Understanding Mendelian randomization

George Davey Smith

MRC Integrative Epidemiology Unit (IEU) at the University of Bristol

Mendelian randomization (MR) was developed to improve identification and understanding of modifiable processes leading to disease, as the title of the first extended presentation of the approach made clear[1]**.** In furthering this aim it leverages the inferential power of employing germline genetic variants as proxy markers (or instrumental variables) for modifiable phenotypes, which may influence risk of disease. The focus of this approach is not on gene discovery or enhancing knowledge of gene function, although developments in these are essential to its success. The response by Joseph Pickrell[2] to my commentary[3] on his original review[4] clarifies what are the remaining areas of disagreement in regard to enhancing the utility of the approach. He retracts his assertion that in 30-years of use MR had not made major contributions to causal understanding - "I agree with Davey Smith that his correction of "30 years" to "<15 years" makes my statement considerably weaker, and I will back off of this critique in a revision." – and thereafter focuses on pleiotropy, as it is a "...fundamental biological reality that is likely leading to spurious causal claims in practice"[2].

One more time on pleiotropy

Pleiotropy has been very extensively discussed as a key issue that could undermine the usefulness of MR since the initial presentations of the method. However some basic issues of particular relevance to MR may not be adequately appreciated. Consider Pickrell's key example of how he considers pleiotropy may influence interpretation of MR studies. In his words "for example, variants in an intron of *FTO* have been used to study the 'causal effects' of obesity on risk of cancer in the MR context[5]. But these variants are also associated with timing of puberty and HDL cholesterol levels (among others), so the standard MR assumption of 'no pleiotropy' is potentially violated"[2].

Rather than reword one of the many previous discussions of why this misrepresents the problem of pleiotropy in MR I will quote from our latest basic review of the field, which coincidentally discusses exactly the same example as Pickrell uses – *FTO* and HDL cholesterol[6].

> Pleiotropy is the phenomenon by which a single locus influences multiple phenotypes[7]. Depending on the form it takes, pleiotropy is a potential limitation to interpretation of MR, so distinguishing between its different types is important. In the context of MR there are two mechanisms by which pleiotropy occurs: a single process leading to a cascade of events (e.g. a locus influences one particular protein product, and this causes perturbations in many

other phenotypes); or a single locus directly influencing multiple phenotypes[8 9]. Amongst its many names, the former has been termed "spurious pleiotropy"[10 12], "mediated pleiotropy"[11] or "type II pleiotropy"[12]; the latter "biological pleiotropy"[13] or "type I pleiotropy"[12]. Type II pleiotropy is not only unproblematic for MR, it is the very essence of the approach, in which the downstream effects of a perturbed phenotype are estimated through the use of genetic variants that relate to this phenotype. Thus the instrument of common variation in *FTO*, known to influence body mass index (BMI)[14], probably through influencing caloric intake[15 16] is associated with a wide range of downstream phenotypes; blood pressure and hypertension[17] , coronary heart disease (CHD)[18], fasting insulin, glucose, HDL cholesterol and trigylcerides[19], bone mineral density[20], chronic renal disease[21], and diabetes[14]. These associations are expected, as higher BMI influences these traits, and it would be an error to consider them "pleiotropic" effects of *FTO* variation that vitiate MR investigations.

Type I pleiotropy, however, is problematic for the interpretation of MR. Estimates of the degree of pleiotropy suggest that type II pleiotropy is the more pervasive form[12 22], with type I pleiotropy being more pronounced at the level of the gene than at the level of single SNPs [12 23]. Greater pleiotropic effects are seen for mutations with larger effects on the primary trait[24 25], as would be anticipated for type II pleiotropic influences that are downstream effects of considerable perturbation of the primary trait.

Potentially erroneous causal inference due to type I pleiotropy can be minimised by restricting instruments to genetic effects which plausibly act directly on the trait (e.g. genetic instruments for CRP levels located within the promoter region of the *CRP* gene). When less well-characterised variants, or combinations of variants, are utilised then the ways of exploring the potential contribution of pleiotropy detailed in this review and elsewhere[26] need to be implemented[6].

Of course, Pickrell's examples are likely to be type II (spurious or mediated) pleiotropy and whilst it is an elementary truth that pleiotropy is a "fundamental biological reality" the particular way this may influence MR inference in any particular context requires some contextual understanding. Take Pickrell's example of why he thinks there is violation of the MR assumptions with respect to *FTO*, it is because "these variants are also associated with timing of puberty and HDL cholesterol levels" [2] However, greater adiposity causally influences both HDL cholesterol and age at puberty. I will not

labour the point as it is easy to access the extensive literature (including RCTs in the case of HDL cholesterol) on this. In the MR context the degree to which *FTO* influences HDL cholesterol is as anticipated by prospective data on change in BMI and change in HDL cholesterol[27]. The illustration that *FTO* and other genetic markers influencing BMI generate the same predicted effects of BMI on HDL cholesterol[27] is a simple example of how use of multiple genetic variants can provide evidence on the degree to which type1 (or biological) pleiotropy may distort interpretation of any particular MR finding. Pleiotropy is, of course, "a fundamental biological reality", but needs informed consideration in each situation, and cannot be simply used as an automatically enunciated hex to damn any MR study.

Genetic correlations are not Mendelian randomization

Pickrell states, with respect to genetic correlation and MR studies, that "mathematically they are identical". Is this true? In the words of Samuel Beckett, "It is not"[28]. Genetic correlation approaches attempt to estimate the full extent of correlation between phenotypes accounted for by all germline genetic variation. MR studies utilise what is generally a very minor component of genetic influence on a trait as an instrumental variable for that trait. Many studies are based on a single genetic variant, which hopefully makes this clear (e.g. studies of alcohol and blood pressure using a single variant in *ALDH2*[29]).

The example Pickrell gives of studies of height and CHD[30 31] are not investigations of genetic correlation – they utilise 180 genetic variants identified in a GWAS of height – not the whole-genome genetic contribution to height – as the instrumental variable. These variants are largely out of linkage disequilibrium and are scattered across the whole genome, and will work through many processes to influence height. It is highly likely that some of these individual variants will involve mechanisms that could influence CHD risk independently of height. However, the magnitude of the observational association of height and CHD is closely similar to the instrumental variable predicted influence of height on CHD[31], which suggests that height itself may influence CHD through mechanisms directly related to height (e.g. lung function and arterial bore)[31]. All 180 variants cannot plausibly influence height and through type 1 (biological) pleiotropy simultaneously influence CHD risk through processes unrelated to height. Indeed we have previously utilized the association of height and lung function (forced vital capacity) in methodological explorations of MR[32 33]. We did so because we thought it obvious that greater height would lead to greater lung volume for mechanical reasons, but perhaps we should have added a link to a web animation of a schematic figure being

elongated and their lungs increasing in volume to back this up. One of the papers Pickrell references[33] applied an approach (MR-Egger) which allows estimation when (type 1**,** biologically) pleiotropic (i.e. invalid) genetic variants are included, and suggested that with less stringent assumptions regarding valid instrumental variables, estimation was robust. Furthermore we showed that the 180 variants individually were estimating a similar causal effect of height on lung function, with the difference between them being as anticipated by sampling variation (see figure 3 in[33]).

Spot the phenotype

Pickrell is confused by the example of proprotein convertase subtilisin/kexin type 9 (PCSK9) that I presented in my earlier commentary. PCSK9 is a protein, one effect of which is to increase LDL cholesterol levels. Like most proteins it has a series of potential and probable other effects as well[34] [35]. Variation in the *PCSK9* gene influences the level of the protein, PCSK9 [36] [37] [38] [39]. MR inference does not depend upon the genetic variant → intermediate phenotype and the genetic variant→ outcome associations being demonstrated in the same sample, with two-sample approaches becoming increasingly widely used[40] [41]. Nor does MR inference require estimation of the magnitude of the causal effect, see, for instance the different categories of potential inference presented in the first extended presentation of the approach1. Thus the demonstration of a robust *PCSK9* genetic variation → PCSK9 association [36] together with the demonstration that the same *PCSK9* genetic variation is robustly associated with CHD risk[42], provides evidence that the PCSK9 protein increases CHD risk. Monoclonal antibodies against PCSK9 (the protein) have been developed, with promising results[43] [44] [45]. The clue to the proximal phenotype for the MR evidence with respect to PCSK9 is given by the fact that the name of the gene and its protein product are the same. Indeed, establishing that the protein PCSK9 causes CHD has led to the protein having been used as surrogate endpoint in RCTs investigating potential methods of reducing CHD risk [46] [47] [48] [49].

The importance of understanding the primary phenotype (and not just when the name gives it away, as in this case) is that, as discussed in my earlier commentary[3], this provides a powerful way of establishing (i) on-target, (ii) mediated (in this case by LDL cholesterol) and (iii) off-target pharmacotherapeutic effects. Studies leveraging inferential power from *PCSK9* genetic variation demonstrate that the protein, PCSK9, influences LDL cholesterol levels, with downstream effects on CHD. As shown in figure 1 it can be demonstrated that the effect of *PCSK9* genetic variation on CHD is as anticipated from its effect on LDL cholesterol. This suggests that the effects of the protein, PCSK9, on CHD is entirely mediated through the effect of the protein on LDL cholesterol, which was

not necessarily the case given additional potential effects of PCSK9[34 35]. There are several issues which require nuanced consideration in this example. PCSK9 genetic variation can be utilised as an instrument for the protein, PCSK9, and this provides evidence with respect to anticipated effects of manipulating the protein levels. With additional assumptions (that the effect of *PCSK9* genetic variation on outcomes is mediated through the effect of the PCSK9 on LDL cholesterol), *PCSK9* genetic variation can be utilised as an instrument for the effects of LDL cholesterol. Interrogating the appropriateness of this usage depends on additional evidence, such as that presented in figure 1. Thus, as in all inference regarding causality, a synthesis of evidence from various sources is required.[50]

Such findings are of some interest to those attempting to develop effective and safe treatments. To take one active area of research, consider the unresolved issue of whether there are mechanism-specific effects of cholesterol lowering on increasing the risk of diabetes (with HMG-CoA reductase being the mechanism which has been highlighted in this regard), or if it is a consequence of LDL cholesterol lowering through any means. A preliminary study has suggested the latter may be the case[51 52], but this is a methodologically challenging question to answer and further exploration of this is a crucial issue for those concerned with patient care.

<15 years of Mendelian randomization

Pickrell now recognises that the implementation of MR studies has been relatively recent, with "<15 years" being his most recent summation[2]. It is perhaps important to recognise that the large majority of MR studies are recent, reflecting the fact that the identification of robust genetic instruments through GWAS was a pre-requisite for widespread implementation of the approach. Consider a recent systematic review which identified 179 MR studies[53] (although this was not a complete listing, as consortia papers, which are increasingly common in the MR field, those using different terminology, etc, were not included). Figure 2 presents the year-by-year number of papers identified in this review, the search for which finished at the end of 2013[53]. It is clear that the very considerable majority of MR papers have been published in the last five years and most of these in the last couple of years. This would be a good starting point for Pickrell to evaluate his expectation that MR studies should have led to many situations in which new causes have been identified and taken through to completed RCTs, perhaps comparing this to the time course through which other innovative methodologies have moved from initial findings through to completed phrase 3 RCTs.

Genes are not magic, but Mendelian randomization is promising

I very much welcome this exchange with Joseph Pickrell, as I hope it will help clarify some issues regarding MR, and I would value the identification of my undoubtedly many misunderstandings of the approach. I think we agree that cautious optimism is justified, and further development of methods is certainly important. As one enthusiastic tweeter summarised Pickrell's last contribution for those who found it "too long: didn't read", "genetics is not magic"[54], something with which we can all agree. This exchange on MR, in which further objections are raised as previous ones appear chimerical, reminds me of lines from Robert Browning's wonderful "Soliloquy of the Spanish Cloister":

There's a great text in Galatians,
Once you trip on it, entails
Twenty-nine distinct damnations,
One sure, if another fails:[55]

When one damnation passes, the 28 remaining ones await (although in my counting from the *Epistle to the Galatians* there are only 17 damnations, but I am no biblical scholar).

There are good and poor MR studies, just as there are good and poor RCTs, or indeed studies using any design. The strongest inferences can be drawn when different lines of evidence converge, and MR studies can be useful contributors to this evidence synthesis. In the meantime evidence from MR studies is being fed into the practical issue of pharmacotherapeutic development[56 57], with this applied work being carried out in parallel with continued theoretical and methodological speculations. As always critiques remain valuable as practical implementation proceeds (figure 3).

Figure 1. Effect of lower LDL-C on risk of CHD [adapted from Ference et al. (2012)[58]]. Boxes represent the proportional risk reduction (1-OR) of CHD for each exposure allele plotted against the absolute magnitude of lower LDL-C associated with that allele (measured in mg/dl).SNPs are plotted in order of increasing absolute magnitude of associations with lower LDL-C. The line (forced to pass through the origin) represents the increase in proportional risk reduction of CHD per unit lower long-term exposure to LDL-C.

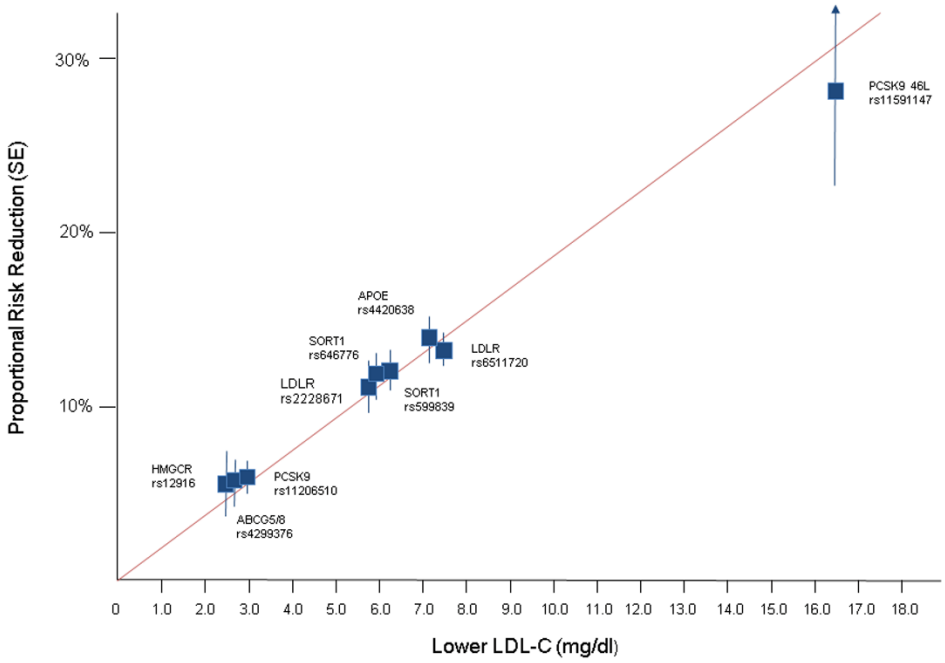

Figure 2: Number of Mendelian randomization studies published per year, from[53] (17 papers located by a search ending 31st December 2013, appeared in journal issues dated 2014).

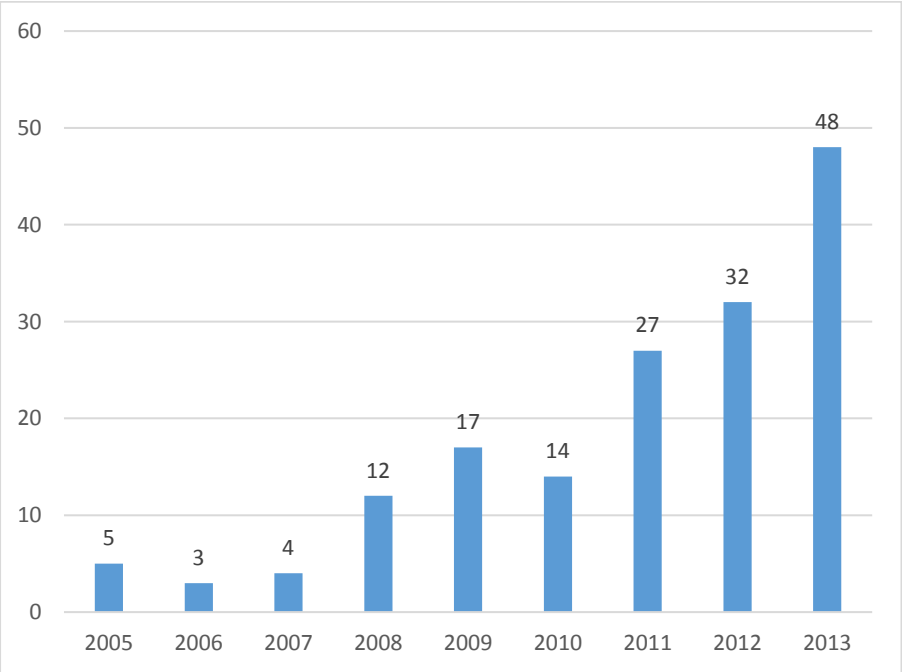

Figure 3: Critics contemplate Mendelian randomization

References

[1] Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiology 2003;32:1-22.
[2] Pickrell J. Some further thoughts on Mendelian randomization (https://www.reddit.com/r/genome/comments/3cf3bw/some_further_thoughts_on_mendelian_randomization/)
[3] Davey Smith G. Mendelian Randomization: a premature burial? BioRxiv 2015; http://biorxiv.org/content/biorxiv/early/2015/06/30/021386.full.pdf
[4] Pickrell J. Fulfilling the promise of Mendelian Randomization. BioRxiv 2015; http://biorxiv.org/content/early/2015/04/16/018150.full.pdf
[5] Brennan P, McKay J, Moore L, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Chow WH, Rothman N, Chabrier A, Gaborieau V, Timpson N, Hung RJ, Smith GD. Obesity and cancer: Mendelian randomization approach utilizing the FTO genotype. Int J Epidemiol. 2009;38:971-5
[6] Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. HMG 2014;23:r89-98 doi: 10.1093/hmg/ddu328.
[7] Stearns FW. One hundred years of pleiotropy: a retrospective. Genetics 2010; 186: 767-773.
[8] Hodgkin J. Seven types of pleiotropy. Int. J. Dev. Biol 1998; 42: 501-505.
[9] Pyeritz RE. Pleiotropy revisited: molecular explanations of a classic concept. American Journal of Medical Genetics 1989; 34: 124-134.
[10] Gruneberg H. An analysis of the "pleiotropic" effects of a lethal mutation in the rat. Proc. R. Soc. Lond. B. 1938; 125: 123-144.
[11]Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet. 2013;14:483-95.

[12] Wagner GP, Zhang J. The pleiotropic structure of the genotype – phenotype map: the evolvability of complex organisms. Nature Reviews Genetics 2011; 12: 204-213.
[13] Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet. 2013;14:483-95.
[14] Frayling T.M., Timpson N.J., Weedon M.N et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 2007;316:889-894.
[15] Timpson NJ, Emmett PM, Frayling TM, Rogers I, Hattersley AT, McCarthy MI, Davey Smith G. The fat mass- and obesity- associated locus and dietary intake in children. Am J Clin Nutr 2008;88:971-978.
[16] Richmond RC, Timpson NJ. Recent Findings on the Genetics of Obesity: Is there Public Health Relevance? Current Nutrition Reports. 2012;DOI 10.1007/s13668-012-0027-x
[17] Timpson N., Harbord R., Davey Smith G., et al. Does greater adiposity increase blood pressure and hypertension risk? Mendelian randomization using Fto/Mc4r genotype.Hypertension 2009;54:84-90.
[18] Nordestgaard BG, Palmer TM, Benn M, Zacho J, Tybjærg-Hansen A, Davey Smith G, Timspon NJ. The Effect of Elevated Body Mass Index on Ischemic Heart Disease Risk: Causal Estimates from a Mendelian Randomisation approach. PLoS Med 2012;9: e1001212. doi:10.1371/journal.pmed.1001212
[19] Freathy RM, Timpson NJ, Lawlor DA, Pouta A, Ben-Shlomo Y, Ruokonen A, Ebrahim S, Shields B, Zeggini E, Weedon WM, Lindgren CM, Lango H, Melzer D, Ferrucci L, Paolisso G, Neville MJ, Karpe F, Palmer CAN, Morris AD, Elliott P, Jarvelin MR, Davey Smith G, McCarthy MI, Hattersley AT, Frayling TM. Common variation in the FTO gene alters diabetes-related metabolic traits to extent expected, given its effect on BMI. Diabetes 2008;57;1419-1426.
[20] Timpson NJ, Sayers A, Davey-Smith G, Tobias JH. How does body fat influence bone mass in childhood? A MRapproach. J Bone Miner Res. 2009;24:522-33.
[21] Hubacek, JA and Viklicky, O and Dlouha, D and Bloudickova, S and Kubinova, R and Peasey, A and Pikhart, H and Adamkova, V and Brabcova, I and Pokorna, E, Bobak, M. The FTO gene polymorphism is associated with end-stage renal disease: two large independent case-control studies in a general population. Nephrol Dial Transplant 2012;27:1030 - 1035.
[22] He X, Zhang J. Toward a molecular understanding of pleiotropy. Genetics 2006; 173: 1885-1891.
[23] Stern DL. Evolutionary developmental biology and the problem of variation. Evolution 2000; 54: 1079-1091.
[24] Wang Z, Liao B-Y, Zhang J. Genomic patterns of pleiotropy and the evolution of complexity. PNAS 2010; doi: 10.7073/pnas.1004666107.
[25] Wagner GP. Pleiotropic scaling of gene effects and the “cost of complexity”. Nature 2008 452; 470-472.
[26] Davey Smith G. Use of genetic markers and gene-diet interactions for interrogating population-level causal influences of diet on health. Genes Nutr.2011;6:27-43.
[27] Würtz P, Wang Q, Kangas AJ et al. Metabolic Signatures of Adiposity in Young Adults: Mendelian randomization Analysis and Effects of Weight Change. PLoS Medicine 2014: doi: 10.1371/journal.pmed.1001765.
[28] Beckett S. More pricks than kicks. Chatto and Windus: London. 1934.
[29] Chen L, Davey Smith G, Harbord R, Lewis S. Alcohol intake and blood pressure: a systematic review implementing MRapproach. PLoS Medicine 2008;5: 461-471.
[30] Nelson CP, Hamby SE, Saleheen D et al. Genetically Determined Height and Coronary Artery Disease. N Engl J Med 2015; 372:1608-1618
[31] Nüesch E, Dale C, Palmer TM et al. Adult height, coronary heart disease and stroke: a multi-locus MRmeta-analysis. Int. J. Epidemiol. 2015 doi:10.1093/ije/dyv074
[32] Davies NM, von Hinke Kessler Scholder S, Farbmacher H, Stephen B, Windmeijer F, Davey Smith G. The many weak instruments problem and Mendelian randomization. Statistics in Medicine 2015;34:454-468.

[33] Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015, DOI: 10.1093/ije/dyv080.
[34] Schulz R, Schlüter KD, Laufs U . Molecular and cellular function of the proprotein convertase subtilisin/kexin type 9 (PCSK9). Basic Res Cardiol. 2015; 110: 4.
[35] Cariou B, Si-Tayeb K, Le May C. Role of PCSK9 beyond liver involvement. Current Opinion in Lipidology 2015;26:155–161
[36] Lakoski SG, Lagace TA, Cohen JC, Horton JD, Hobbs HH. Genetic and metabolic determinants of plasma PCSK9 levels. J Clin Endocrinol Metab. 2009;94:2537-43
[37] Humphries SE, Neely RD, Whittall RA, Troutt JS, Konrad RJ, Scartezini M, Li KW, Cooper JA, Acharya J, Neil A. Healthy individuals carrying the PCSK9 p.R46L variant and familial hypercholesterolemia patients carrying PCSK9 p.D374Y exhibit lower plasma concentrations of PCSK9. Clin Chem. 2009;55:2153-61.
[38] Chernogubova E, Strawbridge R, Mahdessian H, Mälarstig A, Krapivner S, Gigante B, Hellénius ML, de Faire U, Franco-Cereceda A, Syvänen AC, Troutt JS, Konrad RJ, Eriksson P, Hamsten A, van 't Hooft FM. Common and low-frequency genetic variants in the PCSK9 locus influence circulating PCSK9 levels. Arterioscler Thromb Vasc Biol. 2012;32:1526-34.
[39] Postmus I, Trompet S, de Craen AJ, Buckley BM, Ford I, Stott DJ, Sattar N, Slagboom PE, Westendorp RG, Jukema JW. PCSK9 SNP rs11591147 is associated with low cholesterol levels but not with cognitive performance or noncardiovascular clinical events in an elderly population. J Lipid Res. 2013;54:561-6.
[40] Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. Am J Epidemiology 2013;178:1177-84.
[41] Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG; EPIC- InterAct Consortium. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiology 2015.
[42] Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med 2006;354:1264-72.
[43] Sabatine MS, Giugliano RP, Wiviott SD, Raal FJ, Blom DJ, Robinson J, Ballantyne CM, Somaratne R, Legg J, Wasserman SM, Scott R, Koren MJ, Stein EA; Open-Label Study of Long-Term Evaluation against LDL Cholesterol (OSLER) Investigators. Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. N Engl J Med 2015;372:1500-9.
[44] Robinson JG, Farnier M, Krempf M, Bergeron J, Luc G, Averna M, Stroes ES, Langslet G, Raal FJ, El Shahawy M, Koren MJ, Lepor NE, Lorenzato C, Pordy R, Chaudhari U, Kastelein JJ; ODYSSEY LONG TERM Investigators. Efficacy and safety of alirocumab in reducing lipids and cardiovascular events. N Engl J Med. 2015 Apr 16;372(16):1489-99. doi: 10.1056/NEJMoa1501031. Epub 2015 Mar 15.
[45] Navarese EP, Kolodziejczak M, Schulze V, Gurbel PA, Tantry U, Lin Y, Brockmeyer M, Kandzari DE, Kubica JM, D'Agostino RB Sr, Kubica J, Volpe M, Agewall S, Kereiakes DJ, Kelm M. Effects of Proprotein Convertase Subtilisin/Kexin Type 9 Antibodies in Adults With Hypercholesterolemia: A Systematic Review and Meta-analysis. Ann Intern Med 2015; doi: 10.7326/M14-2957.
[46] Graversen CB, Lundbye-Christensen S, Thomsen B, Christensen JH, Schmidt EB. Marine n-3 polyunsaturated fatty acids lower plasma proprotein convertase subtilisin kexin type 9 levels in pre- and postmenopausal women: a randomised study. Vascular Pharmacology 2015: DOI: 10.1016/j.vph.2015.07.001.
[47] Arsenault BJ, Pelletier-Beaumont E, Alméras N, Tremblay A, Poirier P, Bergeron J, Després JP. PCSK9 levels in abdominally obese men: association with cardiometabolic risk profile and effects of a one-year lifestyle modification program. Atherosclerosis. 2014;236:321-6.
[48] Cariou B, Langhi C, Le Bras M, Bortolotti M, Le K, Theytaz F, May C, Guyomarc'h-Delasalle B, Zair Y, Kreis R, Boesch C, Krempf M, Tappy L, Costet P. Plasma PCSK9 concentrations during an oral fat load and after short term high-fat, high-fat high-protein and high-fructose diets. Nutrition & Metabolism 2013;10:4.

[49] Richard C, Couture P, Desroches S, Lamarche B. Effect of the Mediterranean diet with and without weight loss on markers of inflammation in men with metabolic syndrome. Obesity (Silver Spring). 2013; 21:51-7

[50] Lipton P. Inference to the Best Explanation (2nd edition). Routledge: UK, 2004.

[51] Fall T, Xie W, Poon W et al. Using Genetic Variants to Assess the Relationship Between Circulating Lipids and Type 2 Diabetes. Diabetes 2015;64:2676-2684.

[52] Swerdlow DI, Sattar N. Blood Lipids and Type 2 Diabetes Risk: Can Genetics Help Untangle the Web? Diabetes 2015;64:2344-2345.

[53] Boef AGC, Dekkers OM, le Cessie S. Mendelian randomization studies: a review of the approaches used and the quality of reporting. Int J Epidemiol 2015;44:496-511.

[54] MacArthur D. tl;dr genetics is not magic": great post by @joe_pickrell in response to @mendel_random: 7th July 2015. https://twitter.com/dgmacarthur/status/618567779084595201

[55] Robert Browning. Soliloquy of the Spanish Cloister. *Dramatic Lyrics*. London: Moxon, 1842.

[56] Kathiresan S. Developing medicines that mimic the natural successes of the human genome: lessons from NPC1L1, HMGCR, PCSK9, APOC3, and CETP. J Am Coll Cardiol. 2015;65:1562-6.

[57] Catapano AL, Ference BA. IMPROVE-IT and genetics reaffirm the causal role of LDL in Cardiovascular Disease. Atherosclerosis 2015;241:498-501.

[58] Ference BA, Yoo W, Alesh I, Mahajan N, Mirowska KK, Mewada A, Kahn J, Afonso L, Williams KA Sr, Flack JM. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. J. Am. Coll. Cardiol 2012, 60, 2631–2639