From the Department of Medical Epidemiology,
Karolinska Institutet, Stockholm, Sweden

GENETIC EPIDEMIOLOGICAL APPROACHES TO THE STUDY OF RISK FACTORS FOR CARDIOVASCULAR DISEASES

Anastasia Iliadou

Stockholm 2003

Published and printed by Karolinska University Press
Box 200, SE-171 77 Stockholm, Sweden

ISBN 91-7349-548-4

Τα παντα ρει" Ηρακλιτος
"Everything is relative" Heraklitos

To my dearest grandfather

SUMMARY

Cardiovascular disease (CVD) is one of the most common causes of death in most developed countries. Genetic and environmental factors influence CVD and its risk factors in a complex fashion. Parallel to this, new statistical methods are being developed to make use of the new molecular genetic information. The aim of this thesis was focused on developing and implementing these new statistical techniques and applying them to twin data from the population-based Swedish Twin Registry in order to study risk factors of CVD.

In the first study we investigated genetic and environmental effects in the variation of lipids and apolipoproteins in 725 twin pairs. Structural equation modelling, which also tested for age and gender effects were applied to the data. Heritabilities for lipids and apolipoproteins ranged from 35-74%. Consistent sex differences were found in triglycerides. Total phenotypic variation increased across the age groups for Cholesterol and Apolipoprotein B due to an increase in unique environmental variance components, which is probably due to the accumulation of environmental experiences throughout life. In contrast, in Apolipoprotein A1 variance was highest in the middle age group with changes due to differences in genetic variance. This could imply that different genetic mechanisms might act at different time points in life.

In the second study we investigated genetic and environmental effects in repeated measures of diastolic and systolic blood pressure in a sample of 298 twin pairs. With the use of a Cholesky decomposition model we found that genetic influences were stable over time and they explained up to 46% of the phenotypic variance in diastolic and 63% of the phenotypic variance in systolic blood pressure. We also investigated the association between blood pressure and two polymorphisms, the angiotensin-I converting enzyme gene (ACE) and the angiotensin II type 1 receptor (AT_1R) gene with a novel approach that simultaneously tests for linkage and association. The method tests whether the polymorphisms are the true QTLs or in linkage disequilibrium with the true QTLs. No association was found.

Linkage to obesity has been previously reported on chromosome 2 and 10. In the third study we investigated the findings in candidate regions on Chromosome 2 and 10 in a selected (based on their phenotypic value) sample of Swedish twin pairs. It is shown that sampling siblings in that way is a more cost-effective approach of doing linkage. We implemented a "combined" Haseman-Elston approach that is especially powerful and developed for studying selected samples. No linkage was confirmed in the candidate regions of chromosome 2 and 10.

The Barker hypothesis states that metabolic disorders such as cardiovascular disease, blood pressure and diabetes begin as a sub optimal or abnormal development during foetal and early neonatal life. There have been criticisms to the Barker hypothesis, above all because the genetic and early environmental influences could be responsible for previously observed associations. Twins are ideal for testing these alternative hypotheses. We investigated the relationship between self-reported birth weight and Type II diabetes, first in a cohort of 11 226 same-sexed Swedish twins, and secondly in 142 pairs discordant for Type II diabetes by utilizing the co-twin control method. We found a direct effect of low birth weight on Type II diabetes that is independent of genetic and early environmental effects.

In conclusion, the incorporation of molecular data in order to disentangle the complexity underlying biological mechanisms and the statistical models developed to accompany them will in the future be a challenging task.

CONTENTS

LIST OF PUBLICATIONS

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals:

I. Iliadou A, Lichtenstein P, de Faire U, Pedersen NL.
Variation in genetic and environmental influences in serum lipid and apolipoprotein levels across the lifespan in Swedish male and female twins.
Am J Med Genet. 2001 Jul 22;102(1):48-58.

II. Iliadou A, Lichtenstein P, Morgenstern R, Forsberg L, Svensson R, de Faire U, Martin NG and Pedersen NL.
Repeated Blood Pressure Measurements in a Sample of Swedish Twins: Heritabilities and Associations with Polymorphisms in the Renin-Angiotensin-Aldosteron System
Journal of Hypertension 2002, 20:1543-1550.

III. Iliadou A, Lichtenstein P, Ahlberg S, Hoffstedt J, Arner P, Schalling M, Pedersen N L, Lavebratt C
No linkage to obesity in candidate regions of chromosome 2 and 10 in a selected sample of Swedish twins.
Twin Research, 2003 Apr; 6(2):162-169.

IV. Iliadou A, Cnattingius S, Lichtenstein P.
Low birth weight and Type II diabetes: A study on 11 226 Swedish twins.
(*Submitted*)

LIST OF ABBREVIATIONS

MZ	monozygotic
DZ	dizygotic
CVD	cardiovascular disease
QTL	quantitative trait locus
IBD	identity by descent
BMI	body mass index
OR	odds ratio
CI	confidence interval

INTRODUCTION

Although many cardiovascular diseases (CVDs) can be treated or prevented, an estimated 17 million people worldwide die of CVDs each year (WHO). The risk increases with age and is larger for women than for men. According to the WHO Collaborating Centre on Surveillance of Cardiovascular Diseases, 49% of all causes of death in Sweden in 1996 were due to CVD (Fig. 1). The proportional mortality for a range of pathological processes, including ischemic heart diseases, cerebrovascular diseases, hypertensive diseases and other CVDs can also be seen in figure 1.

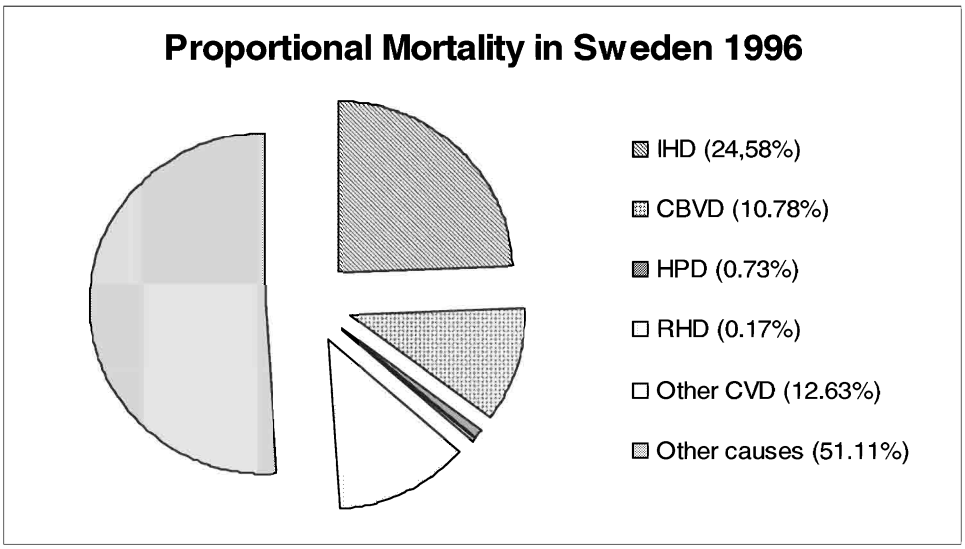


Figure 1. CVD: Cardiovasuclar Diseases, IHD: Ischemic Heart Diseases, CBVD: Cerebrovascular Diseases, HPD: Hypertensive Diseases, RHD: Chronic Rheumatic Heart Diseases
Source: *http://cvdinfobase.ic.gc.ca*

In order to better understand the causes behind the development of CVDs, we need to explore and dissect the genetic and environmental origins of old as well as novel risk factors. Parallel to the rapid development of new molecular techniques, the need for statistical tools to accompany them has become increasingly important. In this thesis I have explored and implemented statistical techniques and applied them to twin data from the population-based Swedish Twin Registry in studies of risk factors for CVD.

BACKGROUND

Several risk factors have been associated with the development of cardiovascular diseases. Age, sex and family history are non-modifiable factors. That is, they are not amenable to intervention. Modifiable factors exist only as the result of a conscious decision about life-style such as cigarette smoking, sedentary living and diet. Factors that are subject to quantitative changes and have been studied in this thesis are lipid levels, blood pressure, obesity, and Type II diabetes.

Cholesterol

Total blood cholesterol is the most common measurement of blood lipid disturbance. Its main particles are Low Density Lipoprotein (LDL), and High Density Lipoprotein (HDL) (American Heart Association, 2003; Goldbourt, de Faire, & Berg, 1994; Nora, Berg, & Nora, 1991).

When too much LDL cholesterol circulates in the blood, it can slowly build up in the walls of the arteries. Together with other substances it can form plaque, a thick, hard deposit that can clog those arteries. This condition is known as atherosclerosis. If a clot forms and blocks a narrowed heart or brain artery, it can cause a myocardial infarction or a stroke. LDL cholesterol of less than 100 mg/dL is the optimal level. A high LDL level (more than 160 mg/dL without other risk factors or more than 130 mg/dL if you have two or more additional risk factors for cardiovascular disease) conveys an increased risk of heart disease (European Societies on coronary prevention, 1998; The National Cholesterol Education Program, 2001).

About one-third to one-fourth of blood cholesterol is carried by high-density lipoprotein (HDL). HDL cholesterol is known as the "good" cholesterol because high levels seem to protect against myocardial infarction. Low HDL cholesterol levels (less than 40 mg/dL) increase the risk of heart disease. The biological mechanism whereby clinical disease is produced is not entirely clear. One prevailing theory is that HDL removes excess cholesterol from plaque in arteries, thus slowing the build-up.

Apolipoprotein B and A1

Lipids are transported in the circulation by lipoproteins, macromolecular complexes that consist of lipids and proteins, termed apolipoproteins (Goldbourt et al., 1994). Apolipoproteins serve a

variety of physiological functions in lipoprotein metabolism. Apolipoprotein B is a primary protein component of LDL particles and hence an important risk factor for CVD. Apolipoprotein A1, on the other hand, is a major component of HDL particles and is inversely associated with risk of CVD (Rader & Brewer, 1994; Schulte, Rothman, & Austin, 1993).

Triglycerides

Triglycerides are fatty substances in the blood carried in lipoproteins, which also carry cholesterol (American Heart Association, 2003). Fatty substances come from food and are also metabolized from other energy sources, like carbohydrates (American Heart Association, 2003). Calories in a meal, that are not used immediately, are converted to triglycerides and transported to fat cells to be stored. When the body needs energy fat is released from tissues by hormones. People with high triglycerides often have low HDL cholesterol levels as well. Many people with heart disease, diabetes and obesity have high triglyceride levels. Triglyceride levels of 200 mg/dL and above are considered high, and may need treatment in some people. However, many studies have examined the association between triglycerides and heart disease, and have not yet been able to determine whether triglyceride levels independently can directly predict heart disease risk (Austin, Hokanson, & Edwards, 1998).

Blood pressure

Blood pressure is the force exerted against artery walls as blood is carried through the circulatory system. The measurement of force is made in relation to the heart's pumping activity, and is measured in millimetres of mercury (mmHg). Systolic pressure is the measurement of pressure that occurs when the heart contracts or beats. Diastolic pressure is the measurement recorded between beats, when the heart muscle is at a relaxed stage.

High blood pressure, or hypertension, is a major health problem affecting about 20% of the adult population in most countries (American Heart Association, 2003). Left untreated, hypertension can cause many complications. Artery walls thicken and harden. The elasticity or stretchiness in the arteries decreases as well, requiring the heart to work harder to pump blood through the arteries. Furthermore, the increased strain on the heart wall leads to hypertrophy of the left ventricle (left ventricular hypertrophy), that is associated with increased mortality. A consistently elevated blood pressure hastens the formation of plaque or fatty deposits within the blood vessels, which causes atherosclerosis (or hardening of the arteries). Atherosclerosis increases the risk of myocardial infarction and/or stroke. The kidneys, which filter waste from

the body, are also vulnerable to damage as a result of high blood pressure. Hypertension can cause the arteries feeding the kidneys to become thickened and effectively constricted. This condition can lead to progressive kidney damage and ultimate failure. In the eye, the retinas may be damaged because of increased pressure in blood vessels in the retinal artery.

Obesity

Obesity contributes to an increased risk for many serious noncommunicable diseases, including cardiovascular disease, hypertension, stroke, non-insulin dependent diabetes mellitus (NIDDM), and various forms of cancer. Body mass index (BMI) has commonly been used as a measure of obesity and is defined as weight in kilograms divided by height in meters squared (kg/m^2). A BMI value of 30 or more should be taken as signifying obesity according to guidelines by WHO (WHO, 1998).

Heritability for BMI has been estimated to be between 50% and 90%, highlighting the importance of genetic factors to the variation in BMI (Maes, Neale, & Eaves, 1997). Environmental contributions are also of significant importance to the increasing prevalence of obesity, since the gene pool has remained stable over the same interval. High fat diets and a sedentary lifestyle are most likely explanations.

Type II Diabetes

Data have shown that approximately 150 million people suffer from diabetes mellitus worldwide, and that this number may well be doubled by the year 2025 (Alberti & Zimmet, 1998). Much of this increase will occur in developing countries and will be due to population growth, ageing, unhealthy diets, obesity and sedentary lifestyle. The term diabetes mellitus describes a metabolic disorder of multiple etiology characterized by chronic hyperglycemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action or both. Diabetes mellitus is classified into two major categories: insulin-dependent diabetes mellitus (IDDM) or Type I, and non- insulin-dependent diabetes mellitus (NIDDM) or Type II. The former includes cases that are attributable to an autoimmune process, as well as those with beta-cell destruction and who are prone to ketoacidosis. Type II diabetes encompasses cases that result from defects in insulin secretion, almost always with a major contribution from insulin resistance. Type II diabetes accounts for around 90% of all diabetes cases worldwide. It occurs most frequently in adults. Type II diabetes is strongly

familial, but some genes have only recently been consistently associated with increased risk for Type II diabetes in certain populations (Hanis et al., 1996; Horikawa et al., 2000).

Genetic and environmental influences

The importance of genetic and environmental effects in the variation of these risk factors for CVD has been examined in previous twin studies (Heller, de Faire, Pedersen, Dahlen, & McClearn, 1993; Hong, de Faire, Heller, McClearn, & Pedersen, 1994; Iselius, Morton, & Rao, 1983; Snieder et al., 2000; Snieder, van Doornen, & Boomsma, 1997, 1999). However, results in most studies are either based on small samples or are ambiguous. Further the long term importance of genetic and environmental effects over the lifespan have not been clearly assessed. Recently variance component models have also been developed to incorporate molecular data in the analysis.

The Swedish Twin Registry is the largest in the world (Lichtenstein et al., 2002) and with this enormous resource in our hands we wanted to investigate the importance of genetic and environmental effects in these risk factors by using the best available methodological approaches to date. More specifically we were interested in investigating the stability of genetic effects for cardiovascular risk factors over time; the most optimal way to detect and localize genes of importance and investigate their association to cardiovascular risk factors; and finally we were interested in best identifying environmental effects for risk factors of cardiovascular diseases.

AIMS OF THE THESIS

The aim of this thesis is to develop and implement the new statistical techniques presented in the last years, and apply them to twin data in order to study risk factors for CVD. Specific aims are:

1. To implement existing sex-limitations models and develop age models in order to explore sex and age differences in the variation of lipids and apolipoproteins (paper *I*).

2. To extend the "Fulker" model to include longitudinal data when modeling linkage and association simultaneously, and to investigate the association between blood pressure and two polymorphisms (paper *II*).

3. To apply powerful selection procedures and investigate, with the combined Haseman-Elston approach, linkage for obesity in candidate regions of chromosome 2 and chromosome 10 (paper *III*).

4. To implement the co-twin control method, controlling for genetic and shared environmental effects, in order to investigate the association between low birth weight and Type II diabetes (paper *IV*).

METHODS

Subjects

The Swedish Twin Registry

The twins used in all four papers were from the population based Swedish Twin Registry (Lichtenstein et al., 2002). The Registry comprises three cohorts (Fig.2).

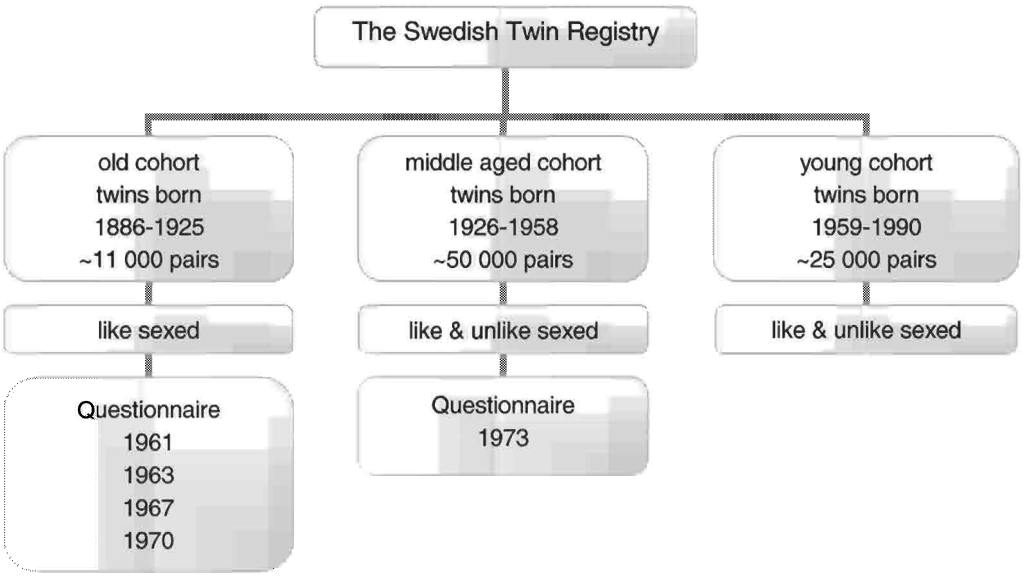


Figure 2. Structure of the Swedish Twin Registry.

Like-sexed twins born 1886-1925 constitute the 'Old Cohort'. In 1961, a questionnaire was sent out to like-sexed twin pairs, where both twins in a pair were alive and living in Sweden. Answers from approximately 11 000 pairs were collected. Questionnaire data has been collected in 1961, 1963, 1967 and in 1970 (for non responders of the 1967 questionnaire).

Like- and unlike-sexed twins born 1926-1958 constitute the 'Middle Cohort', which was compiled in 1970 by using the birth register. In 1973 a questionnaire was sent out only to like-sexed twin pairs that were both alive and living in Sweden. Responses were received from approximately 14 000 twin pairs.

The 'Young Cohort' comprised twin pairs born 1959-1990. They were also identified through the birth register. So far only the cohort born between 1985 and 1986 has been contacted.

Information gathered from questionnaires includes demographic, medical, lifestyle habits, health and psychosocial data.

Sub-studies

In the last two decades, there have been studies based on sub-samples of the Swedish Twin Registry (Fig.3).

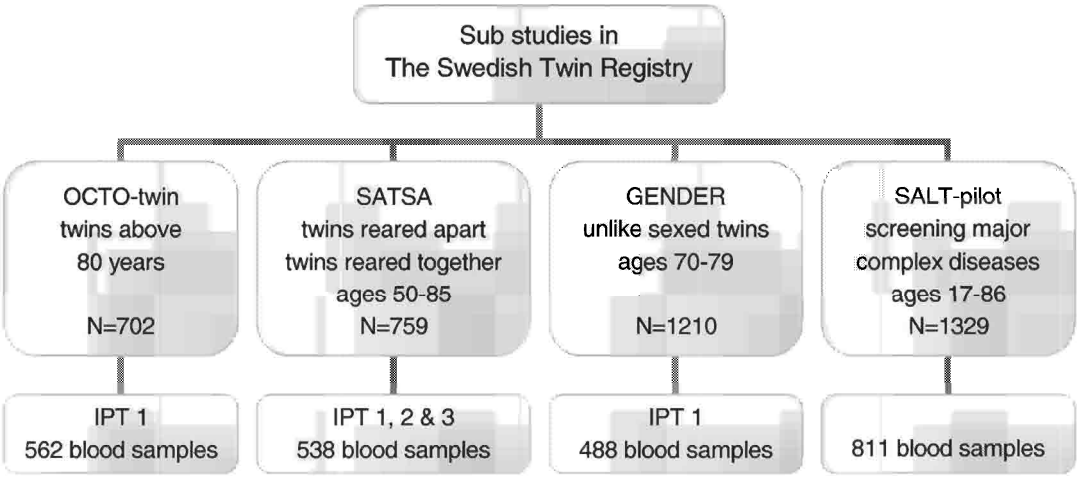


Figure 3. Sub studies in the Swedish Twin Registry with their contents and structure.

OCTO-Twin

OCTO-twin (the name originates from octogenarians) started in 1991 by collecting twin pairs from the old cohort that were alive and above the age of 80 during the three-year period of data collection (Berg et al., 1992). Twins also participated in an in person testing (IPT) including health examinations and interviews. Biological specimens (serum, plasma) were gathered from 562 individuals.

SATSA (Swedish Adoption Twin Study of Aging)

SATSA started in 1984 (Pedersen, Friberg, Floderus-Myrhed, McClearn, & Plomin, 1984; Pedersen et al., 1991). A questionnaire was initially sent out to twin pairs reared apart and a control sample of twin pairs reared together. Over 2 000 pairs have responded to repeated questionnaires sent

out every 3rd year. A sub-sample of approximately 150 pairs reared apart and approximately 150 pairs reared together participated in waves of in person testings, including health examinations and structured interviews (including functional capacity, cognitive abilities and memory). Biological specimens were available from 538 individuals from the 3 waves of testing.

GENDER

A questionnaire was sent out in 1994 to approximately 3 400 unlike-sexed twins between the ages of 70-79, mainly to examine sex differences in health and aging (Malmberg, Berg, McClearn, & Pedersen, 1995). Biological specimens were available from 488 individuals.

SALT (Screening Across the Lifespan Twin – study)

All twins born 1958 or earlier were asked to participate in a telephone interview called the Screening Across the Lifespan Twin Study (SALT) (Lichtenstein et al., 2002). It is aimed at screening all major complex diseases. All screening data are collected over the telephone by trained interviewers (with adequate medical background) using a computer based data collection system. Information was gathered on a range of symptoms and exposures. Twins that were initially contacted as part of a pilot study were also asked to contact their local nurses' office for health tests and blood sampling. Blood samples were collected from 811 individuals.

Zygosity determination

Zygosity in the SALT screening comprising twins from the old and middle cohorts of the registry is based on responses to questions regarding childhood resemblance (i.e., were you and your partner as like as 'two peas in a pod'?). This method of zygosity determination has been proven to correctly diagnose more than 95% of the twins in Sweden (Cederlöf, Friberg, Jonsson, & Kaij, 1961; Lichtenstein et al., 2002). Zygosity classification was evaluated in the SALT-pilot study (N=199 pairs), using 13 DNA markers and was proven to be correct in 99% of the pairs (Lichtenstein et al., 2002). Zygosity in the remaining sub-studies is based on serological markers.

Analytical methods

Quantitative genetic analysis

Twin studies are ideal for estimating genetic and environmental effects of traits and diseases (Martin, Boomsma, & Machin, 1997). Identical (monozygotic or MZ) twins share the same genes whereas, fraternal (dizygotic or DZ) twins share on average half of their segregating genes. A broad measure of the similarity between twins is gained from calculations of the intraclass

correlations (Neale & Cardon, 1992). Comparisons between the intraclass correlations for MZ and DZ twins provide information about the effects (genetic and/or environmental) that are present.

In general, the phenotypic variance is assumed to be due to three latent factors: additive genetic factors (A), shared environmental factors (C), and non-shared environmental factors (E, which also include measurement error):

$$Var(Y) = A + C + E$$

The correlation for MZ twin pairs is due to additive genetic and shared environmental factors (A + C). The correlation in DZ twins is assumed to be due to half the genetic, plus shared environmental factors (½A + C). A genetic effect is indicated if twin similarity is greater among MZ than DZ pairs. Heritability is defined as the proportion of total phenotypic variation directly attributable to genetic effects (Falconer, 1989).

The partitioning of phenotypic variance into genetic and environmental effects is usually illustrated in a path diagram. Genetic, shared, and non-shared environmental components are presented as latent variables (Neale & Cardon, 1992). The genetic correlation (r_g) is set to 1 in MZ twins and to 0.5 in like-sex DZ twins. The shared environment correlation (r_c) is set to 1 for both groups and is based on the equal environment assumption. By definition, there is no correlation for the non-shared environmental factors. Figure 4 illustrates a path diagram for a twin pair.

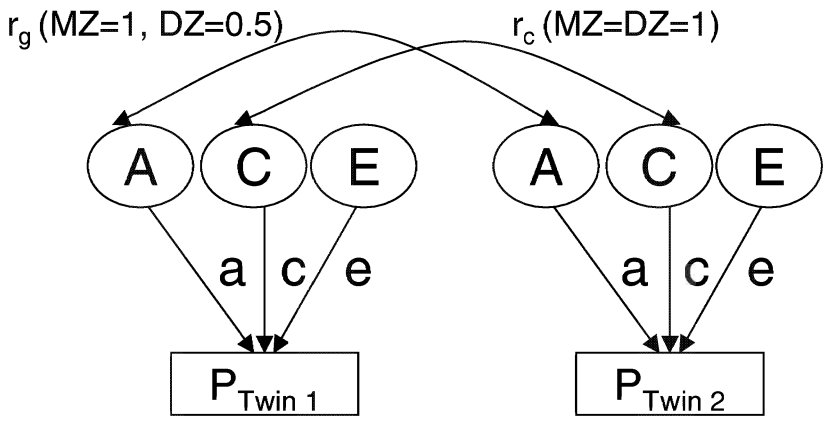


Figure 4. Path diagram for a twin pair. A, C and E are the genetic, shared and non-shared environmental effects, respectively. r_g and r_c are the genetic and shared environmental correlations, respectively.

Structural equation modeling

Mathematically, structural equation modelling combines confirmatory factor analysis and path analysis (Musil, Jones, & Warner, 1998). Confirmatory factor analysis examines causal links between observed variables and latent factors; path analysis examines the causal links among observed variables; a structural equation model examines the causal links both between observed variables and latent factors and among latent constructs themselves. The path diagram illustrates the linear structural model we use to derive predictions about the variance and covariances of the variables under study. The variance-covariance matrices modelled are of the form:

$$\begin{pmatrix} \sigma_A^2+\sigma_C^2+\sigma_E^2 & r_g\sigma_A^2+r_c\sigma_C^2 \\ r_g\sigma_A^2+r_c\sigma_C^2 & \sigma_A^2+\sigma_C^2+\sigma_E^2 \end{pmatrix}$$

where, r_g and r_c are the genetic and environmental correlations respectively and σ_A^2, σ_C^2 and σ_E^2 are the genetic, shared and non-shared environmental variance components. In all structural equation models the phenotypic means can be adjusted for various covariates such as age and sex. A maximum-likelihood estimation procedure is used, with an underlying multivariate

$$\log L = \log |2\pi\Sigma_i|^{-n/2} - \tfrac{1}{2}(x_i-\mu)'\Sigma_i^{-1}(x_i-\mu)$$

normality assumption in order to get unbiased estimates of genetic, shared and non-shared environmental effects.

Structural equation modelling is a powerful approach that also gives confidence intervals for the parameters estimated. For the models fitted, the degrees of freedom and twice the log-likelihood probability is computed by means of the structural equation-modelling package Mx (Neale, 1999b). Models are applied to raw data. To compare two models a likelihood ratio test is used. The difference between twice the log-likelihood can be interpreted as a χ^2 statistic. A significant difference indicates that the model with fewer parameters to be estimated fits data worse.

Testing for sex and age differences (Paper I)

Sex differences

A series of models can be fitted in order to test for sex differences (Neale & Martin, 1989). There are five models that are nested within each other according to figure 5.

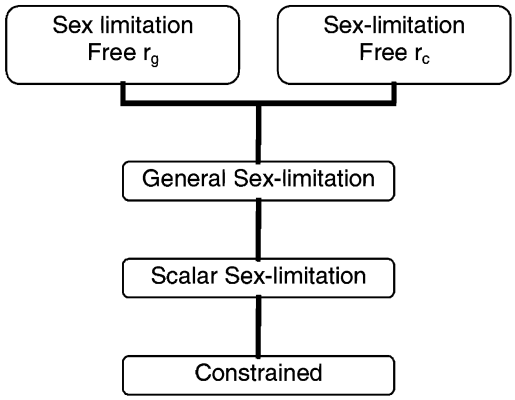


Figure 5. Nested models for testing sex differences. Hierarchy starts from the top with the models with most free parameters and going down to the more constrained models.

In the constrained model we assume equal genetic and environmental variance components for males and females. That is, there are no sex differences in the genetic and environmental influences. A scalar sex-limitation model evaluates whether sex differences in the total phenotypic variance differ only by a scalar component. That is, the genetic and environmental variance in females is a scalar multiple of that in males, $a_f + c_f + e_f = k(a_m + c_m + e_m)$, whereas, for example, heritabilities are the same for both females and males. In the general sex-limitation model there are sex differences in the relative importance of these effects by assuming one set of parameters for males in both like- and unlike-sex twin pairs and similarly another set of parameters for females. The genetic correlation (r_g) is set to 1 for MZ and 0.5 for DZ and unlike-sex twins in the first three models mentioned above. Hence, in the case of the general sex-limitation model a correlation of 0.5 implies that, although there are common genes for the genetic effects in males and females, their magnitude differs across sexes. In the two additional sex-limitation models, we test whether there are different genes or different environmental factors influencing phenotypic variation in the sexes. In the first of these two models we allow not only different variance components for males and females but also the genetic correlation between the members of the unlike-sex twin pairs to vary. For instance, if the genetic correlation is estimated at 0, it indicates that completely different genes influence the trait in

males and females. The next model is similar to the previous model, except that now we allow the shared environmental correlation (r_c) to vary. It should be noted that there is not enough information in the twin design to estimate both the genetic and environmental correlation at the same time.

Age differences

In order to test for age differences in the genetic and environmental components we fitted four models (Fig.6).

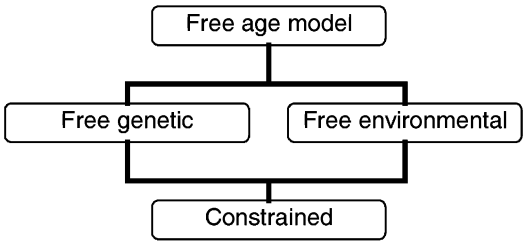


Figure 6. Nested models for testing age differences. Hierarchy starts from the top with the model with most free parameters and going down to the more constrained model.

Initially we test whether a constrained or a free age model fitted the data best in a manner analogous to that for sex differences. In the constrained age model the genetic and environmental components were set equal across the age groups. In the free model, genetic and environmental components between age groups are free to be estimated. That is, we test whether there are differences in phenotypic variance in the magnitude of genetic and environmental influences, or both, between the age groups. If the free model gives a better fit to data, two nested age models are further tested. One model assumes different genetic components across the age groups, keeping the unique environmental components stable across the age groups. In the other age model, only the unique environmental components were allowed to be free for estimation, keeping the genetic components stable across the age groups. That is, we test whether differences in phenotypic variance across the age groups are due to differences in variation in genetic or environmental influences, respectively, across the age groups.

Cholesky decomposition (Paper II)

Longitudinal or repeated data are modeled by means of the Cholesky decomposition (Neale & Cardon, 1992). Figure 7a illustrates a path diagram of the model. According to this, the first latent factor, for instance the genetic (A_1), is loading on all three phenotypes. The second latent factor (A_2) loads on the second and third phenotype. Finally, the third latent factor (A_3) loads only on the third phenotype. If several loadings from one factor are significant this indicates that the latent factor (genetic or environmental) influences the trait over time. On the other hand, if each factor only loads on one of the phenotypes, this indicates that separate (genetic or environmental) influences are operating at each time. First, we tested the significance of the genetic and environmental effects by fixing all genetic, shared environmental, or non-shared environmental loadings, separately. Then we proceeded by reducing the Cholesky model to a common factor model (Fig 7b). That is, one common factor (genetic or environmental) loads on measures at each time point and is therefore proposed to affect the trait over time. We also tested whether specific (genetic or environmental) effects were important at each time point by allowing only for separate factor loadings at each time point (Fig.7c).

Linkage and Association (Paper II)

Recently, models were presented that jointly perform tests of both linkage and association controlling for spurious associations due to population stratification and admixture (Fulker, Cherny, Sham, & Hewitt, 1999; Neale, 1999a). Testing linkage while simultaneously modeling association would provide a test of whether a Quantitative Trait Locus (QTL) is a candidate gene or whether it is in linkage disequilibrium with a QTL. Linkage is modeled in the covariance structure, while the association parameters along with other covariates are modeled in the means. The covariance matrix is of the form:

$$\begin{pmatrix} \sigma_A^2+\sigma_Q^2+\sigma_C^2+\sigma_E^2 & r_g\sigma_A^2+\hat{\pi}_i\sigma_Q^2+\sigma_C^2 \\ r_g\sigma_A^2+\hat{\pi}_i\sigma_Q^2+\sigma_C^2 & \sigma_A^2+\sigma_Q^2+\sigma_C^2+\sigma_E^2 \end{pmatrix}$$

where, σ_A^2, σ_C^2, σ_E^2 are the additive genetic, shared, and non-shared environmental variance components, σ_Q^2 is the proportion of phenotypic variance explained by the additive effects of

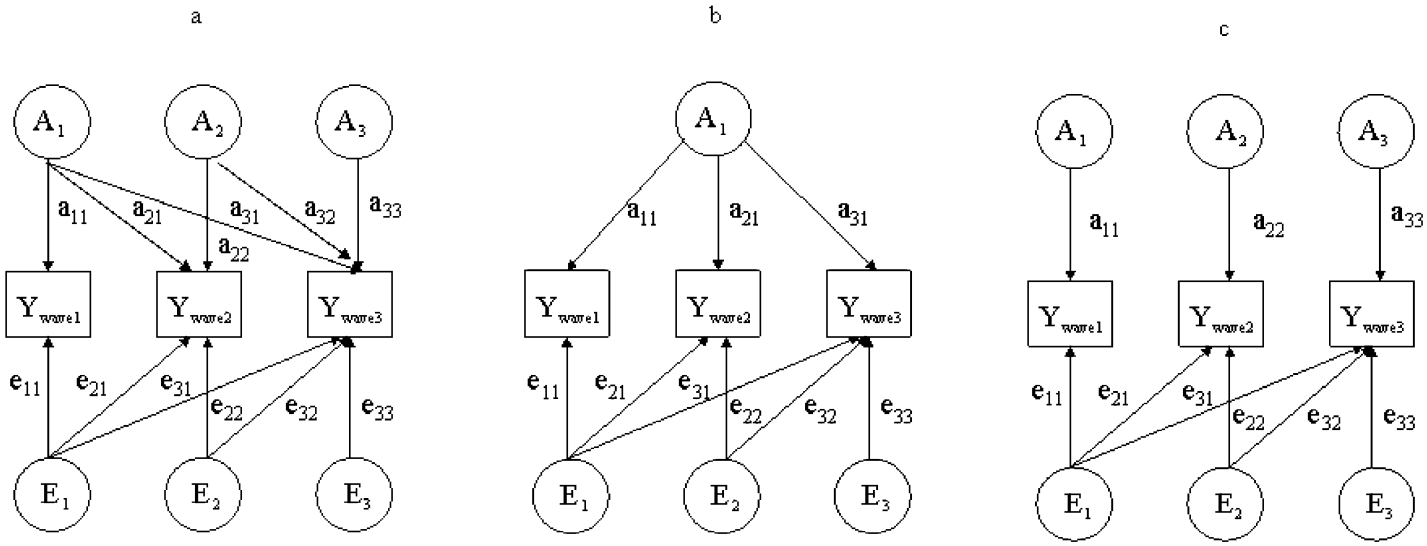


Figure 7 a, b and c. Cholesky decomposition models for three time points, wave 1, 2 and 3. A_1, A_2 and A_3 and E_1, E_2 and E_3 are the genetic and non-shared environmental components at 1st, 2nd and 3rd wave of testing, respectively. The shared environmental component (C) is omitted for simplicity. Y is the phenotype under study. Model a is the Cholesky decomposition model; model b is the common genetic factor; model c is the specific genetic factor model.

the QTL, and $\hat{\pi}_i$ is the average allele sharing at a specific chromosomal location

(Pr(*sharing* 2 *alleles IBD*)+$\frac{1}{2}$Pr(*sharing* 1 *allele IBD*)).

In this case the likelihood is the sum of three likelihoods for $\hat{\pi}_i$ =0, 0.5 and 1

($\hat{\pi}_i$ average allele sharing proportions for sharing i= 0, 1 and 2 alleles respectively). That is the

$$\log L = \log \sum_{\pi=0,0.5,1} P(IBD = \pi)\varphi_{\pi}(x)$$

estimated IBD (identity by descent) probabilities of a pair at a particular chromosomal location are used as weights. The IBD probabilities can be estimated from programs like GENEHUNTER or Merlin.

In the association part the genetic effect of a QTL is partitioned into between and within sib-pair components (Fulker et al., 1999). Consider a QTL with alleles A_1 and A_2 and respective frequencies p and q. Let the effects of the genotypes A_1A_1, A_1A_2 and A_2A_2 be a, 0 and –a. Then there are nine different genotype combinations in sib-pairs. The means and differences of the genetic effects are listed in Table 1.

Table 1. Expected sib pair means and differences, also partitioned into between and within components, for a single biallelic locus.

Genotypes		Additive effect				Partitioned effect	
Sib 1	Sib 2	Sib 1	Sib 2	Mean	Difference/2	Mean	Difference/2
A_1A_1	A_1A_1	a	a	a	0	a_b	0
A_1A_1	A_1A_2	a	0	a/2	a/2	$a_b/2$	$a_w/2$
A_1A_1	A_2A_2	a	- a	0	a	0	a_w
A_1A_2	A_1A_1	0	a	a/2	- a/2	$a_b/2$	$-a_w/2$
A_1A_2	A_1A_2	0	0	0	0	0	0
A_1A_2	A_2A_2	0	- a	- a/2	a/2	$-a_b/2$	$a_w/2$
A_2A_2	A_1A_1	- a	a	0	- a	0	$-a_w$
A_2A_2	A_1A_2	- a	0	- a/2	- a/2	$-a_b/2$	$-a_w/2$
A_2A_2	A_2A_2	- a	- a	- a	0	$-a_b$	0

In the likelihood the mean is then modelled as the overall mean plus the pair mean plus the pair difference. However, an association could occur due to population stratification, which will influence only the pair means. Fulker et al (1999) proposed to partition the genetic effect a to be different for pair means and pair differences and denoted the effects a_b and a_w. The overall mean is then modelled by means of the within and between twin-pair components.

A robust test for association may be obtained by computing the difference between a model with the within sib-pair parameter free and a model with the same parameter set to 0, while the between sib-pair parameter is free in both models. If linkage drops totally while modelling association then all resemblance among pairs is accounted for by the means model (i.e. the QTL). If there is still significant linkage but no association then the locus is not the functional QTL but rather in linkage disequilibrium with the true QTL. A model utilizing phenotypic information from longitudinal or repeated data can be used by constraining the between parameter to be equal at all three time points and similarly for the within parameter.

Linkage analysis in selected samples (Paper III)

There are various methodological approaches in doing linkage analysis. Linkage analysis of selective samples (extreme concordant and/or discordant pairs) violates important assumptions (i.e. normality) and can produce false positives or elevated type I errors when analysed. This can be overcome, either by incorporating the phenotypic values from the individuals not selected for genotyping or by incorporating the selection probabilities into the linkage calculations (Allison et al., 1999). However, assigning of prior identity-by-descent probabilities to unselected sib pairs, in covariance-structure modelling of a quantitative-trait locus has been shown to bias the estimates of genetic effects (Dolan, Boomsma, & Neale, 1999). Therefore, the main focus in modelling selected samples has been to correct for ascertainment procedures. It was shown that in the frame of the variance-components approach, conditioning on trait values leads to a likelihood ratio test that is valid and equal in power to alternative methods for analysing selected samples and hence controlling for ascertainment (Sham, Zhao, Cherny, & Hewitt, 2000). However, a simpler alternative approach to variance-components models, the "combined" Haseman-Elston approach, was introduced by Sham and Purcell (2001). According to this, the dependent variable, which is a function of the weighted sum of squared sums and squared differences, is regressed onto the estimated proportion of identity-by-descent (IBD) sharing at a test locus. The weights are the variances of the squared sums and

squared differences, respectively, and are expressed as functions of the sibling correlation. That is:

$$\frac{(X+Y)^2}{(1+r)^2}-\frac{(X-Y)^2}{(1-r)^2}+\frac{4r}{1-r^2}=\frac{4(1+r^2)}{(1-r^2)^2}Q(\hat{\pi}-.5)+\varepsilon$$

where X and Y are the bivariate normal sibling trait values with mean 0 and variance 1,and sibling correlation r, $\hat{\pi}$ is the proportion of average allele sharing and Q is the proportion of phenotypic variance explained by the additive effects of the QTL (earlier denoted as σ_Q^2). Linkage is usually presented in terms of LOD scores (logarithm of the likelihood ratio test for test of the significance of a QTL). A LOD score above 3 is considered as significant linkage (Lander & Kruglyak, 1995). An asymptotic estimate of the LOD score is given by the square of the t-value obtained from the "combined" Haseman-Elston regression divided by 2 times the natural logarithm of 10.

$$LOD \approx \frac{t^2}{2\ln(10)}$$

The t^2 is expected to be distributed as a χ^2 with 1 df, as is the maximum likelihood test statistic usually employed in linkage analysis (Fulker & Cherny, 1996). Two times the natural logarithm of 10 is a constant for converting from common (base 10) to natural logarithms.

Co-twin control (Paper IV)

The purpose of any type of matched design is to control for confounding. Studying a variable (i.e birth weight) in twins discordant for disease such as Type II diabetes allows matching for confounding factors influencing the variable under study. Any intrapair differences in that variable are indications of other factors, such as intrauterine influences, affecting the variable, after controlling for genetic and shared environmental effects. In MZ twins both genetic and early environmental effects can be controlled for, whereas in DZ twins only early environmental effects are completely controlled for. The analysis of discordant for disease twin pairs is done by conditional logistic regression implemented in the statistical package SAS (SAS/STAT, 1996).

Outliers

Phenotypic

Phenotypic outliers were identified by using estimated z-scores based on the measure of the Mahalanobis distance (Hopper & Mathews, 1982). The estimated z-score is of the form:

$$Z = \left(\left(Q / n_i \right)^{1/3} - 1 + 2/(9n_i) \right) \left(9n_i / 2 \right)^{1/2}$$

where Q is the square root of the Mahalanobis distance, $Q = (x-\mu)' \Sigma^{-1} (x-\mu)$. The z-scores have the standard normal distribution to a good approximation for the typical pedigrees we have studied (twin pairs). A Q-Q plot of the scores can be used to test the assumption that the effects predicted by the structural equation model are similar across all pedigrees.

Genotypic

Genotyping errors and marker mutations often lead to unlikely double recombination in high-resolution maps. This consequently leads to a reduced estimated level of allele sharing IBD between sib pairs, which can profoundly affect linkage information. Therefore, likely genotyping errors and marker mutations were detected by SIBMED (Douglas, Boehnke, & Lange, 2000). It is a multipoint approach based on hidden Markov models and it is designed for sib-pair data when parental genotypes are unavailable. It calculates the posterior probability of genotyping error or mutation for each sib-pair-marker combination, conditional on all marker data and an assumed genotype error rate (0.001). Hence, it removes the errors that have the largest impact on linkage results.

RESULTS

Paper I

The sample of 725 like- and unlike-sexed twin pairs gathered from three sub-studies (SATSA, GENDER and SALT-pilot) was divided into three age groups, 17-49, 50-69 and 70-85. We investigated sex and age differences in the variation of total cholesterol, apolipoprotein B and A1, and triglycerides. There were no apparent sex differences in any of the age groups for cholesterol, apolipoprotein B and A1. However, for triglycerides there were significant sex differences in the middle age group with higher heritabilities (56%) for females than males (35%). Heritabilities for lipids and apolipoproteins ranged from 45-70% in the ages 17-49, 19-74% in the ages 50-69 and 38-49% in the ages 70-85. The shared environmental effect was consistently non-significant through all the age groups for all lipids and apolipoproteins examined.

Further, there were age differences in the phenotypic variance for cholesterol, apolipoprotein B and A1 (Fig. 8). The age differences in cholesterol were due to changes in unique environmental effects. For apolipoprotein B, changes were due to changes in both genetic and unique environmental effects. Finally, for apolipoprotein A1, changes were due to changes in genetic effects. For triglyceride levels, environmental effects were larger in men than in women.

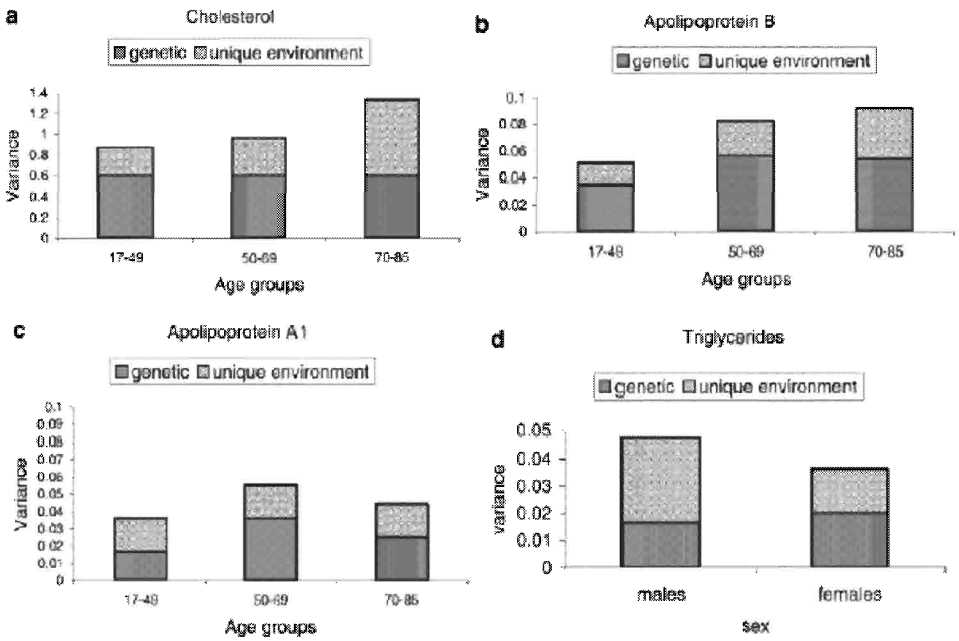


Figure 8. Genetic and unique environmental variance components for cholesterol, apolipoprotein B and A1, and triglycerides.

Paper II

Information on diastolic and systolic blood pressure measured at three time points was used from 298 like-sexed twin pairs and 86 singletons from the SATSA sample. A Cholesky decomposition model was used in order to model the repeated measures of blood pressure adjusting for age and sex. The best fitting Cholesky models are presented in Figure 9. There were no time specific genetic influences for diastolic blood pressure. Rather a common genetic factor loaded on all three time points, indicating that the same set of genes are important for blood pressure over the 6-year period of time.

The non-shared environmental loadings could not be reduced to a common factor. Most of the non-shared environmental variance in diastolic blood pressure at waves 2 and 3 comes from time specific influences, indicating that separate non-shared environmental effects are operating at each time point, with small effects loading from the first to the second (0.17) and third (0.13), and from the second to the third (0.18) (Fig. 9a). A similar pattern was seen in systolic blood pressure (Fig. 9b).

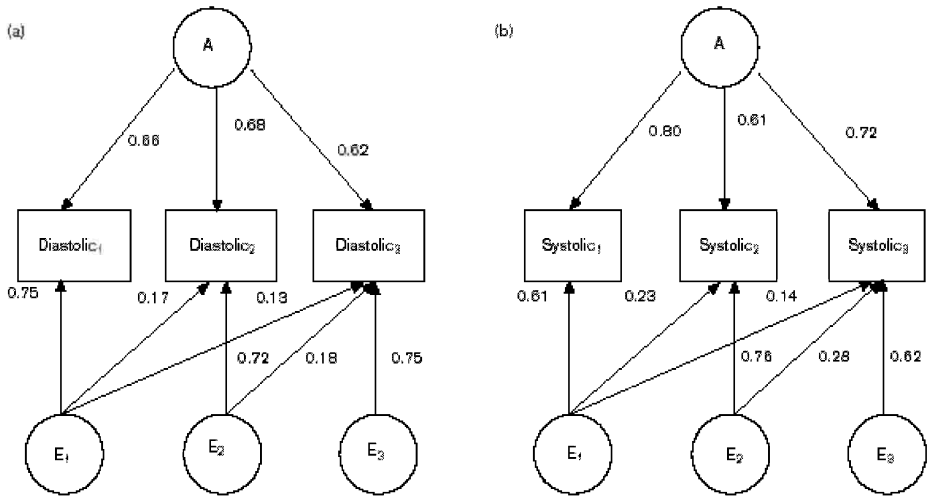


Figure 9 (a) and (b). Cholesky decomposition model for (a) diastolic and (b) systolic blood pressure at three time points. Latent factor loadings are standardized to unit variance.

We also tested for simultaneous linkage and association ("Fulker" test) between blood pressure and two polymorphisms in the renin-angiotensin-aldosterone system: the angiotensin-I converting enzyme insertion/deletion and angiotensin II type 1 receptor polymorphism (A1166C). Testing linkage while simultaneously modeling association would provide a test of whether the QTL is a candidate or whether it is in linkage disequilibrium with a trait locus. No linkage or association could be verified.

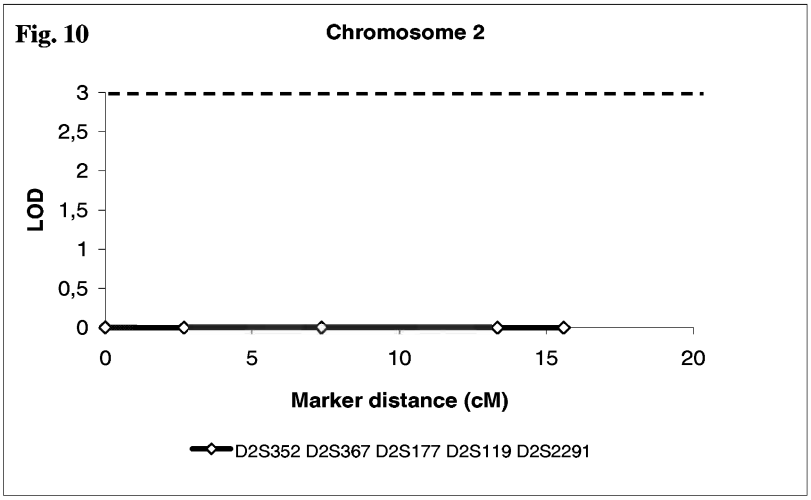
Paper III

A sample of 1422 twin pairs from four sub-studies (SATSA, OCTO-twin, GENDER and SALT-pilot) was used in order to estimate the genetic and environmental effects in the variation of BMI in three age groups: 17-49, 50-69 and 70-94 years. Heritability ranged from 59-70%, implying that genetic effects were of importance for the variation of obesity. However, in the middle age group (50-69 years) significant genetic effects in the variation of BMI only appeared among women (Table 2).

Table 2. Genetic (A), shared (C) and non-shared (E) environmental estimates of variance for BMI (with 95% confidence intervals) derived from the best fitting model in each age group.

	Men			Women		
	A	C	E	A	C	E
17≤ Age <50	0.70	0	0.30	same as men		
	.55-.77	0-.10	.23-.41			
50≤ Age <70	0.18	0.47	0.35	0.68	0.01	0.31
	0-.58	.12-.68	.22-.51	.42-.82	0-.13	.18-.54
Age ≥70	0.59	0	0.41	same as men		
	.46-.67	0-.08	.33-.51			

The results of the linkage analysis with the "combined" Haseman-Elston approach, using genotypic data from 51 extreme concordant and 155 discordant twin pairs are shown in figures 10 and 11 for chromosome 2 and 10, respectively. They are plots of asymptotic LOD scores at each marker. Five markers were selected to cover a 16 cM candidate region on chromosome 2 with an average marker density of 3.9 cM. Eight markers were selected for the 14 cM candidate region on chromosome 10 with an average marker density of 2 cM. The sibling correlation was estimated from the whole sample of twins to 0.35. There were no significant LOD scores, indicating that linkage could not be verified on chromosomes 2 and 10.



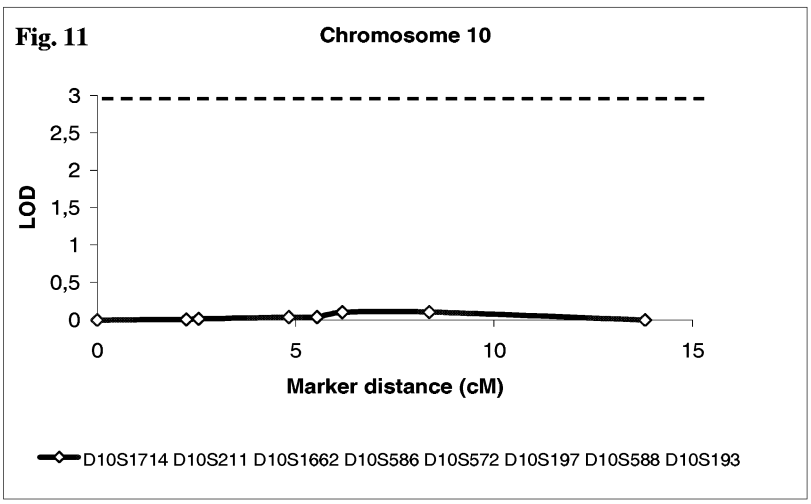


Figure 10 and 11. Plot of asymptotic LOD scores by each marker for chromosome 2 and 10, respectively.

Paper IV

In a cohort of 11 226 same-sexed Swedish twins born before 1958 and included in the SALT study, we found that individuals with a birth weight $\leq$ 2.0 kg faced an almost doubled risk (adjusted OR=1.7; 95% c.i. 1.4-2.2) of developing Type-II diabetes compared to individuals with a birth weight between 2.0 and 3.0 kg (Table 3). The increased risk of Type-II diabetes remained in the intrapair comparison among 142 twin pairs discordant for Type-II diabetes (adjusted OR=1.6; 95% c.i. 0.7-3.5), in which unmeasured genetic and shared environmental effects were accounted for (Table 3). The risk was of the same magnitude for both MZ (OR=1.9; 95% c.i. 0.7-5.3) and for DZ (OR=1.7; 95% c.i. 0.7-4.1) twins.

Table 3. Adjusted odds ratios (with 95% confidence intervals) for the association between birth weight and Type-II diabetes in the cohort and in the co-twin control analysis of Swedish twin pairs.

Birth weight	Cohort analysis	Co-twin control analysis
≤ 2.0 kg	1.7 (1.4-2.2)	1.6 (0.7-3.5)
2.0-3.0 kg	reference	reference
≥ 3.0 kg	0.9 (0.6-1.1)	1.0 (0.4-2.6)
Walds test		
p-value	< 0.0001	0.4

* Adjusted for age, sex, BMI, smoking status and zygosity.

** Adjusted for BMI and smoking status

DISCUSSION

Methodological issues

Twin method

Some of the assumptions of the basic (ACE) twin model is: (i) random mating, (ii) additive genetic effects, (iii) no gene-environment correlation or interaction, and (iv) the equal environment assumption (Plomin, DeFries, McClearn, & McGuffin, 2001).

Non random mating

Non random or assortative mating could bias the estimates of genetic effects downwards. When it occurs, similarity in DZ twins will be increased and hence the difference between MZ and DZ correlations will be smaller, leading to a smaller estimate of heritability. However, the effects of assortative mating could be modeled if appropriate parental information is gathered. Assortative mating is plausible for some traits in the domain of education, religion, attitudes and socioeconomic status and was not considered as an issue in this thesis.

Non additive effects

Quantitative genetic theory assumes that alleles at a locus and across loci "add up" to affect a trait. In the basic twin model we are interested in the average effect of an allele at a locus which is attributed to the additive genetic effects. Nevertheless, the effects can be non-additive if alleles at a locus interact (dominant) or alleles at different loci interact (epistatic). Offspring receive only one allele from each parent and not a combination of alleles at a locus. Therefore, non additive genetic influence will not be transmitted from one generation to another. Models can include both additive and dominance effects independent of each other. However, there is not enough information in the models to differentiate dominance from shared environmental effects; hence the basic twin model only includes additive genetic effects. In our studies we tested for dominance and found no significant effects. Therefore it is highly unlikely that the effects of nonadditivity affected the results of our studies.

Gene-environment correlation and interaction

Gene-environment correlation occurs when certain genes are associated with certain environments. Gene-environment interaction occurs when there is an interaction between genetic effects and environment, and refers more to the genetic susceptibility to environments. The bias generated from omitting gene-environment correlations or interactions is generally hard

to characterize. Although models can be extended to incorporate more complex effects such as interactions, it is not possible to include all of them at the same time (Purcell, 2002). In order to use these models, a priori knowledge of existing interactions is needed and measures of specific environmental effects are often essential. Further, the detection of interaction effects in analysis of variance requires far more subjects than the detection of main effects. The focus of our studies has not been to examine interactions. It should be noted that interaction effects have been found in some studies along with main genetic and environmental effects. However, they explained a small proportion of the variation in traits, indicating that main effects of genes and environments are reasonable assumptions.

Equal environment assumption

One of the most fundamental assumptions in twin studies is the equal environment assumption. It assumes that if MZ twins are treated more similarly than DZ twins it is only due to the MZ's genetic similarity. That is, environmentally caused similarity is roughly the same for both MZ and DZ twins. The equal environment assumption has been tested by examining behavioral similarity in twins with mislabeled zygosity compared to twins with correct zygosity and it was shown that mislabeling had an insignificant effect in the twin study method (Scarr & Carter-Saltzman, 1979). Therefore, it is unlikely that the equal environment assumption has been violated in our studies.

Linkage and Association: the Fulker test

The method presented by Fulker et al (1999) involves simultaneous testing of linkage and association. The linkage test is based on differences in covariances according to the IBD status at a candidate locus. The association test is based on differences in means given the genotypes at the locus. The model for the means is partitioned into between- and within-pairs components and the association test is therefore robust to population stratification and admixture. The power of the "Fulker" test has been shown to be dependent on the amount of variance explained by the QTL, the marker information and the degree of linkage disequilibrium (Sham, Cherny, Purcell, & Hewitt, 2000). The power of linkage is related to the square of the QTL heritability and it attenuates with incomplete marker information or when analysis is performed at a locus that is linked to the QTL rather than being the actual QTL itself. The power of association is related to the QTL heritability and it attenuates with incomplete linkage disequilibrium. However, it was also shown that both tests had increasing power with increasing residual shared variance, which could be accomplished by use of repeated or multivariate data; hence the use of repeated blood pressure measurements in paper *II*. Nevertheless, marker informativeness, measured by observed

heterozygosity, was relatively low (0.50 for the ACE I/D polymorphism and 0.43 for the AT_1R-A1166C polymorphism) in our study. The markers used could also have been in linkage disequilibrium with the actual QTL, which could have reduced the power to find linkage and association. Therefore we could not exclude linkage and association between the polymorphism and blood pressure. Other non-methodological issues related to the power of the study are discussed below.

Linkage in selected samples

Selective sampling has been suggested for mapping quantitative trait loci as a more cost-effective and powerful approach compared to random sampling (Eaves & Meyer, 1994). That is, sib pairs are selected on the basis of their phenotypic distribution. Risch and Zhang (1995) suggested choosing the top and bottom 10^{th} percentile or the top 10^{th} and bottom 30^{th} percentile of the phenotypic distribution as cut off points for the selection procedure (Fig. 12). Initially it was shown that extremely discordant sib pairs are the most powerful study subjects for linkage analysis (Risch & Zhang, 1995; Risch & Zhang, 1996). Further, the importance of extreme concordant sib pairs was also examined (Zhang & Risch, 1996). It was shown that combining extreme discordant and extreme concordant sib pairs appears to provide a more powerful approach to linkage (Dolan & Boomsma, 1998; Gu, Todorov, & Rao, 1996; Gu, Todorov, & Rao, 1997). However, its application is strongly dependent on the genetic model specified, such as the heritability of the locus, the density and informativeness of the markers.

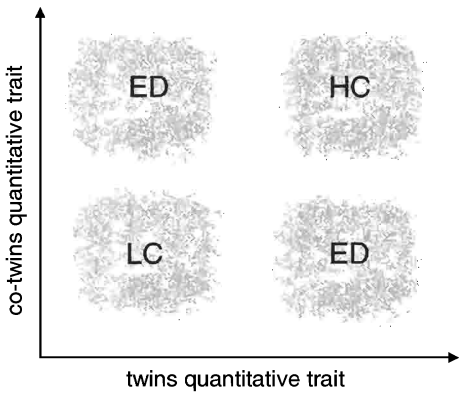


Figure 12. Distribution of extreme discordant and concordant pairs.
ED: extreme discordant
HC: high concordant,
LC: low concordant.

In addition, Gu et al. (1996) suggested that using extreme concordant pairs from both tails of the distribution may not be as powerful as using only pairs from the upper tail (high concordant pairs) especially when the underlying genetic model is unclear (Gu et al., 1996). In paper *III* we used extreme discordant twin pairs and only high concordant twin pairs, using the 10^{th} percentile

as a cut off (a BMI of 29 kg/m^2). The reason for this approach is that, the underlying genetic mechanism for obesity is not known and also, that low concordant sib pairs might have a different genetic origin compared to obese individuals. The use of DZ twins in linkage studies is also advantageous since they are age matched and the occurrence of non-paternity is sufficiently rare that it can be considered unlikely.

However, the Risch and Zhang (1995) percentile selection approach poses a few problems. For instance, a pair where one sibling has a BMI of 22 (normal value) and the other a BMI of 40 (extreme value) will not be included in the selected sample according to the percentile approach, because only one sibling exceeds the cut off point. However, this pair will be included and rated as discordant with the "sliding window rule", a more flexible approach that allows a window of difference to slide through a distribution of difference scores. Nevertheless, this approach does not yield equivalent power to Risch and Zhang's percentile based approach even though it allows a greater proportion of sibling pairs to be eligible (Allison, 1996). In our study (paper *III*), we used a combination of both approaches. Thus, the selected sample of twins contained twins that were either above the top 10th percentile or below the bottom 30th percentile, but were also rated appropriately as concordant or discordant with the "sliding window rule" (i.e. a twin with a BMI of 29 and the co-twin with a BMI of 40, both above the top 10th percentile, were rated discordant). Recently, a more optimal method for sibling selection has been presented by introducing a quantitative index of informativeness based on the sibship's contribution to the non-centrality parameter (Purcell, Cherny, Hewitt, & Sham, 2001). However, the twin pairs included in paper *III*, were already selected according to the best available method at the time, highlighting the rapid methodological development in the area of quantitative trait mapping.

Variance-components techniques have emerged as the most powerful of available methods for quantitative trait loci mapping (Allison et al., 1999; Blangero, Williams, & Almasy, 2001; Goring, Williams, & Blangero, 2001). However, standard variance components linkage analysis can produce elevated type I errors, when applied to selected samples and non-normal data. Although initially suggested, it has become increasingly clear that imputing the expected IBD probabilities for the unselected part of the sample can introduce severe bias in linkage analysis (Dolan et al., 1999). Sham et al. (2000) suggested adjusting the log-likelihood function by conditioning on the trait values as a mean of ascertainment correction. At the same time the combined Haseman-Elston approach was shown to be equivalent to the variance components analysis conditioning on trait values in selected samples. The advantage with the combined Haseman-Elston method is its simplicity in use; it is a regression analysis, which can be applied with most statistical

programs. Compared to the variance components approach it is also a more intuitive and straightforward approach.

Co-twin control

The aim in co-twin control study is to control for genetic and shared environmental effects by choosing disease discordant twins (Spector, Snieder, & MacGregor, 2000). That is, the trait under study is under genetic control for MZ pairs and partly genetic control for DZ pairs, while both are controlled for shared environmental effects. The ideal situation is to present results for each zygosity separately in order to draw conclusions. Although data for paper *IV* were taken from a large telephone interview screening of Swedish twins still alive and born between 1886 and 1958, we had limited power to conduct a co-twin control analysis, which was evident from the large confidence intervals. Analysis was done on the merged sample of MZ and DZ discordant for Type II diabetes twin pairs. We had approximately 30% power to find a risk of 2 for the association of low birth weight and Type II diabetes. Power calculations showed that a sample of approximately 500 pairs would be needed to reach power levels of 80% at 5% significance level (Dupont & Plummer, 1998).

Phenotypic issues

Lipids and Apolipoproteins - Paper I

The most important finding in studying age and sex differences in the variation of lipids and apolipoproteins was the higher variance in the older compared to younger age groups. In general, heritabilities did not deviate much from those reported in previous twin and family studies of total cholesterol, apolipoprotein B and A1, and triglyceride levels (Heller et al., 1993; Iselius, 1979; Snieder et al., 1997; Whitfield & Martin, 1983).

Twin and family studies have also suggested an increased phenotypic variance in lipids and apolipoproteins with increasing age (Beekman et al., 2002; Boomsma et al., 1996; Ericsson et al., 1991; Heller et al., 1993; Reilly, Kottke, & Sing, 1990; Snieder et al., 1997, 1999). We found that phenotypic variance increased across the age groups for cholesterol due to an increase in non-shared environmental variance. Similar increase in non-shared environmental variance was also found for apolipoprotein B above the age of 50. It could be argued that this is entirely due to measurement error, which is included in the non-shared environmental component, since data are derived from different sub-studies. However, biological assessments were done by the same laboratory for all three sub-studies. It is more likely that the results could be the consequence of accumulated environmental experiences over the life course (Nelson & Dannefer, 1992). Exposure

to environmental variations such as smoking, diet and alcohol, may alter lipid metabolism. The age effect could also reflect a weakening of the homeostatic control mechanisms of the lipid system with aging (Reilly et al., 1990; Snieder et al., 1997). For apolipoprotein A1, differences in phenotypic variance were due to differences in genetic variance components between age groups, probably reflecting that different combinations of multiple genes influence lipids and apolipoproteins in different periods of life. This is also supported in a study by Snieder et al (1997), who assessed differential gene expression as a function of age by studying young-adult twins, their middle-aged parents, and a second group of twins of the same age as the parents. Phenotypic variance was lower in the older age group for apolipoprotein A1 compared to the middle age group. This decrease could also reflect cohort differences, as individuals with a genetic predisposition to lipid disorders may die at earlier ages and therefore not be present in the oldest age group (less variability). However, the effects of random variation could not be excluded.

We also found that heritability in the variation of triglyceride levels was higher in women than in men between the ages of 50 and 69. Above the age of 50, major changes take place in production of sex hormones in women, which may cause lipid changes. The effects of hormones on lipid metabolism were examined in pre- and postmenopausal women, and significant changes were found for apolipoprotein B and LDL cholesterol (Schaefer et al., 1994). The shared environmental component of variance was consistently zero for all lipids and apolipoproteins, in accordance with other twin studies, indicating that the shared environment does not play a major role in the variation of lipids and apolipoproteins (Beekman et al., 2002; Boomsma et al., 1996; Snieder et al., 1997, 1999).

One limitation was the nature of our study, comprising a combination of cross-sectional studies, which precluded conclusions about changes over time due to genetic and environmental influences. This limitation highlights the importance of longitudinal approaches in studying the variation of lipid and apolipoproteins throughout the lifespan.

Diastolic and Systolic Blood Pressure - Paper II

This study quantified genetic and environmental sources of variance in repeated measures of diastolic and systolic blood pressure across a 6-year period of time in twins reared together and apart. It showed a genetic stability with the genetic component of variance explaining up to 46% of the total phenotypic variance in diastolic and up to 63% in systolic blood pressure. This may imply that the underlying genetic mechanisms of blood pressure regulation do not change

appreciably during this 6-year period of time. Similar results of genetic stability above the age of 50 have been shown from other twin studies (Colletto, Cardon, & Fulker, 1993; Vinck, Fagard, Loos, & Vlietinck, 2001). Previous studies have also shown that shared environmental effects on the variation of blood pressure are negligible or of minor importance. The importance of shared rearing environment can be estimated from twins reared together and apart. We found from the intraclass correlations that rearing effects could be of importance for diastolic but not necessarily for systolic blood pressure. However, due to small sample sizes in each category, analyses preceded with the merged sample of twins reared together and apart. The non-shared environmental components were characterised by time specific influences with small transmissions from one time point to another, suggesting non-shared environmental influences having a small long-term effect. They may reflect the importance of lifestyle habits such as smoking. However, they could also reflect measurement error, which is also included in the non-shared environmental part.

We found no linkage or association between blood pressure and two known polymorphisms in genes comprising the renin-angiotensin-aldosterone system. The angiotensin-I-converting enzyme insertion/deletion polymorphism has shown a lack of an association with blood pressure in most European countries (Bengtsson et al., 1999; Berge & Berg, 1994, 1998; O'Donnell et al., 1998; O'Malley, Maslen, & Illingworth, 1999). In contrast, African and Jamaican populations have shown an association (Barley et al., 1996; Forrester et al., 1997; Zhu et al., 2001). This indicates that a different mix of genes is important for blood pressure regulation in different ethnic groups. The angiotensin-II type 1 receptor (A1166C) polymorphism has shown an association to hypertension in previous studies (Bonnardeaux et al., 1994). However, the lack of an association between blood pressure and the A1166C polymorphism in our study could be due to the fact that the sample comprised of normotensive subjects. Other Nordic studies have not shown an association either, expect for a Finnish study that found an association in subjects with early onset of hypertension and normal body weight (Kainulainen et al., 1999). However, the sample was selected from a uniform geographical region, and could therefore have a different genetic background compared to other Nordic countries.

There were a few limitations in this study. Variance in blood pressure at the third time point was consistently higher compared to the variance in the first and second time point, which could imply poorer fit in the models tested. This could be due to the fact that nurses changed to electronic devices at the third time point, which yielded more imprecise readings compared to the ordinary manual procedures for blood pressure measurements used in previous testings.

Automated blood pressure devices do not produce accurate blood pressure values and one reason could be that they were designed for self-measurement and not clinical use (Beevers, Lip, & O'Brien, 2001a, 2001b). Individuals with blood pressure medication were omitted from the study. Considering antihypertensive treatment as a covariate for adjustment may create potential bias, because of the correlation between treatment and the genes affecting blood pressure variation. By omitting individuals on antihypertensive medication we may have excluded the subjects with the highest genetic predisposition, and hence reduced the chances of finding an association. The optimal way to handle this would be to adjust blood pressure measurements with the average level of the increase or decrease incurred by medication. However, it would require information exactly before and after medication which was not available in our study. The lack of an association could also be due to small sample size, resulting in low power. Also the mode of inheritance for the polymorphisms is not known. We used the basic model with additive genetic effects. In case of dominance or recessiveness or even interactions and epistasis far more subjects would be needed (Sham, Cherny et al., 2000).

Obesity - Paper III

Heritability estimates for BMI in the present study ranged between 59-70%, which is similar to previous reports (Maes et al., 1997). The apparent effects of sex and age in the variation of BMI have also been reported in previous studies (Korkeila, Kaprio, Rissanen, & Koskenvuo, 1991; Neale & Cardon, 1992). In the present study, the heritability of BMI was significantly higher for women compared to men in the middle age group (50-70 years). Above the age of 50 major changes take place in the production of sex hormones in women. The shared environmental effect in the middle age group was higher among men than among women. This could reflect the accumulation of environmental experiences and changes in lifestyle habits, such as smoking or diet that could be more important in men than in women. Subsequently, one would expect to see a larger shared environmental effect in the older age group. However, in our opinion, a more likely explanation is that this represents a chance finding. Shared environmental effects are rarely found to be a major contributor to the variation of BMI (Maes et al., 1997). Therefore we believe that the shared environmental effect is probably much less than our observed (0.47). Sex differences are not persistent through all age groups, but heritability was lower in the older age group compared to the younger. This may imply that individuals with a genetic predisposition to obesity have a lower survival probability.

In contrast to results from previous studies (Comuzzie et al., 1997; Hager et al., 1998; Hinney et al., 2000; Price et al., 2001), we could not verify linkage in candidate regions on chromosome 2 and

10. The effect of the loci could be different in different samples. The findings on chromosome 2 and 10 were found in extreme obese samples rather than moderately obese (Hager et al., 1998), indicating that genetic effects may be more apparent in extreme obese compared to moderately obese humans. Previous linkage studies on chromosome 2 have shown linkage with leptin levels and not BMI, indicating that genetic mechanisms could act differently in separate obesity related phenotypes (Comuzzie et al., 1997). A genome-scan aimed at identifying QTLs that have a potential influence on BMI found only weak linkage signals on chromosome 2, indicating that the locus may have a smaller effect in some samples compared to others (Feitosa et al., 2002). Cultural and ethnic differences could also have limited our chances of verifying those findings. In Sweden, for example, the prevalence of obesity is less than half that of the United States (WHO 1998). The apparent lack of concordance between studies could also reflect the fact that genetic determinants of inter-individual variation in obesity and related phenotypes are likely to be multiple and interacting, with most single variants producing only a moderate effect, which are hard to detect with linkage.

One limitation of our study was the limited power to detect linkage. In a series of simulation studies on linkage and replication of linkage results, it was shown that 'detectability' was improved as the number of QTLs decreased (Suarez, Hampe, & van Eerdewegh, 1994). The results also suggested that replication of a true linkage claim in an oligogenic disorder, requires a much larger sample. In contrast a false linkage claim is unlikely to replicate. The mode of inheritance for obesity is unknown and power is related to the genetic parameters specified. We conducted our own simulation studies using the same amount of concordant and discordant twins as were available in the study. We altered parameters such as the increaser allele frequency, sibling correlation and the amount of variance explained by the putative QTL and found that power ranged from 10% to 90% depending on the specified parameters.

Type II diabetes and Low Birth Weight - Paper IV

In this study we found an association between low birth weight and Type-II diabetes in a population-based cohort of 11 226 same-sexed Swedish twins (Lichtenstein et al., 2002). The analysis support the 'thrifty phenotype' hypothesis, which proposes that impaired glucose intolerance and other metabolic disorders originate through adaptations, in the malnourished fetus. This theory further proposes that, as long as poor nourishment continues during childhood and adult life, these adaptations are beneficial. However, a sedentary lifestyle, changes in food intake (high calorie food), and the subsequent development of obesity ultimately lead to impaired glucose intolerance, the insulin resistance syndrome and Type-II diabetes (Hales &

Barker, 2001). In the current study, the rate of Type-II diabetes was higher among obese individuals, suggesting this as a possible explanation.

In the within twin pair analysis, the association between low birth weight and Type-II-diabetes in disease discordant pairs remained even after we controlled for genetic and shared environmental effects. The reason for the difference between this study on Type-II-diabetes and studies on other cardiovascular outcomes, where familial effects seem to have been responsible for the association with low birth weight (Christensen, Stovring, & McGue, 2001; Hubinette et al., 2001; Hubinette, Cnattingius, Johansson, Henriksson, & Lichtenstein, 2003; Ijzerman, Stehouwer, & Boomsma, 2000), are not known. We can only speculate that the association between reduced fetal growth and Type-II-diabetes is of a different origin than the association between fetal growth and other cardiovascular outcomes. Animal experiments have shown that growth restricted offspring of protein malnourished rats undergo a greater age-dependent glucose intolerance, which is associated to insulin resistance (Ozanne, 2001). The association with other cardiovascular outcomes, on the other hand, might be due to a common genetic etiology. For instance, both birth weight (Clausson, Lichtenstein, & Cnattingius, 2000; Magnus, Gjessing, Skrondal, & Skjaerven, 2001) and hypertension (Lehtovirta et al., 2000) are genetically influenced. It has been suggested that the genetic influences in preeclampsia could be the common cause, since preeclampsia is associated both with low birth weight in the offspring and later hypertension in the mother (North, Simmons, Barnfather, & Upjohn, 1996; Salonen Ros, Lichtenstein, Lipworth, & Cnattingius, 2000).

One limitation of the current study is that both birth weight and diabetes were self-reported. However, the validity of self-reported birth weight in this cohort of twins has been examined and correlated reasonably well with birth weight derived from medical birth records (r=0.82). The agreement between questionnaire data and medical records has been shown to be good for well-known chronic diseases, such as diabetes and cardiovascular diseases in a study of middle-aged and elderly Finnish men and women (Haapanen, Miilunpalo, Pasanen, Oja, & Vuori, 1997). Another limitation was the small number of disease discordant pairs in the intrapair analyses. However, the interpretation that there was an association also within twin pairs was supported by the fact that the risks of Type-II diabetes related to low birth weight in both MZ and DZ were of similar magnitude to the risks found in the cohort.

Generalisability of twin studies

Twins are on average 0.9 kg lighter than singletons at birth and are delivered 2-3 weeks earlier (Leon, 2001; Spector et al., 2000). In spite their appreciably impaired growth in utero, they undergo a rapid catch-up growth in childhood. In the context of the fetal origin hypothesis, it is interesting to compare the association of low birth weight with cardiovascular diseases between twins and singletons and whether they occur more often in twins. Two thirds of monozygotic twin pairs have a common chorion and placenta and therefore compete for nutrients. Dizygotic twin pairs always have separate chorions and placentas may or may not be fused. Therefore, monozygotic twins may experience a more adverse intrauterine environment than dizygotic twins. However, there is no evidence that twins have increased cardiovascular mortality (Vagero & Leon, 1994) or overall mortality (Christensen, Wienke et al., 2001) compared to the general population. Hence these findings support the hypothesis that results from twin studies can be generalised to the population.

Future directions

With the beginning of the new millennium we have also entered a new era in genetic epidemiology. The completion of the human genome has offered an abundance of valuable information and we are just catching up on the methodological side. "*Big is beautiful*" will not only be used as a fashion statement, but must also be adopted by scientists. Small genetic effects or interactions with other genes and/or environments in which they are expressed, will only be accomplished through powerful methodological techniques and large sample sizes.

The power of our molecular studies has been relatively low. Not only were we haunted by the small sample sizes but also of the limited knowledge regarding the possible genetic mechanisms involved. The elucidation of such mechanisms has now approached reachable levels with the access of results from the human genome project. For traits where genetic stability is present, as was the case in blood pressure and some lipids (cholesterol and Apolipoprotein B), gene localization may be easier compared to traits where their genetic effects may be expressed differently in different ages (such as in apolipoprotein A1). An approach to deal with statistical and epidemiological issues is reflected in an integrated project called GENOMEUTWIN, that combines genetic and epidemiological data from the Scandinavian twin registries, the Netherlands, and Italy to form a promising scientific resource (Boomsma, Busjahn, & Peltonen, 2002).

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all colleagues and friends who have contributed to this thesis. In particular, I would like to thank:

Paul Lichtenstein, my principal tutor, mentor and very good friend. Your guidance, support, understanding, kindness and friendship, through my ups and downs, made this thesis possible. You always made me feel like an "equal" although my journey has been an educative one in research and life. Thank you is not enough to express my gratitude and appreciation!

Ulf de Faire, my co-tutor for guidance and support, and enthusiasm for my skills. Your help and advice has meant a lot to me.

Nancy Pedersen, my other co-tutor for welcoming me to the Swedish Twin Registry and making my PhD studies a possibility. Thank you for support and guidance.

Sven Cnattingius, a co-author and enthusiastic researcher. I admire your kindness, straight forwardness and sharpness in research. Thank you for all the help you have given me.

Nick Martin, a co-author and enthusiastic researcher who made part of this thesis a reality, by "housing" me in his department and giving me abundant help, from small things to bigger issues (access to the library, scientific help: analytical, epidemiological). It has always been a pleasure to work with you and your tremendous team. Many thanks to: David Duffy, Kathy Kirk, Andrew Birley, Sue Treloar, Gu Zhu, Grant Montgomery, Nick McGregor, Veronica Sellin, Coral Pink, David Smyth, Olivia Zheng and the rest of the Genetic Epidemiology department. My friends and colleagues at the department: David Evans, Kate Morley, Anna Marsh, Sarah Medland, Christine Smit, Michelle Luciano, Narelle Hansell, Juli Condon, Claire Sullivan, Belinda Castellano, and the rest of you guys!

Professor Hans-Olov Adami for being such a tremendous leader of the Department of Medical Epidemiology and Biostatistics.

I would also like to thank everyone at MEB for making it a wonderful working place (we are way too many to mention you all).

Everyone from the "twin" group: Kerstin Linderholm, Pia Svedberg, Måns Flensburg, Karin Dellenvall, Mårten Jansson, Chayna Davis, Erica Spotts, Lorelei Mucci, Birgitta Andersson, Carola Bardage, Mia Wallin, Kecia Watari, Margy Gatz and all others in the group.

Especially my parents Eleni and Nikolaos, who fought all their lives to make sure their children got educated, never had to suffer and get exploited in "dirty factories"! I owe you what I have achieved and what I have become.

My dearest brother Simon, who is always there for me and supports me in every possible way. I could not have done this without you. "Σ΄αγαπω πολυ γιαβρι μου "!

To my big Greek family: grandmothers, aunts and uncles and beloved cousins, who always make sure my trips to Greece, are well remembered and lived to the fullest.

To my other family Liisa, Anders and Fabian Fredmark, who adopted me into their family and gave me the extra pleasure of becoming a godmother!

To all of my friends who have stood by me one way or another and enriched my life in many ways: Nina Waldenström, Kim Ikonen, Catherine Tuvblad, Maria Katsafiloudi, Henrik Larsson, Emma Nilsson, Natasa Simi, Anna, Mary and Vicki, Nelli, Anna Mikaelsson, Anna Gatzoflia, Hossein, Nabil, Ulla Forssén, Vasso, Sofia Papathoma, Renee and Bengt with family, Roseanna, Susane Berghauer, Cia and Tanja, and Tyron Ridgway.

I would also like to acknowledge the funding sources that supported my research and all the twins in the Swedish Twin Registry.

REFERENCES

Alberti, K. G., & Zimmet, P. Z. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med, 15*(7), 539-553.

Allison, D. B. (1996). The use of discordant sibling pairs for finding genetic loci linked to obesity: practical considerations. *Int J Obes Relat Metab Disord, 20*(6), 553-560.

Allison, D. B., Neale, M. C., Zannolli, R., Schork, N. J., Amos, C. I., & Blangero, J. (1999). Testing the robustness of the likelihood-ratio test in a variance- component quantitative-trait loci-mapping procedure. *Am J Hum Genet, 65*(2), 531-544.

American Heart Association, h. w. a. o. (2003). *Diseases And Conditions*. Available: http://www.americanheart.org/.

Austin, M. A., Hokanson, J. E., & Edwards, K. L. (1998). Hypertriglyceridemia as a cardiovascular risk factor. *Am J Cardiol, 81*(4A), 7B-12B.

Barley, J., Blackwood, A., Miller, M., Markandu, N. D., Carter, N. D., Jeffery, S., Cappuccio, F. P., MacGregor, G. A., & Sagnella, G. A. (1996). Angiotensin converting enzyme gene I/D polymorphism, blood pressure and the renin-angiotensin system in Caucasian and Afro-Caribbean peoples. *J Hum Hypertens, 10*(1), 31-35.

Beekman, M., Heijmans, B. T., Martin, N. G., Pedersen, N. L., Whitfield, J. B., DeFaire, U., van Baal, G. C., Snieder, H., Vogler, G. P., Slagboom, P. E., & Boomsma, D. I. (2002). Heritabilities of apolipoprotein and lipid levels in three countries. *Twin Res, 5*(2), 87-97.

Beevers, G., Lip, G. Y., & O'Brien, E. (2001a). ABC of hypertension. Blood pressure measurement. Part I-sphygmomanometry: factors common to all techniques. *Bmj, 322*(7292), 981-985.

Beevers, G., Lip, G. Y., & O'Brien, E. (2001b). ABC of hypertension: Blood pressure measurement. Part II-conventional sphygmomanometry: technique of auscultatory blood pressure measurement. *Bmj, 322*(7293), 1043-1047.

Bengtsson, K., Orho-Melander, M., Lindblad, U., Melander, O., Bog-Hansen, E., Ranstam, J., Rastam, L., & Groop, L. (1999). Polymorphism in the angiotensin converting enzyme but not in the angiotensinogen gene is associated with hypertension and type 2 diabetes: the Skaraborg Hypertension and diabetes project. *J Hypertens, 17*(11), 1569-1575.

Berg, S., Johansson, B., Plomin, R., Ahern, F. M., Pedersen, N. L., & McClearn, G. E. (1992). Origins of variance in the old-old: The first presentation of the OCTO-Twin Study in Sweden. *Behavior Genetics*(22), 708-709.

Berge, K. E., & Berg, K. (1994). No effect of insertion/deletion polymorphism at the ACE locus on normal blood pressure level or variability. *Clin Genet, 45*(4), 169-174.

Berge, K. E., & Berg, K. (1998). Polymorphisms at the angiotensinogen (AGT) and angiotensin II type 1 receptor (AT1R) loci and normal blood pressure. *Clin Genet, 53*(3), 214-219.

Blangero, J., Williams, J. T., & Almasy, L. (2001). Variance component methods for detecting complex trait loci. *Adv Genet, 42*, 151-181.

Bonnardeaux, A., Davies, E., Jeunemaitre, X., Fery, I., Charru, A., Clauser, E., Tiret, L., Cambien, F., Corvol, P., & Soubrier, F. (1994). Angiotensin II type 1 receptor gene polymorphisms in human essential hypertension. *Hypertension, 24*(1), 63-69.

Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nat Rev Genet, 3*(11), 872-882.

Boomsma, D. I., Kempen, H. J., Gevers Leuven, J. A., Havekes, L., de Knijff, P., & Frants, R. R. (1996). Genetic analysis of sex and generation differences in plasma lipid, lipoprotein, and apolipoprotein levels in adolescent twins and their parents. *Genet Epidemiol, 13*(1), 49-60.

Cederlöf, R., Friberg, L., Jonsson, E., & Kaij, L. (1961). Studies on similarity diagnosis in twins with the aid of mailed questionnaires. *Acta Geneticae Medicae et Gemellologiae (Roma), 11*, 338-362.

Christensen, K., Stovring, H., & McGue, M. (2001). Do genetic factors contribute to the association between birth weight and blood pressure? *J Epidemiol Community Health, 55*(8), 583-587.

Christensen, K., Wienke, A., Skytthe, A., Holm, N. V., Vaupel, J. W., & Yashin, A. I. (2001). Cardiovascular mortality in twins and the fetal origins hypothesis. *Twin Res, 4*(5), 344-349.

Clausson, B., Lichtenstein, P., & Cnattingius, S. (2000). Genetic influence on birthweight and gestational length determined by studies in offspring of twins. *Bjog, 107*(3), 375-381.

Colletto, G. M., Cardon, L. R., & Fulker, D. W. (1993). A genetic and environmental time series analysis of blood pressure in male twins. *Genet Epidemiol, 10*(6), 533-538.

Comuzzie, A. G., Hixson, J. E., Almasy, L., Mitchell, B. D., Mahaney, M. C., Dyer, T. D., Stern, M. P., MacCluer, J. W., & Blangero, J. (1997). A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2. *Nat Genet, 15*(3), 273-276.

Dolan, C. V., & Boomsma, D. I. (1998). Optimal selection of sib pairs from random samples for linkage analysis of a QTL using the EDAC test. *Behav Genet, 28*(3), 197-206.

Dolan, C. V., Boomsma, D. I., & Neale, M. C. (1999). A simulation study of the effects of assignment of prior identity-by- descent probabilities to unselected sib pairs, in covariance-structure modeling of a quantitative-trait locus. *Am J Hum Genet, 64*(1), 268-280.

Douglas, J. A., Boehnke, M., & Lange, K. (2000). A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am J Hum Genet, 66*(4), 1287-1297.

Dupont, W. D., & Plummer, W. D. (1998). Power and Sample Size Calculations for Studies Involving Linear Regression. *Controlled Clinical Trials*(19), 589-601.

Eaves, L., & Meyer, J. (1994). Locating human quantitative trait loci: guidelines for the selection of sibling pairs for genotyping. *Behav Genet, 24*(5), 443-455.

Ericsson, S., Eriksson, M., Vitols, S., Einarsson, K., Berglund, L., & Angelin, B. (1991). Influence of age on the metabolism of plasma low density lipoproteins in healthy males. *J Clin Invest, 87*(2), 591-596.

European Societies on coronary prevention. (1998). Prevention of coronary heart disease in clinical practice. Recommendations of the Second Joint Task Force of European and other Societies on coronary prevention. *Eur Heart J, 19*(10), 1434-1503.

Falconer, D. S. (1989). *Introduction to quantitative genetics.* (3rd ed.). Essex. England: Longman Scientific and Technical.

Feitosa, M. F., Borecki, I. B., Rich, S. S., Arnett, D. K., Sholinsky, P., Myers, R. H., Leppert, M., & Province, M. A. (2002). Quantitative-trait loci influencing body-mass index

reside on chromosomes 7 and 13: the National Heart, Lung, and Blood Institute Family Heart Study. *Am J Hum Genet, 70*(1), 72-82.

Forrester, T., McFarlane-Anderson, N., Bennett, F. I., Wilks, R., Cooper, R., Rotimi, C., Morrison, L., & Ward, R. (1997). The angiotensin converting enzyme and blood pressure in Jamaicans. *Am J Hypertens, 10*(5 Pt 1), 519-524.

Fulker, D. W., & Cherny, S. S. (1996). An improved multipoint sib-pair analysis of quantitative traits. *Behav Genet, 26*(5), 527-532.

Fulker, D. W., Cherny, S. S., Sham, P. C., & Hewitt, J. K. (1999). Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet, 64*(1), 259-267.

Goldbourt, U., de Faire, U., & Berg, K. (1994). *Genetic factors in coronary heart disease.* Dordrecht, Netherlands: Kluwer.

Goring, H. H., Williams, J. T., & Blangero, J. (2001). Linkage analysis of quantitative traits in randomly ascertained pedigrees: comparison of penetrance-based and variance component analysis. *Genet Epidemiol, 21*(Suppl 1), S783-788.

Gu, C., Todorov, A., & Rao, D. C. (1996). Combining extremely concordant sibpairs with extremely discordant sibpairs provides a cost effective way to linkage analysis of quantitative trait loci. *Genet Epidemiol, 13*(6), 513-533.

Gu, C., Todorov, A. A., & Rao, D. C. (1997). Genome screening using extremely discordant and extremely concordant sib pairs. *Genet Epidemiol, 14*(6), 791-796.

Haapanen, N., Miilunpalo, S., Pasanen, M., Oja, P., & Vuori, I. (1997). Agreement between questionnaire data and medical records of chronic diseases in middle-aged and elderly Finnish men and women. *Am J Epidemiol, 145*(8), 762-769.

Hager, J., Dina, C., Francke, S., Dubois, S., Houari, M., Vatin, V., Vaillant, E., Lorentz, N., Basdevant, A., Clement, K., Guy-Grand, B., & Froguel, P. (1998). A genome-wide scan for human obesity genes reveals a major susceptibility locus on chromosome 10. *Nat Genet, 20*(3), 304-308.

Hales, C. N., & Barker, D. J. (2001). The thrifty phenotype hypothesis. *Br Med Bull, 60*, 5-20.

Hanis, C. L., Boerwinkle, E., Chakraborty, R., Ellsworth, D. L., Concannon, P., Stirling, B., Morrison, V. A., Wapelhorst, B., Spielman, R. S., Gogolin-Ewens, K. J., Shepard, J. M., Williams, S. R., Risch, N., Hinds, D., Iwasaki, N., Ogata, M., Omori, Y., Petzold, C., Rietzch, H., Schroder, H. E., Schulze, J., Cox, N. J., Menzel, S., Boriraj, V. V., Chen, X., & et al. (1996). A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat Genet, 13*(2), 161-166.

Heller, D. A., de Faire, U., Pedersen, N. L., Dahlen, G., & McClearn, G. E. (1993). Genetic and environmental influences on serum lipid levels in twins. *N Engl J Med, 328*(16), 1150-1156.

Hinney, A., Ziegler, A., Oeffner, F., Wedewardt, C., Vogel, M., Wulftange, H., Geller, F., Stubing, K., Siegfried, W., Goldschmidt, H. P., Remschmidt, H., & Hebebrand, J. (2000). Independent confirmation of a major locus for obesity on chromosome 10. *J Clin Endocrinol Metab, 85*(8), 2962-2965.

Hong, Y., de Faire, U., Heller, D. A., McClearn, G. E., & Pedersen, N. (1994). Genetic and environmental influences on blood pressure in elderly twins. *Hypertension, 24*(6), 663-670.

Hopper, J. L., & Mathews, J. D. (1982). Extensions to multivariate normal models for pedigree analysis. *Ann Hum Genet, 46*(4), 373-383.

Horikawa, Y., Oda, N., Cox, N. J., Li, X., Orho-Melander, M., Hara, M., Hinokio, Y., Lindner, T. H., Mashima, H., Schwarz, P. E., del Bosque-Plata, L., Oda, Y., Yoshiuchi, I., Colilla, S., Polonsky, K. S., Wei, S., Concannon, P., Iwasaki, N., Schulze, J., Baier, L. J., Bogardus, C., Groop, L., Boerwinkle, E., Hanis, C. L., & Bell, G. I. (2000). Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat Genet, 26*(2), 163-175.

Hubinette, A., Cnattingius, S., Ekbom, A., de Faire, U., Kramer, M., & Lichtenstein, P. (2001). Birthweight, early environment, and genetics: a study of twins discordant for acute myocardial infarction. *Lancet, 357*(9273), 1997-2001.

Hubinette, A., Cnattingius, S., Johansson, A., Henriksson, C., & Lichtenstein, P. (2003). Birthweight and risk of Angina Pectoris: Analysis in Swedish Twins. *Eur J Epidemiol (In Press).*

Ijzerman, R., Stehouwer, C. D., & Boomsma, D. I. (2000). Evidence for genetic factors explaining the birth weight-blood pressure relation. Analysis in twins. *Hypertension, 36*(6), 1008-1012.

Iselius, L. (1979). Analysis of family resemblance for lipids and lipoproteins. *Clin Genet, 15*(4), 300-306.

Iselius, L., Morton, N. E., & Rao, D. C. (1983). Family resemblance for blood pressure. *Hum Hered, 33*(5), 277-286.

Kainulainen, K., Perola, M., Terwilliger, J., Kaprio, J., Koskenvuo, M., Syvanen, A. C., Vartiainen, E., Peltonen, L., & Kontula, K. (1999). Evidence for involvement of the type 1 angiotensin II receptor locus in essential hypertension. *Hypertension, 33*(3), 844-849.

Korkeila, M., Kaprio, J., Rissanen, A., & Koskenvuo, M. (1991). Effects of gender and age on the heritability of body mass index. *Int J Obes, 15*(10), 647-654.

Lander, E., & Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet, 11*(3), 241-247.

Lehtovirta, M., Kaprio, J., Forsblom, C., Eriksson, J., Tuomilehto, J., & Groop, L. (2000). Insulin sensitivity and insulin secretion in monozygotic and dizygotic twins. *Diabetologia, 43*(3), 285-293.

Leon, D. A. (2001). The foetal origins of adult disease: interpreting the evidence from twin studies. *Twin Res, 4*(5), 321-326.

Lichtenstein, P., De Faire, U., Floderus, B., Svartengren, M., Svedberg, P., & Pedersen, N. L. (2002). The Swedish Twin Registry: a unique resource for clinical, epidemiological and genetic studies. *J Intern Med, 252*(3), 184-205.

Maes, H. H., Neale, M. C., & Eaves, L. J. (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behav Genet, 27*(4), 325-351.

Magnus, P., Gjessing, H. K., Skrondal, A., & Skjaerven, R. (2001). Paternal contribution to birth weight. *J Epidemiol Community Health, 55*(12), 873-877.

Malmberg, B., Berg, S., McClearn, G. E., & Pedersen, N. L. (1995). *A twin register of unlike sexed elderly Swedes.* Paper presented at the 8th International Congress of Twin Studies, Richmond VA.

Martin, N., Boomsma, D., & Machin, G. (1997). A twin-pronged attack on complex traits. *Nat Genet, 17*(4), 387-392.

Musil, C. M., Jones, S. L., & Warner, C. D. (1998). Structural equation modeling and its relationship to multiple regression and factor analysis. *Res Nurs Health, 21*(3), 271-281.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families.*: Dordrecht: Kluwer Academic Publications.

Neale, M. C., & Martin, N. G. (1989). The effects of age, sex, and genotype on self-report drunkenness following a challenge dose of alcohol. *Behav Genet, 19*(1), 63-78.

Neale, M. C., S.S. Cherny, P.C. Sham, J.B. Whitfield, A.C. Heath, A.J. Birley and N.G. Martin. (1999a). Distinguishing Population Stratification from Genuine Allelic Effects with Mx: Association of ADH2 with Alcohol Consumption. *Behav Genet, 29*(4), 233-243.

Neale, M. C. (1999b). Mx: Statistical modeling. (Version 5th). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Nelson, E. A., & Dannefer, D. (1992). Aged heterogeneity: fact or fiction? The fate of diversity in gerontological research. *Gerontologist, 32*(1), 17-23.

Nora, J. J., Berg, K., & Nora, A. H. (1991). *Cardiovascular Diseases, Genetics, Epidemiology and Prevention*. New York: Oxford University Press.

North, R. A., Simmons, D., Barnfather, D., & Upjohn, M. (1996). What happens to women with preeclampsia? Microalbuminuria and hypertension following preeclampsia. *Aust N Z J Obstet Gynaecol, 36*(3), 233-238.

O'Donnell, C. J., Lindpaintner, K., Larson, M. G., Rao, V. S., Ordovas, J. M., Schaefer, E. J., Myers, R. H., & Levy, D. (1998). Evidence for association and genetic linkage of the angiotensin-converting enzyme locus with hypertension and blood pressure in men but not women in the Framingham Heart Study. *Circulation, 97*(18), 1766-1772.

O'Malley, J. P., Maslen, C. L., & Illingworth, D. R. (1999). Angiotensin-converting enzyme and cardiovascular disease risk. *Curr Opin Lipidol, 10*(5), 407-415.

Ozanne, S. E. (2001). Metabolic programming in animals. *Br Med Bull, 60*, 143-152.

Pedersen, N. L., Friberg, L., Floderus-Myrhed, B., McClearn, G. E., & Plomin, R. (1984). Swedish early separated twins: identification and characterization. *Acta Genet Med Gemellol, 33*(2), 243-250.

Pedersen, N. L., McClearn, G. E., Plomin, R., Nesselroade, J. R., Berg, S., & DeFaire, U. (1991). The Swedish Adoption Twin Study of Aging: an update. *Acta Genet Med Gemellol, 40*(1), 7-20.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001). *Behavioral Genetics* (Fourth ed.). United States of America: Worth Publishers.

Price, R. A., Li, W. D., Bernstein, A., Crystal, A., Golding, E. M., Weisberg, S. J., & Zuckerman, W. A. (2001). A locus affecting obesity in human chromosome region 10p12. *Diabetologia, 44*(3), 363-366.

Purcell, S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin Res, 5*(6), 554-571.

Purcell, S., Cherny, S. S., Hewitt, J. K., & Sham, P. C. (2001). Optimal sibship selection for genotyping in quantitative trait locus linkage analysis. *Hum Hered, 52*(1), 1-13.

Rader, D., & Brewer, B. J. (1994). Lipids, apolipoproteins and lipoproteins. In U. Goldbourt & U. de Faire & K. Berg (Eds.), *Genetic factors in Coronary Heart Disease*.

Reilly, S. L., Kottke, B. A., & Sing, C. F. (1990). The effects of generation and gender on the joint distributions of lipid and apolipoprotein phenotypes in the population at large. *J Clin Epidemiol, 43*(9), 921-940.

Risch, N., & Zhang, H. (1995). Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science, 268*(5217), 1584-1589.

Risch, N. J., & Zhang, H. (1996). Mapping quantitative trait loci with extreme discordant sib pairs: sampling considerations. *Am J Hum Genet, 58*(4), 836-843.

Salonen Ros, H., Lichtenstein, P., Lipworth, L., & Cnattingius, S. (2000). Genetic effects on the liability of developing pre-eclampsia and gestational hypertension. *American Journal of Medical Genetics, 91*(4), 256-260.

SAS/STAT, S. (1996). Changes and Enhancements through Release 6.11. Cary, NC: SAS Institute Inc.

Scarr, S., & Carter-Saltzman, L. (1979). Twin method: defense of a critical assumption. *Behav Genet, 9*(6), 527-542.

Schaefer, E. J., Lamon-Fava, S., Cohn, S. D., Schaefer, M. M., Ordovas, J. M., Castelli, W. P., & Wilson, P. W. (1994). Effects of age, gender, and menopausal status on plasma low density lipoprotein cholesterol and apolipoprotein B levels in the Framingham Offspring Study. *J Lipid Res, 35*(5), 779-792.

Schulte, P. A., Rothman, N., & Austin, M. A. (1993). Cardiovascular Disease. In P. A. Schulte & P. F.P. (Eds.), *Molecula Epidemiology : Principles and Practises* (pp. 341-363). San Diego, California: Academic Press, Inc.

Sham, P. C., Cherny, S. S., Purcell, S., & Hewitt, J. K. (2000). Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data. *Am J Hum Genet, 66*(5), 1616-1630.

Sham, P. C., Zhao, J. H., Cherny, S. S., & Hewitt, J. K. (2000). Variance-Components QTL linkage analysis of selected and non-normal samples: conditioning on trait values. *Genet Epidemiol, 19*(Suppl 1), S22-28.

Snieder, H., Hayward, C. S., Perks, U., Kelly, R. P., Kelly, P. J., & Spector, T. D. (2000). Heritability of central systolic pressure augmentation: a twin study. *Hypertension, 35*(2), 574-579.

Snieder, H., van Doornen, L. J., & Boomsma, D. I. (1997). The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am J Hum Genet, 60*(3), 638-650.

Snieder, H., van Doornen, L. J., & Boomsma, D. I. (1999). Dissecting the genetic architecture of lipids, lipoproteins, and apolipoproteins: lessons from twin studies. *Arterioscler Thromb Vasc Biol, 19*(12), 2826-2834.

Spector, T. D., Snieder, H., & MacGregor, A. (2000). *Advances in Twin and Sib-pair Analysis*: Oxford University Press.

Suarez, B. K., Hampe, C. L., & van Eerdewegh, P. (1994). *Genetic Approaches to Mental Disorders*: American Psychiatric Press, Inc.

The National Cholesterol Education Program. (2001). Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *Jama, 285*(19), 2486-2497.

Vagero, D., & Leon, D. (1994). Ischaemic heart disease and low birth weight: a test of the fetal-origins hypothesis from the Swedish Twin Registry. *Lancet, 343*(8892), 260-263.

Whitfield, J. B., & Martin, N. G. (1983). Plasma lipids in twins. Environmental and genetic influences. *Atherosclerosis, 48*(3), 265-277.

WHO. (1998). *World Health Organization, Obesity: Preventing and Managing the Global Epidemic*. Geneva.

Vinck, W. J., Fagard, R. H., Loos, R., & Vlietinck, R. (2001). The impact of genetic and environmental influences on blood pressure variance across age-groups. *J Hypertens, 19*(6), 1007-1013.

Zhang, H., & Risch, N. (1996). Mapping quantitative-trait loci in humans by use of extreme concordant sib pairs: selected sampling by parental phenotypes [published erratum appears in Am J Hum Genet 1997 Mar;60(3):748-9]. *Am J Hum Genet, 59*(4), 951-957.

Zhu, X., Bouzekri, N., Southam, L., Cooper, R. S., Adeyemo, A., McKenzie, C. A., Luke, A., Chen, G., Elston, R. C., & Ward, R. (2001). Linkage and association analysis of angiotensin i-converting enzyme (ace)-gene polymorphisms with ace concentration and blood pressure. *Am J Hum Genet, 68*(5), 1139-1148.