1</sup>

<sup>1</sup>MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK, <sup>2</sup>MRC Biostatistics Unit, Cambridge Institute of Public Health, Cambridge, UK and <sup>3</sup>Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

\*Corresponding author. MRC Integrative Epidemiology Unit, University of Bristol, School of Social and Community Medicine, Oakfield Grove, Bristol BS8 2BN, UK. E-mail: jack.bowden@bristol.ac.uk

*International Journal of Epidemiology*, 2016, 1679–1680

doi: 10.1093/ije/dyw252

Advance Access Publication Date: 20 September 2016



Thank you for the opportunity to respond to the thoughtful and timely letter by Hartwig and Davies.<sup>1</sup> They do indeed raise a very important practical issue with the implementation of MR-Egger regression in the single-sample setting, or with the use of so called ‘internal’ weights, namely weak instrument bias. They are right to point out the unsatisfactory nature of our analysis of the height data in our original publication,<sup>2</sup> and that our use of weak instruments had the likely effect of biasing the MR-Egger estimate towards that of the observational association. Their re-analysis of these data with external weights appears to provide a much more satisfactory answer and, when such weights are available, it is both a simple and an attractive way to circumvent the problem.

Although previous simulation studies have highlighted this fact, further research is needed to completely understand the issue of weak instrument bias for MR-Egger in

the single-sample context. What is clear however, is that the standard notion of instrument strength, as quantified by the F statistic, cannot naively be applied to estimate the magnitude of this bias; new (or at least newly borrowed) theory is required. Before covering initial progress in this vein, we now briefly discuss related (and more mature) work in the two-sample context.

## Recent work on weak instrument bias in the two-sample context

A strength of MR-Egger regression, along with the weighted median<sup>3</sup> and inverse-variance weighted (IVW)<sup>4</sup> methods, is that they do not require gene, exposure and outcome data on a single sample of subjects at the individual level. MR-Egger regression can be implemented with only summary data estimates of the single nucleotide polymorphism (SNP)-exposure and SNP-outcome associations, making it most