CORRECTION

Correction to: Is population structure in the genetic biobank era irrelevant, a challenge, or an opportunity?

Daniel John Lawson[1] · Neil Martin Davies[1] · Simon Haworth[1] · Bilal Ashraf[1] · Laurence Howe[2] · Andrew Crawford[1] · Gibran Hemani[1] · George Davey Smith[1] · Nicholas John Timpson[1]

Correction to: Human Genetics
https://doi.org/10.1007/s00439-019-02014-8

In the original article publication, there is an incorrect impression that Fig. 1 formed a formal Directed Acyclic Graph (DAG) by describing it as a causal model. However, it was not correct if interpreted in this way. The correct Fig. 1 is with a model that is correct when interpreted as a DAG, by making the differences between the *model representation of ancestry* and the *true ancestry*, which is always unobserved, explicit. We stress that there are other ways to introduce these biases, and it need not act through the separation between observed and true ancestry.
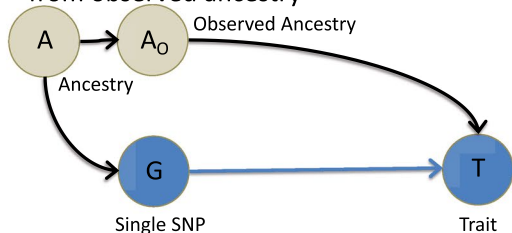

The original article can be found online at https://doi.org/10.1007/s00439-019-02014-8.
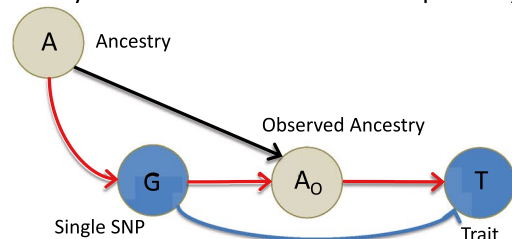

✉ Daniel John Lawson
dan.lawson@bristol.ac.uk

1   MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK

2   Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, Gower Street, London WC1E 6BT, UK
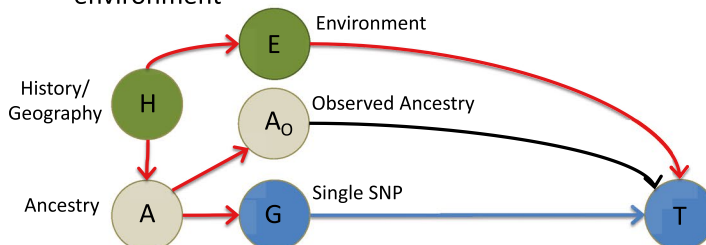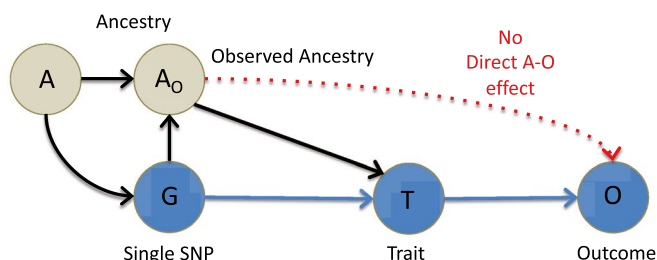
Fig. 1 Causal models including ancestry for the effect of a SNP (G) on a trait (T). The true ancestry (A) is never observed, but is measured through A_0. **a** Correction for stratification will be accurate when observed ancestry (A_0) is confounding T. **b** Correction for structure may give biased inference when observed ancestry A_0 is associated with the causal pathway by which the SNP acts, leading to a correlation in the effect of the SNP on ancestry and the trait. For example, the same SNPs are predictive of height and ancestry. Alternatively, ancestry may be associated with a trait T_A (e.g. skin tone) that causes a trait (e.g. skin cancer). **c** Correction for structure will be incomplete when ancestry is associated with the environment (E) due to shared history and geography (H), for example $T =$ BMI with $E =$ diet choice. **d** Correction for structure when using causal inference is robust to complexity, provided the assumptions of Mendelian randomization (see text) are met; particularly all remaining effects of ancestry go through the trait (T) so there is no direct effect of ancestry (A) on the outcome (O)