Mendelian randomization: a powerful method to determine causality of biomarkers in diseases

Vinicius Tragante[1], Folkert W. Asselbergs[1,2,3,4]

[1] Department of Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht University, the Netherlands;
[2] Durrer Center for Cardiovascular Research, Netherlands Heart Institute, Utrecht, the Netherlands;
[3] Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, United Kingdom;
[4] Farr Institute of Health Informatics Research and Institute of Health Informatics, University College London, London, United Kingdom

Assessing the influence of a marker or exposure on a given outcome is a challenging task, since a multitude of factors can influence levels of given marker, and therefore confound correlation-causation inferences. Notable exceptions are genetic markers, since those are innate and not bound to change through an individual's lifetime. Furthermore, inheritance of each trait is independent of every other according to Mendel's law of inheritance [1]. That's the principle behind Mendelian randomization (MR), a method proposed in the 80's by Katan [2] to understand if alterations in plasma cholesterol levels in cancer patients were the cause or consequence of the disease (later shown through MR that there is no causal relation of cholesterol in cancer development) [3]. More recently the method has become ever more popular due to the continuously increasing availability of cohorts with genetic information, well-determined levels of biomarkers and deep-phenotyped outcome data. A simple PubMed query for the term "Mendelian randomization" shows a continuous increase from one publication in 2003 to 252 papers in 2017, demonstrating that the method has become a common practice in the field of genetic epidemiology.

Yeoung and Schooling [ref from this issue] report in this issue of *International Journal of Cardiology* a bi-directional MR to study the correlation between adiponectin levels and coronary artery disease (CAD). Taking advantage of the fact that many genetic markers have been significantly and independently associated with adiponectin levels and CAD, it is possible to group signals of one trait and test against the other with the possibility of giving a more definitive answer about the causality between adiponectin and CAD in both directions (a graphical demonstration of the method is in Figure 1). Using only significant and independent SNPs as genetic instrument, the researchers found no causal association of adiponectin on CAD (OR 1.02, 95% CI 0.85 - 1.24), confirming a previous study using a smaller sample size and set of SNPs [4]. Results were achieved using the MR-Egger method [5], which accounts for possible pleiotropic effects of the SNPs. The approach is recommended, given several SNPs associated with adiponectin are also associated with obesity, which in turn is a known risk factor for CAD. In addition, no causal association of CAD on adiponectin

levels was found either (beta 0.030, 95% CI -0.041 - 0.102), using a subset of 28 SNPs (out of 56 CAD-associated SNPs by the time of submission) available in the biggest GWAS on adiponectin to date.

These results provide further evidence that there is no causal effect of adiponectin on CAD development, and therefore, adiponectin does not seem a good target for drug development with the goal of preventing CAD. Such information is valuable for decision-making before running clinical trials that may lead to negative results. Indeed, in a textbook example, Voight *et al.* [6] demonstrated through MR that when using genetic information as instrument to assess the effect of HDL levels in development of CAD, there was no significant association between those genetic proxies for HDL and CAD. The pharma industry would have benefited from this information if it were available before clinical trials of CETP inhibitor torcetrapib had to be halted due to an increase in cardiovascular mortality [7]. Turns out that therapies aimed to increase HDL levels will not lead to reduction in CAD risk despite the strong inverse association between HDL-cholesterol and CAD in epidemiological studies. With proper MR being performed beforehand, such failures may be avoided, reducing the costs and risks of drug development.

In conclusion, MR studies are helping identify and confirm or reject incidental causation inferences found in clinical trials, animal models and other studies. Given the availability of well-powered genetic studies and the identification of thousands of genetic markers associated with a myriad of traits and diseases, one can expect a fast increase in identification of measurable biomarkers, via MR, that contribute to disease, hopefully leading to better therapies to the most appropriate targets. This approach has attracted the interest of the pharmaceutical industry, with the intention of finding suitable candidates for drug repurposing and development. An example comes from EU-funded project BigData@Heart [8], within the Innovative Medicines Initiative framework, which has partners from academia and industry working together to identify novel targets using MR.

Acknowledgements

Folkert W. Asselbergs is supported by UCL Hospitals NIHR Biomedical Research Centre.

References

1 Smith, G. D. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* **32**, 1-22 (2003).

2 Katan, M. B. Apolipoprotein E isoforms, serum cholesterol, and cancer. *Lancet* **1**, 507-508 (1986).

3 Trompet, S. *et al.* Apolipoprotein e genotype, plasma cholesterol, and cancer: a Mendelian randomization study. *Am J Epidemiol* **170**, 1415-1421, doi:10.1093/aje/kwp294 (2009).
4 Borges, M. C. *et al.* Role of Adiponectin in Coronary Heart Disease Risk: A Mendelian Randomization Study. *Circ Res* **119**, 491-499, doi:10.1161/CIRCRESAHA.116.308716 (2016).
5 Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* **44**, 512-525, doi:10.1093/ije/dyv080 (2015).
6 Voight, B. F. *et al.* Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* **380**, 572-580, doi:10.1016/S0140-6736(12)60312-2 (2012).
7 Johns, D. G., Duffy, J., Fisher, T., Hubbard, B. K. & Forrest, M. J. On- and off-target pharmacology of torcetrapib: current understanding and implications for the structure activity relationships (SAR), discovery and development of cholesteryl ester-transfer protein (CETP) inhibitors. *Drugs* **72**, 491-507, doi:10.2165/11599310-000000000-00000 (2012).
8 Anker, S. *et al.* Big Data in Cardiovascular Disease. *Eur Heart J* **38**, 1863-1865, doi:10.1093/eurheartj/ehx283 (2017).

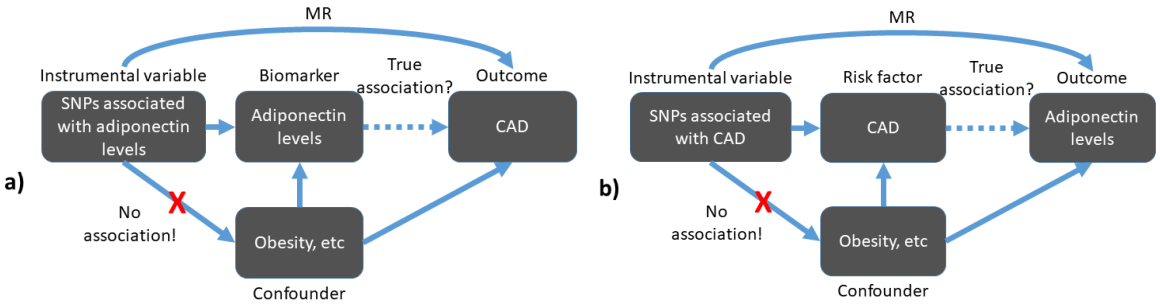

Figure 1. Graphical representation of the MR approach. In a), SNPs significantly associated with alterations in adiponectin levels are used as an instrument to causally associate with incidence of CAD. SNPs that influence adiponectin levels are not confounded by any other factors, while the biomarker can be influenced by other risk factors, such as obesity. In b), the same logic is employed in the opposite direction, in other words, SNPs associated with incidence of CAD are used as instrument to test the causal association with adiponectin levels.