Heart Online First, published on November 10, 2012 as 10.1136/heartjnl-2012-303001

Understanding coronary artery disease using twin studies

Massimo Mangino, Tim Spector

Department of Twin Research & Genetic Epidemiology, King's College London, London, UK

Correspondence to
Dr Tim Spector, Department of Twin Research & Genetic Epidemiology, King's College London, St Thomas' Hospital Campus, Lambeth Palace Road, London SE1 7EH, UK; tim.spector@kcl.ac.uk

MM and TS contributed equally.

ABSTRACT

Twins have fascinated human communities since the beginning of recorded history. In cardiovascular research, twin studies played a pivotal role in the discovery of the genetic basis of coronary artery disease, myocardial infarction and most of their associated risk factors. Matched for age, prenatal and postnatal environmental factors, discordant monozygous twin pairs provide a means to enhance the power to analyse epigenetic and metagenomic mechanisms. In the near future, combining the monozygous discordant twin study design with the recent advance of the sequencing technologies it will be possible to explore the complexity of the gene-environment relationships and individual variability to provide important insights into the pathogenesis of coronary artery disease/myocardial infarction.

One of the main objectives of modern medicine is to use genetic information in clinical practice to predict, diagnose and treat common human diseases.

Coronary Artery Disease (CAD) and its main complication Myocardial Infarction (MI) are the leading cause of death worldwide. In USA, one in every six deaths is currently caused by CAD.[1] By 2030, 40.5% of the US population is expected to suffer from some form of CAD disease, with an estimated direct medical cost of $818 billion and an increase of 61% of indirect costs, due to lost productivity.[2]

The first recognition of a genetic contribution to CAD was reported by an Irish physician, Samuel Black in the early 1800s[3]:

> We have seen that the disease appears to be connected with a plethoric state of the system and with obesity: that the great majority of the subjects of it have belonged to the better ranks of society, who were in the habit of sitting down everyday to a plentiful table, in the pleasures of which they may have indulged to a greater extent than was suitable to the tendency of their constitution …

Indeed, the observation that CAD/MI 'runs in families' alone is not enough to infer its genetic aetiology.

Twins have fascinated human communities since the beginning of recorded history and feature in the legends and myths of many cultures.[4] In the context of medical research, twins provide invaluable opportunities to identify the genetic component of diseases. In fact, classical twin studies supplied the first evidence of a genetic basis of CAD/MI and its associated risk factors (eg, obesity, hypertension, lipids levels and diabetes).[5–7]

Twin studies compare the concordance of a trait/disease between monozygous (MZ) twin pairs, which are genetically identical (as they result from the division of a single fertilised egg), and non-identical or dizygous (DZ) twin pairs who share, on average, 50% of their genome (they are formed from the separate fertilisation of two eggs). Because both MZ and DZ twin pairs are exposed to similar prenatal and postnatal environmental factors (equal environment assumption),[8 9] if the MZ are more similar than the DZ when compared for a particular trait/disease, then a genetic origin can be inferred. A very simple method to estimate the heritability (h^2) of a trait is to double the difference between MZ and DZ correlations ($h^2=2*(r_{MZ}-r_{DZ})$).[10] It is also possible to determine the proportion of the variance that is due to a shared environment for MZ ($r_{MZ}-h^2$) and DZ ($r_{DZ}-h^2/2$).

The twin design incorporates several important properties which are particularly relevant to the analysis of complex traits/diseases. For example, for CAD, as for most complex diseases, age is crucially important for the onset/expression. Twins' studies, compared with family studies, are more powerful because they can take advantage of a unique characteristic of twins: their matching age (and date of birth). Twins are exposed to higher degree of family environment compared with sib pair (eg, lifestyle, diet).[8 9] Sharing a range of environmental variables, both measured and unmeasured, do contribute to the expression of complex traits/diseases. Therefore, it is much easier to attribute phenotypic differences between twins to genetic rather than to environmental factors. Finally, an additional advantage of twin studies is that non-paternity, an important cause of error in sib pair analysis is reduced to almost nil.

In 1994 Marenberg *et al*[5] studied the concordance on 10 502 Swedish twin pairs to investigate the genetic basis of CAD mortality. Their results showed that, if the co-twin had died from early-onset (presenting before the age of 55) CAD, the relative hazard (RH) of death by CAD was double in male MZ twins (RH=8.1) compared to male DZ twins (RH=3.8). In females the effect was even stronger with a RH nearly six times higher ($RH_{MZ}=15$; $RH_{DZ}=2.6$). These findings clearly highlighted the presence of a genetic component in CAD/MI. The authors validated their results in a 36 years follow-up of the same cohort, showing that CAD/MI mortality is heritable, with h^2 values ranging from 0.38 in females to 0.57 in males.[11] Although the estimates of male h^2 were consistent with the Swedish findings in a subsequent study of 7955 pairs from a Danish twin

cohort (h^2=0.53), the female twins had equivalent heritability to males (h^2=0.53).[12] This discrepancy may be in part explained by different methods of sample selection in the two twin cohorts.[11]

In the last 40 years twin studies have also highlighted the genetic component of a number of CAD/MI risk factors including lipid levels[7] blood pressure,[13] smoking,[14] body mass index,[15] physical activity,[16] C reactive protein,[17] plasma homocysteine[18] and diabetes[19] (table 1).

Thus, twin studies provided enough evidence for a genetic aetiology of CAD/MI to launch the hunt for the underlying disease susceptibility genes for the disease and the intermediate traits.

In the last 5 years geneticists have exploited high-throughput technologies to simultaneously genotype hundreds of thousands of single nucleotide polymorphisms in large case-control datasets. Since 2007, these genome wide association studies (GWAS) have identified and validated at least 36 susceptibility loci for CAD/MI.[20]

Twins, who played a pivotal role in the discovery of the genetic basis of CAD/MI, unlike for linkage studies, were not maximally exploited during the GWAS era because MZ twins are genetically identical (and therefore not independent) and DZ twins are not different from ordinary siblings.

Table 1 Heritability of the most common risk factors associated with CAD/MI (Coronary Artery Disease/myocardial infarction)

	Heritability (%)	Reference
High-density lipoprotein		
Male	69	7
Females	69	
Low-density lipoprotein		
Male	77	7
Females	69	
Triglycerides		
Male	69	7
Females	51	
Total cholesterol		
Male	77	7
Female	66	
Apolipoprotein AI		
Male	46	7
Females	55	
Apolipoprotein B		
Male	65	7
Females	59	
Lipoprotein(a)		
Male	89	7
Females	89	
Systolic blood pressure	53	13
Diastolic blood pressure	48	13
Smoking (cigarettes per day)	86	14
Body mass index		
Male (17 years)	83	15
Females (17 years)	74	
Physical Activity		
Leisure time	55	16
Occupational	54	
C reactive protein	52	17
Plasma homocysteine	57	18
Diabetes		
Type I	75	19
Type II	75	

Does this mean that twin studies have no additional value in medical genetics?

The answer is: 'not at all'.

The GWAS era has left open a considerable number of questions. Most of the susceptibility loci identified by GWAS studies typically account for only a small part of the known heritability—usually less than 5% (the so called 'case of missing heritability').[21] So far, a number of explanations have been proposed to account for this phenomenon including: (1) the effect of rare low frequency variants and structural variation such as copy number variants that are not captured by GWAS; (2) the occurrence of gene-gene interaction and epigenetic modifications.[21]

In this scenario twin studies acquire novel importance, as they may allow the analysis of epigenetic and metagenomic mechanisms which may offer a partial explanation for the phenomenon of 'missing heritability'.

The often overlooked result of twin studies is that, despite high heritability a common disease and similar lifestyles, most identical twins don't usually both contract coronary heart disease and will on average die of different causes.[5 22] This shows that structural gene variants alone, even if all were known, cannot explain or predict CAD. Other processes must be at work. The modern definition of the term 'epigenetics' is the study of the heritable changes in gene expression which are not due to changes of the DNA sequence.[23] Epigenetic mechanisms, including DNA methylation and modification of histone proteins, are essential and reversible regulators of gene transcription in complex organisms. The study of epigenetics is one of the fastest growing fields.

In one of the first studies based on multi-site DNA methylation, Fraga *et al*[24] showed that MZ twins have different epigenetic profiles and that on average these differences increase as they become older. Some of these differences arise as genes for growth and development are less tightly controlled in older people—adding variability. However other key genes involved in aging processes may get more methylated as people age—as has been seen with older twins.[25]

Importantly a number of studies have reported that CAD/MI and some of the associated risk factors are the result of altered patterns of DNA methylation.[26–29]

With the advance of the sequencing technologies, it is now possible to analyse genome wide epigenetic profiles at a high resolution and at a relatively low cost.[30] In the near future, the whole genome epigenetic profiling of phenotypically discordant MZ twins, as is ongoing in the EpiTwin study (http://www.epitwin.eu), will allow us to obtain a more complete understanding of the functional impact of epigenetic modifications on complex diseases in general and CAD/MI in particular.

Another new emerging omics field, in which MZ discordant twin studies could be an invaluable resource is metagenetics—the study of the metagenome.

Humans are permanently populated by complex site-specific microbial communities and/or infected by viruses without obvious negative effects. The combined genes of host (human) and pathogens (microbiome) represent our metagenome and these 100 trillion microbes have 20 times more genes than us humans.

There is growing evidence that the interaction between microbiome and the host immune system contributes to the pathogenesis of inflammatory diseases in individuals carrying specific genetic variants.[31]

A recent study of MZ twins discordant for obesity, showed that: (1) the gut microbiome is shared among families, (2) the

microbial community varies in each person for specific bacterial lineage and (3) obesity is associated with phylum-level changes in the microbiome (the obese host microbiome was enriched in gene categories involved in carbohydrate and lipid metabolism) and reduced bacterial diversity.[32]

Two recent studies indicated that type 2 diabetes in humans is associated with compositional changes in intestinal microbiome[33] and in obese mice metabolic diseases are connected with the presence of Gram-negative bacteria in the gut.[34]

Furthermore, a prospective study on 15 273 Swedish twins highlighted the link between oral bacteria and CAD.[35]

Thanks to recent advances, it is now possible to explore the metagenome biodiversity at a scale that allows for the discovery of virtually all of the microbiome present in an individual—whether in the bowels, skin or oral cavities. Combining this powerful new technology with the MZ discordant twin study design will allow us to better understand the link between metagenome and common multifactorial diseases.

In conclusion, over the course of the last century, the classical twin design has provided an invaluable tool to evaluate the genetic basis of CAD/MI. The challenge is now to dissect out the complexity of the gene-environment relationships and individual variability. In the future, discordant twin studies in particular are expected to provide important insights into the pathogenic role of epigenetic modifications and metagenetic interactions in heart disease.

Acknowledgements MM receives support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. TS is an NIHR senior Investigator and is holder of an ERC Advanced Principal Investigator award.

Funding National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London.

Competing interests None.

Provenance and peer review Commissioned; internally peer reviewed.

REFERENCES

1. **Roger VL,** Go AS, Lloyd-Jones DM, *et al*. Heart disease and stroke statistics—2012 update: a report from the American Heart Association. *Circulation* 2012;**125**:e2–220.
2. **Heidenreich PA,** Trogdon JG, Khavjou OA, *et al*. Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association. *Circulation* 2011;**123**:933–44.
3. **Evans A.** Dr Black's favourite disease. *Br Heart J* 1995;**74**:696–7.
4. **Stewart EA.** *Exploring twins: towards a social analysis of twinship*. Palgrave Macmillan, 2000.
5. **Marenberg ME,** Risch N, Berkman LF, *et al*. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med* 1994;**330**:1041–6.
6. **Samuels LE,** Samuels FS, Thomas MP, *et al*. Coronary artery disease in identical twins. *Ann Thorac Surg* 1999;**68**:594–600.
7. **Snieder H,** van Doornen LJ, Boomsma DI. The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am J Hum Genet* 1997;**60**:638–50.
8. **Derks EM,** Dolan CV, Boomsma DI. A test of the equal environment assumption (EEA) in multivariate twin studies. *Twin Res and Hum Genet: Offic J Int Soc Twin Stud* 2006;**9**:403–11.
9. **Kendler KS,** Neale MC, Kessler RC, *et al*. A test of the equal-environment assumption in twin studies of psychiatric illness. *Behav Genet* 1993;**23**:21–7.
10. **Falconer DS.** *Introduction to quantitative genetics*. Harlow, England: Longman, 1989.
11. **Zdravkovic S,** Wienke A, Pedersen NL, *et al*. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J Int Med* 2002;**252**:247–54.
12. **Wienke A,** Holm NV, Skytthe A, *et al*. The heritability of mortality due to heart diseases: a correlated frailty model applied to Danish twins. *Twin Res: Offic J Int Soc Twin Stud* 2001;**4**:266–74.
13. **Evans A,** Van Baal GC, McCarron P, *et al*. The genetics of coronary heart disease: the contribution of twin studies. *Twin Res: Offic J Int Soc Twin Stud* 2003;**6**:432–41.
14. **Koopmans JR,** Slutske WS, Heath AC, *et al*. The genetics of smoking initiation and quantity smoked in Dutch adolescent and young adult twins. *Behav Genet* 1999;**29**:383–93.
15. **Lajunen HR,** Kaprio J, Keski-Rahkonen A, *et al*. Genetic and environmental effects on body mass index during adolescence: a prospective study among Finnish twins. *Int J Obes* 2009;**33**:559–67.
16. **Eriksson M,** Rasmussen F, Tynelius P. Genetic factors in physical activity and the equal environment assumption—the Swedish young male twins study. *Behav Genet* 2006;**36**:238–47.
17. **MacGregor AJ,** Gallimore JR, Spector TD, *et al*. Genetic effects on baseline values of C-reactive protein and serum amyloid a protein: a comparison of monozygotic and dizygotic twins. *Clin Chem* 2004;**50**:130–4.
18. **Siva A,** De Lange M, Clayton D, *et al*. The heritability of plasma homocysteine, and the influence of genetic variation in the homocysteine methylation pathway. *QJM: Monthly J Assoc Physicians* 2007;**100**:495–9.
19. **Kaprio J,** Tuomilehto J, Koskenvuo M, *et al*. Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia* 1992;**35**:1060–7.
20. **Roberts R,** Stewart AF. Genetics of Coronary Artery Disease in the 21st Century. *Clin Cardiol* 2012;**35**:536–40.
21. **Maher B.** Personal genomes: The case of the missing heritability. *Nature* 2008;**456**:18–21.
22. **Zdravkovic S,** Wienke A, Pedersen NL, *et al*. Genetic influences on CHD-death and the impact of known risk factors: comparison of two frailty models. *Behav Genet* 2004;**34**:585–92.
23. **Richards EJ.** Inherited epigenetic variation—revisiting soft inheritance. *Nat Rev Genet* 2006;**7**:395–401.
24. **Fraga MF,** Ballestar E, Paz MF, *et al*. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 2005;**102**:10604–9.
25. **Bell JT,** Tsai PC, Yang TP, *et al*. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet* 2012;**8**:e1002629.
26. **Guay SP,** Brisson D, Munger J, *et al*. ABCA1 gene promoter DNA methylation is associated with HDL particle profile and coronary artery disease in familial hypercholesterolemia. *Epigenet: Offic J DNA Methylation Soc* 2012;**7**:464–72.
27. **Friso S,** Lotto V, Choi SW, *et al*. Promoter methylation in coagulation F7 gene influences plasma FVII concentrations and relates to coronary artery disease. *J Med Genet* 2012;**49**:192–9.
28. **Sharma P,** Kumar J, Garg G, *et al*. Detection of altered global DNA methylation in coronary artery disease patients. *DNA Cell Biol* 2008;**27**:357–65.
29. **Stenvinkel P,** Karimi M, Johansson S, *et al*. Impact of inflammation on epigenetic DNA methylation—a novel risk factor for cardiovascular disease? *J Int Med* 2007;**261**:488–99.
30. **Baranzini SE,** Mudge J, van Velkinburgh JC, *et al*. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature* 2010;**464**:1351–6.
31. **Virgin HW,** Todd JA. Metagenomics and personalized medicine. *Cell* 2011;**147**:44–56.
32. **Turnbaugh PJ,** Hamady M, Yatsunenko T, *et al*. A core gut microbiome in obese and lean twins. *Nature* 2009;**457**:480–4.
33. **Larsen N,** Vogensen FK, van den Berg FW, *et al*. Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PloS One* 2010;**5**:e9085.
34. **Cani PD,** Neyrinck AM, Fava F, *et al*. Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia. *Diabetologia* 2007;**50**:2374–83.
35. **Mucci LA,** Hsieh CC, Williams PL, *et al*. Do genetic factors explain the association between poor oral health and cardiovascular disease? A prospective study among Swedish twins. *Am J Epidemiol* 2009;**170**:615–21.

Understanding coronary artery disease using twin studies

Massimo Mangino and Tim Spector

Heart published online November 10, 2012
doi: 10.1136/heartjnl-2012-303001