Understanding coronary artery disease using twin studies

Massimo Mangino, Tim Spector

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 monzygous twin pairs provide a means to enhance the power to analyse epigenetic and metagenomic mechanisms. In the near future, combining the monzygous 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. By 2050, 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.

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

We have seen that the disease appears to be connected with a plethoraic 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. 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).

Twin studies compare the concordance of a trait/disease between monzygous (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), 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 ($h^2=2(\rho_{MZ}-\rho_{DZ})$). It is also possible to determine the proportion of the variance that is due to a shared environment for MZ ($\rho_{MZ}-h^2$) and DZ ($\rho_{DZ}-h^2/2$).

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). 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 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. Although the estimates of male $h^2$ were consistent with the Swedish findings in a subsequent study of 7955 pairs from a Danish twin...
In the last 40 years twin studies have also highlighted the genetic component of a number of CAD/MI risk factors including lipid levels, blood pressure, smoking, body mass index, physical activity, C reactive protein, plasma homocysteine and diabetes (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 56 susceptibility loci for CAD/MI.

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)

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Heritability (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-density lipoprotein</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Male</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Apolipoprotein Al</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Male</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Lipoprotein(a)</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Male</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>53</td>
<td>13</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>48</td>
<td>13</td>
</tr>
<tr>
<td>Smoking (cigarettes per day)</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>Body mass index</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>Male (17 years)</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>Females (17 years)</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Physical Activity</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Leisure time</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>Occupational</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>C reactive protein</td>
<td>52</td>
<td>17</td>
</tr>
<tr>
<td>Plasma homocysteine</td>
<td>57</td>
<td>18</td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>Type I</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Type II</td>
<td>75</td>
<td></td>
</tr>
</tbody>
</table>

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’). 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.

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. 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. 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 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.

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.

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. 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 metagenomics—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.

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.

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

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

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


Understanding coronary artery disease using twin studies

Massimo Mangino and Tim Spector

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

Updated information and services can be found at:
http://heart.bmj.com/content/early/2012/11/09/heartjnl-2012-303001.full.html

*These include:*

**References**
This article cites 33 articles, 8 of which can be accessed free at:
http://heart.bmj.com/content/early/2012/11/09/heartjnl-2012-303001.full.html#ref-list-1

**Email alerting service**
Published online November 10, 2012 in advance of the print journal.
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/