Common Variants for CardioVascular Disease:

Clinical utility confirmed!

Steve E Humphries

Centre for Cardiovascular Genetics, Institute of Cardiovascular Science,

5 University Street, University College London, London, UK WC1E 6JF.

email: steve.humphries@ucl.ac.uk

Introduction

The increase in sample size of Genome-Wide Association Study (GWAS) meta-analyses, has led to the identification now of more than 150 common variants/single nucleotide polymorphisms (SNPs) that are robustly associated with cardiovascular disease (CVD), or Coronary Heart Disease (CHD), or with CVD traits such as coronary calcification (http://www.ebi.ac.uk/gwas/). For example in 2013, the CARDIoGRAMplusC4D Consortium, with 63,746 CAD cases and 130,681 controls[1] identified 15 novel loci, taking the number of statistically robust CVD SNPs to 46, with a further 104 SNPs associated with CVD at a 5% false discovery rate. Together, these variants explained approximately 10.6% of CVD heritability. Many of these loci identified are known to be involved in lipid metabolism[2], as would be expected from our knowledge of the importance of dyslipidaemia in the development of CVD, with 12/46 CARDIoGRAMplusC4D SNPs showing a significant association with a lipid trait.

These SNPs, located throughout the human genome, are common (frequency all >5%), with the Odds Ratio (OR) for disease ranging from ~30% higher risk for carriers of top ranking GWAS risk variants at the chromosome 9p *CDKN2A/2B* locus, to being only 7-8% higher risk

for carriers of SNPs at the loci for *PCSK9* (chromosome 2) or for *HNF1A* (chromosome 12). Since all of these OR are modest, the question arises as to what is the potential clinical utility of these SNPs and how can we use these genetic tools to explore the inherited contribution to CVD?

Given the relatively small effect sizes associated with these risk loci, it is unsurprising that the addition of one variant into a Classical Risk Factor (CRF) risk score does not result in improved predictive ability[3]. This has led to the development of so-called "genetic risk scores" (GRSs) where SNPs at independent loci are combined. A GRS can be unweighted, where simply the number of risk alleles carried by an individual at each locus is summed, but this assumes that the risk effect associated with each SNP is equal (and additive) and this clearly is not the case. A more accurate GRS can be constructed if carriage of the individual SNPs are weighted using the published effect size. An individual's GRS can then be combined with a Classical Risk Factor (CRF) score such as the Framingham score or QRISK2[4] to give an individual's overall CVD risk estimate.

Statin benefit in those at high genetic risk

The paper published this week in Circulation[5] uses ~60 SNPs that were selected from GWAS studies to be significantly associated with CHD. A weighted GRS was calculated for each subject. The main analysis was carried out in 4910 subjects from The West of Scotland Coronary Prevention Study (WOSCOPS) randomized controlled trial of primary prevention with pravastatin (40mg daily) therapy. Primary outcome was non-fatal MI or death from CHD, and subjects were followed up for a mean of 4.8 years in trial and a further mean 8.7 years out of trial. Subjects designated at high genetic risk were those in the top quintile of the score and were compared to those in the lower four quintiles.

In the placebo group, compared to the low genetic risk subjects, the HR for CHD in those in the high genetic risk quintile was 1.62 (95% CI 1.29-2.05) after adjustment for CRFs. A one SD increase in score was associated with a 25% increase in CHD incidence. In the treated group, statin treatment reduced risk for a first CHD event in the high score subjects by 44% compared to only a 25% reduction in the other subjects. This translated to a 7.9% risk

reduction in the high risk group compared to a 2.7% reduction in the others (p = 0.04 for heterogeneity), with a number needed to treat of 13 in the high score group vs 38 among others. It is worth noting that this difference was achieved with a similar (22%) reduction of LDL-C in both groups.

There was confirmatory analysis in two observational cohort studies. Each standard deviation increase in the GRS was associated with 1.32-fold (95% CI, 1.04-1.68) greater likelihood of having coronary artery calcification in the Coronary Artery Risk Development in Young Adults study (CARDIA) of 1154 participants, and with a 9.7% higher (95% CI, 2.2-17.8%) burden of carotid plaque in the BioImage Study of 4392 subjects. The authors concluded that those at highest genetic risk have a higher burden of subclinical atherosclerosis, and with statin therapy will experience greater relative and absolute benefit to prevent a first CHD/CVD event.

So can these data be extrapolated to other cohorts of patients? The subjects in WOSCOPS were all selected to have high LDL-C (mean untreated levels 192 mg/dl (5.3mmol/l) and as such would be expected to show a large benefit from statin therapy. Interestingly, the GRS score was modestly associated with a family history of early CHD, but not with baseline lipid levels, which is surprising since at least 15 of the included SNPs are in lipid genes (eg *APOB, PCSK9, LDLR, APOE* etc)). This perhaps reflects the utility of including so many SNPs which influence CVD through different pathways. WOSCOPs were all males, but there is no reason to believe that the score is not equally applicable in women, as shown in the other two cohorts examined here were there was no gender difference in effects.

Also there are several studies which support the association of the GRS with risk of CVD in general population cohorts both in the US[6] and in Europe[7,8]. In the UCLEB study[7], data were drawn from seven UK prospective studies including 11, 851 individuals initially free of CVD, with 1444 incident CVD events over 10 years' follow-up. Using 53 CVD GWAS SNPs and the QRISK-2 CRF algorithm the GRS showed only modest improvement in risk stratification, with additional benefit mainly in those at intermediate risk. Applying the GRS only to those with QRISK-2 risk of 10%-<20% and prescribing statins where risk exceeded 20% suggested that genetic information could prevent one additional event for every 462 people screened. Morris et al [7] proposed that tailored prediction using a GRS for those at intermediate risk

may have clinical utility. In a similar study of five prospective population cohorts from Finland and the US[8] and a GRS which took into account over 49,000 SNPs, addition of the GRS to CRF scores significantly improved the 10 years risk prediction ($p < 0.001$), particularly for individuals ≥60 years old ($p < 0.001$). The GRS captured substantially different trajectories of absolute risk, with men in the top 20% of the GRS reaching a 10% CHD risk 12-18 y earlier than those in the bottom 20%.

Motivation to initiate lipid-lowering therapy

One useful consequence of informing at-risk subjects of their genetic risk would be if such information motivates lifestyle changes or adherence to prescribed medication to a greater extent than CRF information alone, and a trial from the US has examined this[9]. Two hundred and three subjects (mean age 59 years), at intermediate risk for CHD, and not on statins were randomly assigned to receive their 10-year probability of CHD based either on a Framingham CRF score or CRFs+GRS. The GRS included 28 CHD GWAS hit SNPs, with each weighted by its published OR effect size. Subjects were told their risk as "high" or "average" or "low" by a genetic counsellor, followed by a discussion about starting statin therapy with a physician. After 6 months follow-up the CRF+GRS group had 9% lower LDL-C than the CRF only group ($p=0.04$), with particular benefit seen in those given a high overall risk. This was because more subjects in the high risk group started statin therapy than in the CRF only group (39% vs 22%, $p<0.01$), and was not due to differences between the groups in dietary fat intake or levels of physical activity. These data support the view that in a clinical setting, a high genetic risk may be particularly motivating to start (and possible adhere to) lipid lowering medication.

Future Prospects

It is unlikely that there are any additional common SNPs to be found by GWAS meta-analysis that will outrank the OR of those currently known, since SNPs with larger effects that are common would have been detected by the size of the datasets currently available.

However, currently for the majority of the risk SNPs, the actual functional variant(s) at the locus is unknown. Therefore, one area of progress over the next five years would be the identification of the functional SNP/SNPs at each risk locus, and molecular approaches[10] and guidelines for this process have recently been published[11]. Of the GWAS hits only 7% are exonic missense variants[10] and therefore may be directly affecting gene function, with the variants creating the well-known *APOE* variants (Ɛ2, Ɛ3, Ɛ4) being a notable example. Another 7% are in the promoter region and thus may potentially be affecting gene expression, while over 70% are intronic or intergenic and are therefore of unknown or possibly of no functional consequence. For example, using luciferase assays of promoter strength we have recently shown that both the *LDLR* GWAS lead SNP rs6511720 located in intron 1, plus only one of several other SNPs in the intron in complete linkage disequilibrium (LD) with this SNP, are functional[12]. Similarly, for the chromosome 9p21 locus, although molecular studies have reported interesting findings[13,14], the precise mechanism of action in the development of CVD remains unclear, almost a decade after discovery. In particular, the actual functional SNP/SNPs at this locus have not been definitively identified. Thus most of the SNPs included in the score are GWAS hits where the lead SNP is unlikely to be the functional SNP at that risk locus. LD between the lead and functional SNPs may differ between ethnicities, meaning that some SNPs will be better proxies than others. This will reduce the ability of the weighted GRS to accurately reflect CVD risk, particularly in different ethnic groups if the LD is less. While analysing every GWAS CVD hit locus in this way constitutes a considerable amount of work, ultimately this will provide the most accurate panel of functional SNPs for CVD risk prediction.

References

1.   CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups K, König IR, Cazier JB, Johansson A, Hall AS, Lee JY, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A, Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K, Lundmark P, Lyytikäinen LP, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altshuler D, Balmforth AJ, Barroso I, Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R; DIAGRAM Consortium; CARDIOGENICS Consortium, Doney AS, El Mokhtari N, Eriksson P, Fischer K, Fontanillas P, Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han BG, Hunt SE, Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C, Leander K, Lokki ML, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S, Morris AD, Müller-Nurasyid M; MuTHER Consortium, Nikus K, Peden JF, Rayner NW, Rasheed A, Rosinger S, Rubin D, Rumpf MP, Schäfer A, Sivananthan M, Song C, Stewart AF, Tan ST, Thorgeirsson G, van der Schoot CE, Wagner PJ; Wellcome Trust Case Control Consortium, Wells GA, Wild PS, Yang TP, Amouyel P, Arveiler D, Basart H, Boehnke M, Boerwinkle E, Brambilla P, Cambien F, Cupples AL, de Faire U, Dehghan A, Diemert P, Epstein SE, Evans A, Ferrario MM, Ferrières J, Gauguier D, Go AS, Goodall AH, Gudnason V, Hazen SL, Holm H, Iribarren C, Jang Y, Kähönen M, Kee F, Kim HS, Klopp N, Koenig W, Kratzer W, Kuulasmaa K, Laakso M, Laaksonen R, Lee JY, Lind L, Ouwehand WH, Parish S, Park JE, Pedersen NL, Peters A, Quertermous T, Rader DJ, Salomaa V, Schadt E, Shah SH, Sinisalo J, Stark K, Stefansson K, Trégouët DA, Virtamo J, Wallentin L, Wareham N, Zimmermann ME, Nieminen MS, Hengstenberg C, Sandhu MS, Pastinen T, Syvänen AC, Hovingh GK, Dedoussis G, Franks PW, Lehtimäki T, Metspalu A,

Zalloua PA, Siegbahn A, Schreiber S, Ripatti S, Blankenberg SS, Perola M, Clarke R, Boehm BO, O'Donnell C, Reilly MP, März W, Collins R, Kathiresan S, Hamsten A, Kooner JS, Thorsteinsdottir U, Danesh J, Palmer CN, Roberts R, Watkins H, Schunkert H, Samani NJ. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet. 2013 Jan;45(1):25-33. doi: 10.1038/ng.2480. Epub 2012 Dec 2.

2.   Global Lipids Genetics Consortium., Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Do R, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkilä K, Hyppönen E, Isaacs A, Jackson AU, Johansson A, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikäinen LP, Magnusson PK, Mangino M, Mihailov E, Montasser ME, Müller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney AS, Döring A, Elliott P, Epstein SE, Eyjolfsson GI, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJ, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimäki T, Lin SY, Lindström J, Loos RJ, Mach F, McArdle WL, Meisinger C, Mitchell BD, Müller G, Nagaraja R, Narisu N, Nieminen TV, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruokonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stancáková A, Stirrups K, Swift AJ, Tiret L, Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YD, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrières J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllensten U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Järvelin MR, Jula A, Kähönen M, Kaprio J, Kesäniemi A, Kivimaki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, März W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njølstad I,

Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PE, Sheu WH, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BH, Ordovas JM, Boerwinkle E, Palmer CN, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Kathiresan S, Mohlke KL, Ingelsson E, Abecasis GR. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013 Nov;45(11):1274-83

3. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. Heart. 2008 Jan;94(1):34-9.

4. Talmud PJ, Cooper JA, Palmen J, Lovering R, Drenos F, Hingorani AD, Humphries SE. Chromosome 9p21.3 coronary heart disease locus genotype and prospective risk of CHD in healthy middle-aged men. Clin Chem. 2008 Mar;54(3):467-74.

5. Natarajan P, Young R, Stitziel NO, Padmanabhan S, Baber U, Mehran R, Sartori S, Fuster V, Reilly DF, Butterworth A, Rader DJ, Ford I, Sattar N, Kathiresan S. Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. Circulation 2017 in press.

6. Mega JL, Stitziel NO, Smith JG, Chasman DI, Caulfield MJ, Devlin JJ, Nordio F, Hyde CL, Cannon CP, Sacks FM, Poulter NR, Sever PS, Ridker PM, Braunwald E, Melander O, Kathiresan S and Sabatine MS. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet*. 2015;385:2264-71.`

7. Morris RW, Cooper JA, Shah T, Wong A, Drenos F, Engmann J, McLachlan S, Jefferis B, Dale C, Hardy R, Kuh D, Ben-Shlomo Y, Wannamethee SG, Whincup PH, Casas JP, Kivimaki M, Kumari M, Talmud PJ, Price JF, Dudbridge F, Hingorani AD, Humphries SE; UCLEB Consortium.. Marginal role for 53 common genetic variants in cardiovascular disease prediction. Heart. 2016 Oct 15;102(20):1640-7.

8. Abraham G, Havulinna AS, Bhalala OG, Byars SG, De Livera AM, Yetukuri L, Tikkanen E, Perola M, Schunkert H, Sijbrands EJ, Palotie A, Samani NJ, Salomaa V, Ripatti S,

Inouye M. Genomic prediction of coronary heart disease. Eur Heart J. 2016 Nov 14;37(43):3267-3278.

9. Kullo IJ, Jouni H, Austin EE, Brown SA, Kruisselbrink TM, Isseh IN, Haddad RA, Marroush TS, Shameer K, Olson JE, Broeckel U, Green RC, Schaid DJ, Montori VM, Bailey KR. Incorporating a Genetic Risk Score Into Coronary Heart Disease Risk Estimates: Effect on Low-Density Lipoprotein Cholesterol Levels (the MI-GENES Clinical Trial). Circulation. 2016 Mar 22;133(12):1181-8

10. Smith AJ, Humphries SE, Talmud PJ. Identifying functional noncoding variants from genome-wide association studies for cardiovascular disease and related traits. Curr Opin Lipidol. 2015 Apr;26(2):120-6

11. MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, Adams DR, Altman RB, Antonarakis SE, Ashley EA, Barrett JC, Biesecker LG, Conrad DF, Cooper GM, Cox NJ, Daly MJ, Gerstein MB, Goldstein DB, Hirschhorn JN, Leal SM, Pennaccio LA, Stamatoyannopoulos JA, Sunyaev SR, Valle D, Voight BF, Winckler W, Gunter C. Guidelines for investigating causality of sequence variants in human disease. Nature. 2014 Apr 24;508(7497):469-76.

12. Fairoozy RH, White J, Palmen J, Kalea AZ, Humphries SE. Identification of the Functional Variant(s) that Explain the Low-Density Lipoprotein Receptor (LDLR) GWAS SNP rs6511720 Association with Lower LDL-C and Risk of CHD. PLoS One. 2016 Dec 14;11(12):e0167676.

13. Miller CL, Pjanic M, Wang T, Nguyen T, Cohain A, Lee JD, Perisic L, Hedin U, Kundu RK, Majmudar D, Kim JB, Wang O, Betsholtz C, Ruusalepp A, Franzén O, Assimes TL, Montgomery SB, Schadt EE, Björkegren JL, Quertermous T. Integrative functional genomics identifies regulatory mechanisms at coronary artery disease loci. Nat Commun. 2016 Jul 8;7:12092.

14. Almontashiri NA, Antoine D, Zhou X, Vilmundarson RO, Zhang SX, Hao KN, Chen HH, Stewart AF. 9p21.3 Coronary Artery Disease Risk Variants Disrupt TEAD Transcription Factor-Dependent Transforming Growth Factor β Regulation of p16 Expression in Human Aortic Smooth Muscle Cells. Circulation. 2015 Nov 24;132(21):1969-78.