

COHORT PROFILE

Cohort Profile: The Corogene study

Satu Vaara,¹ Markku S Nieminen,¹ Marja-Liisa Lokki,² Markus Perola,³ Pirkko J Pussinen,⁴ Jaakko Allonen,¹ Olavi Parkkonen¹ and Juha Sinisalo^{1*}

¹Division of Cardiology, Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland, ²Transplantation Laboratory, Haartman Institute, University of Helsinki, Helsinki, Finland, ³Public Health Genomics, National Institute for Health and Welfare, Helsinki, Finland and ⁴Institute of Dentistry, University of Helsinki, Helsinki, Finland

*Corresponding author. Division of Cardiology, Department of Medicine, Helsinki University Central Hospital, PO Box 340, FI-00029 HUS, Finland. E-mail: juha.sinisalo@hus.fi

Accepted 5 May 2011

How did the study come about?

Coronary artery disease (CAD) risk factor profiles change over time. The death rate from coronary disease has been decreasing mainly due to changes in levels of risk factors, and improvement in CAD treatments.¹ However, there are risk factors, whose prevalence is increasing, such as diabetes and obesity.² When the risk factor profile changes, we can assume that gene profiles and epigenetics also change.³ This has not been studied widely.

The Corogene study was designed as a large cohort to study mainly CAD, but also other related heart diseases such as heart failure and aortic valve disease. We selected the patients from the CAD point of view, and decided to include over 5000 consecutive patients assigned for coronary angiogram. In Finland, coronary angiogram is performed to practically all patients assigned for invasive heart examination. Despite technical developments in diagnostics, coronary angiogram is still the gold standard for evaluating coronaries.⁴ The purpose of this study is to follow contemporary trends in coronary heart disease, and related heart disease risk factors, genetics and epigenetics by collecting cohorts referred to heart examination. New cohorts will be collected at 5-year intervals in order to see trends in CAD, its risk factors and epigenetics.

What does it cover?

The objective was to increase our knowledge of (i) how genetics and epigenetics affect the development and degree of heart diseases, (ii) use of medication over time: what and how often medications are prescribed, (iii) what is the adherence over time to medication and how it affects morbidity and

mortality, (iv) pharmacogenetics; who benefits from the use of medication, (v) to uncover the inflammatory pathways linked to these diseases and (vi) finally, contemporary trends in morbidity, mortality and risk factors in patients assigned to coronary angiogram.

Who is in the sample?

The Corogene study is a prospective cohort study where 5000 consecutive Finnish patients assigned to coronary angiogram in the region of Helsinki University Central Hospital are included. The goal was to study more than 5000 patients in the first cohort, which was achieved during a 20-month period (June 2006 to March 2008). A new cohort of 5000 patients will be collected every 5 years within a period of 12 months in order to study the trends in heart disease and coronary risk factors. The next cohort is scheduled to be collected in 2013.

The Hospital District of Helsinki and Uusimaa comprises a population of 1.5 million, and includes four hospitals with coronary angiogram laboratories, three of which (Helsinki University Central Hospital, Jorvi Hospital and Peijas Hospital) are supervised by the Helsinki University Central Hospital. The fourth unit is in a private hospital. During the study period, over 4800 coronary angiograms were performed annually within the region. Helsinki University Central Hospital covers 75%, Jorvi Hospital 16%, Peijas Hospital 4% and the private hospital 5% of these angiograms. ST-elevation myocardial infarction patients are referred exclusively to Helsinki University Central Hospital. During off-duty hours, the region sending ST-elevation myocardial infarction patients for angiogram is wider; it includes about two million inhabitants.

Criteria for inclusion

Patient was assigned to coronary angiogram. Finnish origin was required for follow-up reasons. Criteria for exclusion comprised: non-Finnish origin, previous heart transplantation, low haemoglobin or previous blood transfusion during the same hospitalization. Of the excluded patients, we registered sex, age and the reason for exclusion.

Participation rate for the study was very high: 93.4% of consecutive patients of Finnish origin were included. For patient recruitment see Figure 1.

For further evaluation, the patients were assigned to four groups as follows: (i) Patients without significant CAD ('non-CAD' patients). Angiograms were performed on these patients for diagnostic reasons. These patients had stable or atypical chest pains, valvular heart disease, cardiomyopathy or other conditions requiring diagnostic coronary angiogram. At examination the coronaries were either completely normal or there were non-significant obstructions ($\leq 50\%$). (ii) Patients with stable CAD with $>50\%$ stenosis in at least one coronary artery.⁵ (iii) Acute coronary syndrome (ACS) defined as an episode of typical chest pain for ischaemia, and $>50\%$ stenosis

in at least one coronary artery. These patients were classified as having unstable angina, non-ST-elevation myocardial infarction or ST-elevation myocardial infarction according to ischaemic ECG changes, and elevated levels of a cardiac biomarkers.⁴ Those patients who had clear symptoms and ST elevation, but no significant coronary obstructions ($\leq 50\%$) were still considered as ST-elevation myocardial infarction patients, if they had large cardiac enzyme release ($n=7$): (creatin kinase-MB_{mass} >50 mg/l). (iv) Patients with other ischaemic event defined as an ACS-like episode of chest pain with or without ischaemic changes in ECG, and without significant ($>50\%$) CAD and elevated levels of cardiac biomarkers. This group included patients with type 2 myocardial infarction⁶ and conditions with troponin elevation without overt ischaemic heart disease such as myocarditis and apical ballooning syndrome.

How often have they been followed up?

Nationwide registers in Finland are used for follow-up of the subjects. We have a link to the national Hospital Discharge Registry maintained by the Ministry of Social Affairs and Health,⁷ where date of every hospitalization in Finland has been filed. At this point, we have follow-up information till the end of 2008. The cause of death statistics are obtained from Statistics Finland in Helsinki. This registry covers deaths and causes of deaths of all Finnish residents.⁷ The follow-up covers information from the beginning of the study up to 5 May 2010, median follow-up being 36 months (range 25–48 months). The information of medication usage is based on the Finnish prescription register maintained by the Social Insurance Institution of Finland.⁸ This registry covers, at this stage, all patients' medication purchases from 1 January 2005 (>1 year before the initial hospitalization) to 31 March 2009. Information from all of these registries and of all patients will be collected at 1-, 5-, 10- and thereafter at 5-year bases. New cohorts will be followed up similarly.

What has been measured?

Each patient filled a two-page questionnaire, which included anthropometry (height), medication, heart symptoms (which kind and for how long, rhythm problems), other diseases (hypertension, diabetes, hypercholesterolaemia, claudication, brain insults, myocardial infarctions) and heart diseases or sudden death of relatives (who, when, at what age and which disease). A computerized database was used for the collection of extensive information from hospital records. Trained nurses, supervised by physicians, collected the information from hospital documents (patient records and coronary-angiogram database).

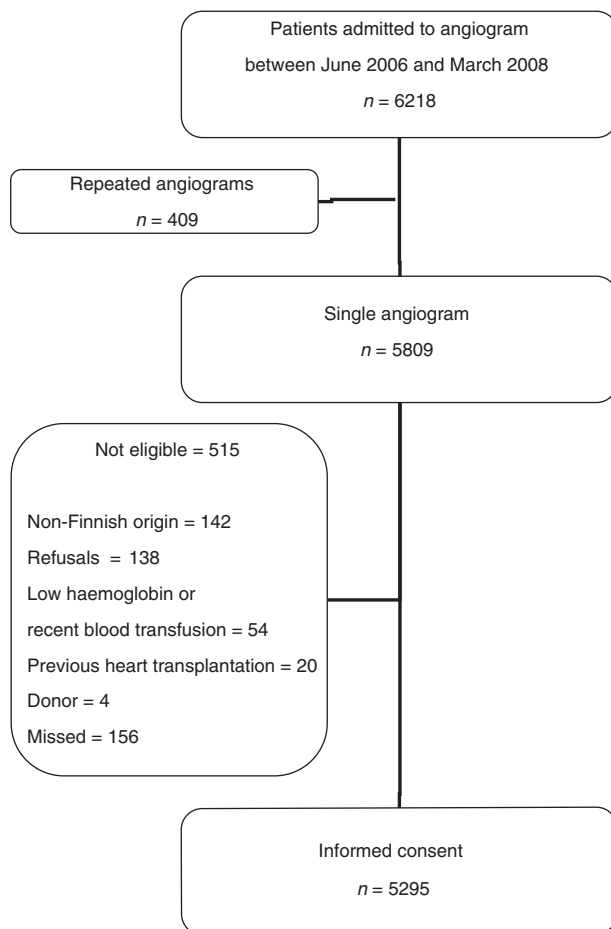


Figure 1 Patient selection process in the first cohort

Table 1 Summary of data items in the Corogene study

Anthropometry
Weight
Patient questionnaire
Height, cardiac symptoms, duration of these symptoms, other vascular disease symptoms, other diseases symptoms, family history of heart diseases, medication in use: including over-the-counter use of pain medication, smoking, use of alcohol, drug abuse, dental health, allergies, special diets, social situation, home nursing
Patient records
Height, risk factors, prior cardiac diseases: including myocardial infarctions, arrhythmias, heart failure, pacemakers and prior cardiac operations, other operations, co-morbidities, medication
Laboratory values: complete blood count, C-reactive protein, total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, glucose, haemoglobin-A1C, alkaline phosphatase, aspartate aminotransferase, alanine aminotransferase, bilirubin, activated partial thromboplastin time, INR, potassium, sodium, creatinine, pro-brain natriuretic peptide, creatine kinase, creatine kinase-MBmass, troponin-T, D-dimer, urea
Chest X-ray
Echocardiogram
Cardiological data log
Angiography, percutaneous coronary interventions, bypass operations
Cardiac surgery database
Electrocardiogram
Laboratory samples
Plasma, serum, DNA, RNA
Aortic valve evaluation by echocardiogram (sub-study of 2200 patients for aortic valve classification)
Dental health of sub-population of 500 (Parogene study)
Oral examination by dentist
Orthopantomography
Sputum sample for bacteria cultures
Questionnaire of dental health
Medication purchase register from the Social Insurance Institution of Finland (KELA)
Use of medication ≥ 1 year prior to index hospitalization
Mortality data from Statistics Finland
Follow-up data from Hospital Discharge Register (HILMO)
Hospitalizations starting ≥ 1 year prior to index hospitalization
Information requested yearly bases

All relevant information on previous medical conditions, medication used, cardiovascular risk factors, laboratory measurements, coronary angiography, ECG, echocardiography and thorax X-ray was collected (Table 1). After completing the Corogene registry, the data were crosschecked with several methods, for example, by using the Finnish prescription register to check how many patients had bought statins before initial coronary angiogram. This number was compared with the Corogene-registry numbers of statin users. Contradictive cases were re-evaluated.

Weight was measured during hospitalization and height was asked from the patient. Body mass index (BMI; kg/m^2) was calculated from these values.

Patients were considered as having hypertension, dyslipidaemia or diabetes (insulin dependent and

non-insulin dependent combined), if they (i) had medication for these disorders or (ii) were told previously by a doctor that they suffer from these conditions, but had refused to use medication prescribed.

Individuals were defined as current smokers if they smoked tobacco during the previous 6 months or had quit smoking within the past 6 months. Those who had quit smoking >6 months ago were defined as ex-smokers.

Information on prior myocardial infarction, cardiomyopathy, arrhythmia, other vascular disease symptoms, other co-morbidities and vascular operations was obtained from patient history.

Information about prior percutaneous coronary interventions was obtained from the patients' cardiology data log. Coronary artery bypass graft (CABG)

surgery data were obtained from the surgery records of Helsinki University Central Hospital. This hospital is the only unit in the hospital district that performs CABG procedures. In a few cases when percutaneous coronary interventions or CABG surgery was done outside the Hospital District of Helsinki and Uusimaa, the patient history was used to confirm the operations.

Laboratory samples were drawn for gene, serum and plasma analyses from the arterial line during the angiogram. The blood samples were handled according to the laboratory standards of the Helsinki University Central Hospital (accredited laboratory). Altogether 70 ml of blood was drawn to EDTA (3 x 10 ml), citrate plasma (2 x 5 ml) and serum (2 x 10 ml) vacuum tubes.

Buffy coats for DNA and RNA extraction were taken from two fuded EDTA tubes. The third EDTA tube was used as whole blood for DNA extraction. EDTA plasma was separated and stored at -80°C in six aliquots for plasma and for proteomic analyses.

Serum was stored in four aliquots and citrate plasma in two aliquots at -80°C for later laboratory determinations, such as: lipids [total and HDL cholesterol, triglycerides, ApoA1, ApoA2, ApoB, Lp(a)], inflammation markers (high-sensitivity CRP, TNF- α , fibrinogen, complement system) and infection markers (antibodies against different pathogens).

Genome-wide association analyses

Of the Corogene cohort, 2500 patients with either ACS or previous myocardial infarction have been genotyped with the Illumina's Human 660 000 BeadChip that contains 660 000 markers in the Sanger Institute (Hinxton, Cambridge, UK). An equal number of controls, age-, geographical area- and sex-matched, has been selected from the FINRISK surveys^{9,10} for the genome-wide association studies.¹¹

RNA extracted from peripheral blood leucocytes is used to discover the genetic and epigenetic changes and modifications of the genes associated with cardiovascular disease risk factors or the disease itself. The study of epigenetic modifications such as DNA methylation, acetylation and microRNAs regulating the gene activities might open novel tools for disease prevention.

Human leucocyte antigen analyses

Human leucocyte antigen-classes I, II and III genes are studied using the most suitable method for this region involving SSP-PCR, genomic real-time PCR or direct sequencing. Genome-wide association data of this area are used to obtain additional genetic information. These data are being used to study immunology of heart diseases.

Oral health

From the whole cohort, we selected a 10% random sample for complete clinical oral examinations.

We investigated the association between CAD and oral health concentrating mainly on the parameters describing periodontal diseases. The oral examination was performed 6 weeks but ≤ 6 months after the initial hospitalization by two specialized periodontists and one radiologist, since all patients had both clinical and radiographic oral examinations, which included collecting saliva and subgingival samples for bacteriological analyses. The present 506 subjects comprise a subgroup called Parogene. Our goal is to invite these 500 subjects to re-examination 5 years after the initial hospitalization to evaluate the role of periodontal disease activity and putative changes in the periodontal microbiota on CAD and recurrent ACS.

For 2200 consecutive patients, an aortic valve echocardiogram was done separately just before or after the angiogram for aortic valve calcification study.

What has been found? Key findings and key publications

A vast amount of data has been collected, checked and stored in a database, and is now available for evaluation. Patient collection and exclusion data are shown in Figure 1. Baseline data are presented in Table 2 and [Supplementary Tables S1 and S2](#). Several manuscripts are in preparation issuing different aspects of the use of medication, reasons for death/death rate and genetics.

At the moment, the Corogene study has been used as a replication cohort in several large genome-wide association study consortia.^{11–17} These studies have revealed several new genetic loci of CAD, human height, weight, BMI and haematological parameters.

What are the main strengths and weaknesses?

The main strength of the study is the large population, and the quality of the data. The Corogene study is a very large, coronary angiogram-evaluated cohort. The coronary status information is solid with a potential to provide important insights into the mechanisms of CAD. An extensive amount of information has been registered in the baseline. We have an excellent estimation of eligible patients for the study, and our sample covers $>91\%$ of all eligible patients. The follow-up registers (Hospital Discharge Registry, Finnish prescription register and death statistics of Statistics Finland) in Finland are comprehensive, covering virtually all Finnish patients. Cumulative data of cohorts to be collected will not only increase the size of population, but also give depth to the data, answering the questions of possible risk profile and epigenetic changes.

Table 2 Patient characteristics in men and women

	Total	Men	Women	P-value	Risk ratio (95% CI)	NA
Number of patients	5295	3379	1916			0
Age groups (years)						
<40	81 (1.5)	68 (2.0)	13 (0.7)	<0.001	2.966 (1.643–5.354)	0
40–49	363 (6.9)	290 (8.6)	73 (3.8)	<0.001	2.253 (1.754–2.894)	0
50–59	1152 (21.8)	813 (24.1)	339 (17.7)	<0.001	1.360 (1.214–1.524)	0
60–69	1657 (31.3)	1123 (33.2)	534 (27.9)	<0.001	1.192 (1.094–1.300)	0
70–79	1556 (29.4)	861 (25.5)	695 (36.3)	<0.001	0.702 (0.647–0.763)	0
≥80	486 (9.2)	224 (6.6)	262 (13.7)	<0.001	0.485 (0.409–0.574)	0
BMI (kg/m ²)	27.4 ± 4.8	27.5 ± 4.5	27.1 ± 5.2	0.002	(0.164–0.722) ^a	25
Hypertension	3515 (66.4)	2138 (63.3)	1377 (71.9)	<0.001	0.880 (0.847–0.914)	3
Dyslipidaemia	3982 (75.3)	2516 (74.5)	1466 (76.6)	0.091	0.973 (0.942–1.004)	7
Diabetes	1074 (20.3)	714 (21.1)	360 (18.8)	0.042	1.125 (1.004–1.260)	0
Smoking						
Current smoker	1287 (24.5)	954 (28.4)	333 (17.5)	<0.001	1.622 (1.451–1.813)	33
Ex-smoker	1873 (35.6)	1445 (43.0)	428 (22.5)	<0.001	1.911 (1.743–2.095)	33
CAD groups						
No CAD	1202 (22.7)	600 (17.8)	602 (31.4)	<0.001	0.565 (0.512–0.623)	0
Stable CAD	1799 (34.0)	1264 (37.4)	535 (27.9)	<0.001	1.340 (1.232–1.457)	0
ACS	2090 (39.5)	1454 (43.0)	636 (33.2)	<0.001	1.296 (1.203–1.396)	0
Other ischaemic events	204 (3.9)	61 (1.8)	143 (7.5)	<0.001	0.242 (0.180–0.325)	0
Prior AMI	1099 (20.9)	776 (23.1)	323 (16.9)	<0.001	1.362 (1.211–1.531)	26
Cardiomyopathy	263 (5.1)	201 (6.1)	62 (3.3)	<0.001	1.830 (1.385–2.419)	124
Arrhythmia						
Atrial flutter	825 (15.7)	527 (15.7)	298 (15.6)	0.970	1.003 (0.880–1.142)	24
Other	582 (11.0)	373 (11.1)	209 (11.0)	0.886	1.012 (0.862–1.187)	24
Other vascular disease symptoms						
Claudication	536 (10.2)	353 (10.5)	183 (9.6)	0.291	1.095 (0.925–1.298)	30
Stroke/TIA	662 (12.6)	413 (12.3)	249 (13.1)	0.429	0.942 (0.814–1.092)	29
Prior cardiovascular operations						
Prior PCI	676 (12.8)	480 (14.2)	196 (10.2)	<0.001	1.389 (1.188–1.624)	0
Prior CABG	500 (9.4)	396 (11.7)	104 (5.4)	<0.001	2.159 (1.753–2.660)	0
Other	186 (3.5)	115 (3.4)	71 (3.7)	0.569	0.919 (0.687–1.229)	2
Prior vascular operations						
Lower extremity	184 (3.5)	131 (3.9)	53 (2.8)	0.034	1.402 (1.024–1.919)	20
Other	106 (2.0)	85 (2.5)	21 (1.1)	<0.001	2.295 (1.428–3.688)	19
Co-morbidities						
Renal disease	234 (4.4)	162 (4.8)	72 (3.8)	0.078	1.276 (0.972–1.674)	2
Malignancy	462 (8.7)	269 (8.0)	193 (10.1)	0.009	0.790 (0.663–0.943)	3
Rheumatoid disease	231 (4.4)	97 (2.9)	134 (7.0)	<0.001	0.410 (0.318–0.530)	2
Psychiatric disorder	121 (2.3)	59 (1.7)	62 (3.2)	<0.001	0.540 (0.380–0.768)	1

Mean ± standard deviation is shown for continuous variables (BMI) and number (%) for categorical variables (all others). Percentage is calculated of those with information available in each group. P-values for difference between men and women are calculated by Student's *t*-test (continuous variables) or chi-square test (categorical variables).

CI, confidence interval; NA, information not available; AMI, acute myocardial infarction; TIA, transient ischaemic attack; PCI, percutaneous coronary intervention.

^a95% CI of the difference in Student's *t*-test.

The main weakness of the study is a collection bias. The method used rules out patients passing away before hospitalization. We have overcome this problem by collaborating with the Department of Forensic Medicine, where all patients who have died outside the hospital without diagnosis are autopsied. Also, we do not get the full picture of coronary disease, because we do not see the patients left who receive only conservative treatment. Generalization of the study findings is limited to White Caucasian subjects.

Another weakness is the self-reported data. We did not measure patient height, although the weight was measured. Smoking habit was also based on the history and therefore ascertained with some error. Other risk factors such as hypertension, diabetes and dyslipidaemia were both self-reported and obtained from hospital records. The reason for that was to cross-check self-reported data and to verify its reliability.

A third weakness is in the genome-wide association study. Our initial plan was to use a case-control design, which is potentially open to confounding if risk factor information is collected in a different way between cases and controls. Now we have the possibility to overcome this problem by using over 15 000 Finnish individuals from population-based controls.

Can I get hold of the data? Where can I find out more?

The Corogene study welcomes proposals for new collaborations, which should be made to Dr Juha Sinisalo (juha.sinisalo@hus.fi). Specific parts of the unidentified data can be transmitted to collaboration partners.

Supplementary Data

Supplementary Data are available at *IJE* online.

Funding

Aarno Koskelo Foundation (partial); the Finnish Foundation for Cardiovascular Research (partial); the Paulo Foundation (partial); Jenny and Antti Wihuri Foundation (partial); EVO funds of Helsinki University Central Hospital; Wellcome Trust, UK (the genetic part).

Acknowledgements

The authors greatly appreciate the technical assistance of Mervi Pietilä, RN, and Sirpa Stick, RN, and the National Institute for Health and Welfare DNA logistics group, especially Ms Minttu Jussila for DNA handling and also Ms Leena Saraste for

language editing. The authors wish to commemorate the late Professor Leena Peltonen-Palotie, whose role was pivotal in the initiation of the molecular genetic part of the study.

Conflicts of interest: None declared.

References

- 1 Ford ES, Ajani UA, Croft JB *et al.* Explaining the decrease in U.S. deaths from coronary disease, 1980–2000. *N Engl J Med* 2007;**356**:2388–98.
- 2 Capewell S, Hayes DK, Ford ES *et al.* Life-years gained among US adults from modern treatments and changes in the prevalence of 6 coronary heart disease risk factors between 1980 and 2000. *Am J Epidemiol* 2009;**170**:229–36.
- 3 Ordovás JM, Smith CE. Epigenetics and cardiovascular disease. *Nat Rev Cardiol* 2010;**7**:510–19.
- 4 Bassand J-P, Hamm CW, Ardissino D, Boersma E. Guidelines for the diagnosis and treatment of non-ST-segment elevation acute coronary syndromes. The Task Force for the Diagnosis and Treatment of Non-ST-Segment Elevation Acute Coronary Syndromes of the European Society of Cardiology. *Eur Heart J* 2007;**28**:1598–660.
- 5 Fox K, Angeles M, Garcia A, Ardissino D, Buszman P. Guidelines on the management of stable angina pectoris. The Task Force on the Management of Stable Angina Pectoris of the European Society of Cardiology. *Eur Heart J* 2006;**27**:1341–81.
- 6 Thygesen K, Alpert JS, White HD. Joint ESC/ACCF/AHA/WHF Task Force for the Redefinition of Myocardial Infarction. Universal definition of myocardial infarction. *Eur Heart J* 2007;**28**:2525–38.
- 7 Mahonen M, Salomaa V, Keskimäki I, Moltchanov V. The feasibility of routine mortality and morbidity register data linkage to study the occurrence of acute coronary heart disease events in Finland. The Finnish Cardiovascular Diseases Registers (CVDR) Project. *Eur J Epidemiol* 2000;**16**:701–11.
- 8 Helin-Salmivaara A, Lavikainen PT, Korhonen MJ *et al.* Pattern of statin use among 10 cohorts of new users from 1995 to 2004: A register-based nationwide study. *Am J Manag Care* 2010;**16**:116–22.
- 9 Vartiainen E, Jousilahti P, Alftan G, Sundvall J, Pietinen P, Puska P. Cardiovascular risk factor changes in Finland, 1972–1997. *Int J Epidemiol* 2000;**29**:49–56.
- 10 Alemao E, Yin D, Sintonen H, Salomaa V, Jousilahti P. Evaluation of lipid-lowering therapy and cholesterol goal attainment in Finland. The National FINRISK Study. *Am J Cardiovasc Drugs* 2006;**6**:349–55.
- 11 Soranzo N, Spector TD, Mangino M *et al.* A meta-analysis of eight hematological parameters identifies 22 associated loci and extensive disease pleiotropy on chromosome 12q24. *Nat Genet* 2009;**41**:1182–90.
- 12 Ripatti S, Tikkanen E, Orho-Melander M *et al.* A multi-locus genetic risk score and risk for incident coronary heart disease. *Lancet* 2010;**376**:1393–400.
- 13 Heid IM, Jackson AU, Randall JC *et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio

- and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet* 2010;**42**:949–60.
- ¹⁴ Lango AH, Estrada K, Lettre G *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010;**467**:832–38.
- ¹⁵ Speliotes EK, Willer CJ, Berndt SI *et al.* Genome-wide association study of 249,796 individuals reveals eighteen loci associated with body mass index. *Nat Genet* 2010;**42**: 937–48.
- ¹⁶ Teslovich TM, Musunuru K, Smith AV *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010;**466**:707–13.
- ¹⁷ Peden JF, Hopewell JC, Saleheen D. A genome-wide association study in Europeans and South Asians identifies five novel loci for coronary artery disease. The Coronary Artery Disease (C4D) Genetics Consortium. *Nat Genet* 2011;**43**:339–44.