Twins and the genetic architecture of osteoarthritis

Clustering of disease in families is either the result of shared genetic influences (‘nature’) or the shared family environment (‘nurture’). By studying twins—specifically by comparing the occurrence of disease in monozygotic (MZ) and dizygotic (DZ) twins—the contribution of these two influences can be readily separated. Expressed in the terminology of quantitative genetics, twin data allow the variation in the propensity (or liability) to disease in a population to be attributed to: (i) genetic variance in the population; (ii) variance that might be due to environmental influences that are shared by twins (such as age, year of birth and the family environment); (iii) random environmental variation attributable to factors that are unique to each twin [1]. The relative contribution of genetic variance to the liability to disease is referred to as ‘heritability’ [2]. Identifying and measuring heritability is the traditional contribution of the twin study. Over the last century, the classical twin study has changed clinicians’ perceptions of many diseases, including many of the common rheumatic diseases, formerly considered as having a wholly environmental basis.

The contribution of the twin study design, however, extends beyond measuring theoretical concepts of genetic and environmental variation. Data from twins can lead directly to the identification of genes. Twin data are of particular value in providing a sound model for identifying the influence of genes of small relative effect against a background of environmental variation [3]. The twin study design, through the exact background genetic matching of MZ twins, can also uniquely assess specific types of gene–environment interaction [1]. The common rheumatic diseases do not show clear Mendelian patterns of inheritance and factors in the external environment influence disease occurrence and expression. Data from twins will, therefore, be particularly useful in uncovering their precise genetic basis. In this editorial, as an example, we focus on the contribution of the twin design to the genetic epidemiology of osteoarthritis (OA) and its future potential for defining the genetic architecture of this complex disease.

The first formal study of the familial clustering of Heberden’s nodes was conducted by Stecher in the 1940s [4]. The mothers and sisters of 64 female Caucasian probands with Heberden’s nodes were shown, respectively, to have a 2- and 3-fold risk of developing nodes when compared with the general population. This work was confirmed and extended in epidemiological surveys in the UK by Kellgren et al. in the 1950s [5, 6]. In families of probands with generalized OA involving six or more joint groups, a 2-fold excess of OA was found among first-degree relatives compared with population controls. The recurrence risk was highest in the relatives of female probands, at older ages, and in the relatives of the subgroup of probands with Heberden’s nodes.

The familial clustering of hand OA alone and in combination with knee OA has recently been confirmed in contemporary data from two large population cohorts in the USA. Data from the Baltimore Longitudinal Study of Aging relating to radiographs obtained between 1978 and 1991 in 167 families have confirmed a significant sibling–sibling intraclass correlation for OA at the DIP, PIP and first carpometacarpal joints, and for ‘polyarticular’ OA [7]. The latter was defined as the involvement of two of three hand joints plus at least one knee. No significant clustering was observed for knee OA alone, although the number of cases was small and patellofemoral OA was not recorded. The Framingham Offspring Study has also reported the results of radiographic studies of hand and knee OA from 337 nuclear families in 313 extended pedigrees [8]. The study showed sibling–sibling and parent–offspring correlations in ‘OA count’ (a combined measure of the number of hand and knee joints affected) ranging from 0.12 to 0.31. The correlations were higher among pairs that included females than in male–male pairs alone. Unlike the earlier studies which identified affected relatives on the basis of having one affected proband, both these population studies ascertained all subjects irrespective of disease status. This approach potentially introduces less bias in estimating familial aggregation. Both studies also adjusted for possible confounding variables including body mass index, sex and age. In the Framingham Offspring Study, the longitudinal design allowed both parents and offspring to be assessed at adult ages.

There are no adequate population data on the familial occurrence of OA at other peripheral joint sites and in the spine. Familial clustering of hip OA has, however, been suggested in two studies which have examined the prevalence of OA among relatives of cases who had undergone total hip replacement surgery. Lindberg [9] showed that the frequency of radiological OA among siblings of 184 probands was increased 2-fold over the general population. Chitnavis et al. [10] have also reported a slight but non-significant increase of a recalled diagnosis of hip OA among relatives of probands with total hip replacement compared with spouse controls. A higher degree of clustering was also demonstrated for knee OA in the Chitnavis study. In the spine, a number of small studies have reported familial clustering for scoliosis [11, 12], spondylolisthesis [13, 14], sciatica [15], cervical spondylosis [16] and herniated discs in both adolescents [17, 18] and adults [19].
While family studies point to a genetic contribution to OA, without information from twins it is difficult to exclude the possibility that familial clustering is the result of the shared environmental factors within families. Contributory environmental factors, such as patterns of exercise, that may influence OA risk are recognized to cluster in families and are difficult to control for adequately in epidemiological studies [8]. An additional difficulty in studying parent–offspring and sibling–sibling correlation is in accounting for differences in age-related disease expression; this issue is readily addressed by the natural age matching of twin samples.

The first systematic twin studies were conducted in the 1960s [6, 20, 21]. Although numbers were small and the results were not published in full, a genetic influence on both Heberden’s nodes and generalized OA (defined as involvement of three or more joint groups) was suggested by greater concordance for disease in MZ than DZ twins. A number of case series of twins have also suggested a genetic influence on herniated discs in juveniles [22], hyperostosis of the spinal ligaments [23] and Schmorl’s nodes [24, 25]. Interpreting these data is difficult because of the small sample sizes involved and the fact that a number studied only MZ twins. Inflated estimates of the genetic contribution may also have been obtained through over-ascertainment of concordant MZ pairs in case series and among volunteer populations with disease [26–28].

Adequate twin data in OA have emerged only relatively recently. In 1996, we reported the first results of studies in OA in the St Thomas’ UK Adult Twin Registry [29], a group of healthy female–female twin pairs recruited from the UK population by media campaigns and from an existing twin register. One of the strengths of the data of the St Thomas’ cohort is that none of the twins was aware that they were the subject of an investigation of OA before volunteering to take part in the study. Further, disease definition was not based on symptoms, but on radiological grounds.

An analysis of hand and knee radiographs obtained from 130 MZ and 120 DZ twin pairs aged between 44 and 77 yr, and assessed using standard atlas definitions, demonstrated a consistently higher concordance for OA among MZ than among DZ twins in all joint areas [30]. A genetic effect was found for both joint space narrowing and osteophytes. These results translate into estimates of the heritability of OA at the hand and knee of between 39 and 65%, and indicate that a substantial part of the variance in liability to OA in the population can be explained by genetic variation. Analysis of the data on radiological OA at the hip in the St Thomas’ cohort has required substantially larger twin numbers due to its lower prevalence and has only recently been completed. These data also indicate a significant genetic influence at that site with a heritability of ~50% [31].

Magnetic resonance images of the lumbar and cervical spine have also been evaluated for features of disc degeneration in a group of 87 MZ and 78 DZ twin pairs from the St Thomas’ cohort in combination with twins ascertained from the Australian NHMRC Twin Registry [32]. Greater similarity in MZ compared with DZ twins was observed for a score comprising the sum of disc space narrowing, disc bulge, osteophytes and disc signal change at both sites. The results corresponded to a heritability of 74% in the lumbar spine and 73% in the cervical spine. These findings confirmed earlier observations by Bätte et al. [33] in a study of lumbar MRI scans in MZ twins alone which showed a marked similarity in appearance that could not be accounted for by environmental variation.

Frequently cited criticisms of the twin method include the concern that extracting genetic variance from twin data requires the assumption that the shared environments of MZ and DZ twins are similar (the ‘equal environments’ assumption) [34]. Should environmental determinants of disease be shared more commonly among MZ than DZ twins, violation of this assumption may lead to inflated estimates of heritability. The observation that a number of adult diseases show an association with intrauterine growth and development has added weight to this criticism, given the greater sharing of the intrauterine environment in MZ when compared to DZ twins [35]. An additional concern is that twins may be unrepresentative of the population with disease.

While it is impossible to exclude fully a biased effect on heritability resulting from unequal environmental sharing in MZ and DZ twins, it is likely that this influence is negligible in OA. The strength of association required for an environmental variable to act as a confounder for a genetic association is large [36]. Multivariate path analysis in our own data, accounting for differences in the distribution of known confounding variables (including obesity, social class and smoking) among MZ and DZ twins, showed these variables to have no important influence on heritability [30, 32]. It is highly unlikely that differential distribution of unknown confounders or the effect of fetal growth could have a sufficiently strong association with OA to influence these results. With respect to the representativeness of twins, the frequency of OA among the St Thomas’ twins was comparable to that observed in an age-matched population female sample [37]. This is in keeping with twin data for other diseases suggesting that morbidity and mortality among adult twins are comparable to the general population.

Twin studies in OA confirm that genetic factors are important in determining the pattern of occurrence of disease in the population. Furthermore, the magnitude of the genetic influence detected in twin studies, with heritabilities for disease as high as 74% for lumbar disc degeneration and 65% for hand and knee OA, indicate that it is feasible to design studies that would detect specific genes that have a quantitative influence on the risk of disease. The possibility that individual genes may have a substantial influence in the occurrence of OA is plausible. Segregation analysis of population data in OA has consistently suggested the influence of a major gene [5, 6, 38, 39]. This has been explored most fully in
the Framingham Offspring Study where analysis has suggested that the distribution of disease could be best explained by a single Mendelian recessive gene in combination with polygenic effects [8]. Specific genes have been implicated in the pathogenesis of disease in animal models of both peripheral joint disease and disc degeneration [40, 41]. In humans, OA of the peripheral joints and the spine is a feature of several chondrodysplasias for which single-gene defects have been identified [42, 43]. Rare examples of familial OA also exist that have been linked to an autosomal dominant mutation in type II collagen [44–46].

Contemporary methods used to implicate genes and genetic regions with disease involve either tracking the segregation of genetic markers with disease in families or testing for associations with diseases in a case–control setting using either population or family-based controls. These methods are widely applicable to the rheumatic diseases and have been the subject of a recent review in this journal [47]. Genetic data from siblings with OA are invaluable to both these strategies. Linkage analysis using affected sibling pairs, for example, is particularly appropriate for complex diseases such as OA where the mode of inheritance is unknown. Studying allelic association in siblings (either supplemented with data on parental genotype to assess transmission distortion [48] or alone [49]) is increasingly recognized as an important strategy for eliminating the confounding effects of population stratification which may lead to inconsistent results in association studies using population controls.

The special contribution of twins to these analytical approaches is not at first obvious. DZ twins share only the genetic likeness of ordinary siblings. The genetic identity of MZ twins means they cannot contribute to standard linkage analysis. In association studies, MZ twins appear to offer no advantage over singletons, other than the cost benefit of genotyping only one member of each pair. This, however, overlooks many of the characteristics of twins that are of importance for conditions such as OA where the expression of disease is a function of age and modified by the environment [3]. To date, published studies of linkage in OA using affected sibling pairs have been limited in their statistical power [50, 51] and the results of population association studies may be difficult to assess because of stratification effects [52–54]. Studying twins offers the potential to overcome a number of the methodological weaknesses inherent in these other study designs.

By carrying out linkage and within-family association analyses in DZ pairs, the risk of misclassification of disease status through varying disease expression with age is reduced. The contemporaneous birth of DZ twins abolishes any influence that year of birth (i.e., cohort effects) may have on the development of the disease. The risk of non-paternity in twins is reduced to almost zero. The matching of the shared environment of twins may also serve to reduce the effects of confounding.

Intriguingly, data from MZ twins also have the potential to enhance the search for genetic linkage and association. If phenotype data from MZ twins are included in analytical designs that also include data on the phenotype and genotype of DZ twins, multivariate analysis of the full twin model increases the precision with which genetic and environmental effects can be partitioned [3]. This full twin model, therefore, has the potential to increase the power to detect genetic linkage in diseases such as OA through better control of environmental confounding and background genetic variance [55]. Furthermore, including data from MZ twins in this way also allows an assessment of the overall size of an individual gene’s influence on the occurrence of the disease, having accounted for shared environmental influences. Estimating the overall contribution of individual genes detected through linkage and association studies has implications for the use of genetic tests for screening future disease risk.

One difficulty that must be faced in identifying the genetic basis of OA is the need for an adequate disease definition. A common thread among reported family studies has been that the genetic influences on OA appear to be strongest for generalized OA. This is intuitively appealing as it provides a constitutional explanation for a tendency to develop disease at multiple joint sites. The definition of ‘generalized OA’, however, remains to be clarified by clinicians and epidemiologists [56]. Whilst clustering of hand and knee OA in populations has been firmly established, involvement of the hip, other peripheral joint sites and the spine is less clear [57]. Other features, such as asymmetrical involvement of the hands or medial involvement at the hips, may be important in indicating a tendency to generalized disease [56, 58, 59]. Clinical characteristics such as the presence of chondrocalcinosis and calcium pyrophosphate deposition disease may further distinguish subgroups with a specific genetic risk [60–62].

The choice of methods used to combine data from joint sites in genetic studies has in the past been empirical, and may introduce problems in interpretation and comparison between studies. For example, data from MCP joints are sometimes included in the definition of hand OA without obvious justification [8]. Summing the number of joints involved to generate a global OA score weighs analyses in favour of hand involvement. This approach also does not consistently take into account disease severity. An allied problem in defining disease at individual joint sites is the use of individual features that have been chosen for convenience of reproducibility rather than necessarily reflecting the underlying pathological process. In our MRI spine data [32], for example, genetic factors appeared to have a bearing on disc bulge, but not on disc signal change, indicating that these two common features of disc degeneration may have a quite different aetiological basis.

Examining the pattern of disease in twins provides a means by which the genetic basis of disease heterogeneity in OA can be explored. Statistical software has been developed recently that allows complex multivariate modelling of the genetic and environmental variance
components that simultaneously determine multiple disease phenotypes [1, 63]. These methods should allow the common genetic basis of individual disease features of OA and of clustering at multiple joint sites to be identified. This approach can also be used to assess genetic heterogeneity by gender in OA. Applying these methods to study designs that include opposite-sex DZ twin pairs will also allow an understanding of whether the proposed difference in heritability of OA in males and females [8, 64] is due to different genetic influences or to differences in the expression of common genes. These analyses should allow investigators to focus methods of gene detection on specific groups of individuals and on combinations of traits that have a clear genetic basis. Clarifying the OA phenotype in this way should enhance the power of studies searching for disease genes.

Twin data will in the future also provide an important contribution to understanding the precise mechanisms of disease and the way in which specific genes interact with the environment to cause disease. The strength of the twin design lies in the background genetic matching of MZ twins which makes differences in gene expression in disease discordant pairs easier to detect. Studies of the differential display of mRNA from tissue samples in disease discordant MZ pairs, for example, should allow identification of genes that are active in the disease process. Such differential gene expression is likely to be the result of different environmental influences that could be identified in a case–control setting. The twin model also provides several ways in which to study how genes and the environment interact to cause disease [1]. One example, unique to twins, is the ability of MZ twin data to differentiate clearly between genotypes that are associated with different degrees of environmental sensitivity. As phenotypic differences in MZ twin pairs are the result of differences in environment alone, MZ genotypes associated with different levels of MZ intra-pair variance may reflect a different response in each genotype to environmental change. Identifying ‘environmental sensitivity genotypes’ may prove to be important in understanding the mode of action of disease genes [65]. In the clinical setting, for example, they may determine an individual’s response to a treatment intervention [66].

Twin studies have had an important contribution to the understanding of OA by confirming the view that there are substantial genetic determinants of the disease, and indicating that the search for specific genes in the population is feasible. In the future, the contribution of twin data will be to define better the genetically relevant phenotype, to complement other family studies in OA by focusing the search for individual genes, to provide a model for uncovering the genes that are active in the causal process, and to provide methods of identifying the ways in which genes and the environment interact to cause disease. Twins are becoming more common and currently represent about 1 in 80 live births in Caucasian populations. The ‘natural experiment’ of twinning provides a unique method for defining the genetic architecture of OA and other common complex diseases.

A. J. MacGregor and T. D. Spector  
Twin Research and Genetic Epidemiology Unit, St Thomas’ Hospital, London SE1 7EH, UK

References


35. Phillips DIW. Twin studies in medical research: can they tell us whether diseases are genetically determined? Lancet 1993;341:1008–9.
Virus-associated vasculitides

The role of viral infections in the development of systemic vasculitides was demonstrated many years ago and a large panel of viruses has been implicated [1–5]. The attention of clinicians arose from the demonstration of hepatitis B virus (HBV) in classic polyarteritis nodosa (c-PAN) [1] and, more recently, from the role of hepatitis C virus (HCV) in nearly every case of mixed cryoglobulinaemia [5]. The attention accorded viral infections also comes from their therapeutic and clinical consequences, discussed in this editorial, in the context of systemic vasculitides.

In most cases, the virus-infection relationship was limited to the description of this association and coincidences could not be excluded. Although anecdotal, the probable role of the virus has been shown for B19 parvovirus [4, 6] and some cytomegalovirus infections.

A second group of viral infections concerns a larger number of vasculitides, in particular HBV, which has been incriminated in c-PAN. Since 1986, and in 1990–1992, the most recent period for which data have been published. HBV-related PAN accounted for <1% of the PAN group of vasculitides. Now, due to the greatly improved safety of blood transfusion and an extensive vaccination programme for high-risk people, and teenagers, the disease has become rare and only nine cases were registered in France in 1998 (personal data).

Human immunodeficiency virus (HIV) has been shown to be one of the aetiologies of systemic vasculitis, but, for unknown reasons, this association remains rare [3, 8].

Over the last 10 yr, HCV became the emerging virus in patients with vasculitis. It has been detected in patients with type II or type III cryoglobulinaemia, formerly called ‘essential’ mixed cryoglobulinaemia, and is the pathogenic agent for >80% of mixed cryoglobulinaemia cases [5, 9]. The disease was first observed in Mediterranean countries, but is now present worldwide and its prevalence is increasing.

Viral infections have been demonstrated to be the aetiology of some cases of vasculitides. For others, bacteria have been suggested, like, for example, Staphylococcus aureus in Wegener’s granulomatosis [10].

The idea that every vasculitis is the consequence of an infection is supported by recent data: 10 yr ago, the viral aetiology of mixed cryoglobulinaemia was only a hypothesis. For the patients in whom the virus has not been detected, we can propose that a virus is probably present, but not detected with the usual techniques, or these techniques are not adapted to unknown virus(es) or to virus mutants. It is, therefore, necessary to evaluate systematically newly discovered viruses as potential aetiological agents for vasculitis, as was recently done, with no success, for GBV-C virus which cannot be considered responsible for cryoglobulinaemia [11] or PAN, microscopic polyangiitis or Churg–Strauss syndrome [12].

In an experimental murine model, herpes virus was shown to be the cause of large-vessel vasculitis [13], but no such association has ever been proven in humans.

Certain viruses seem to be preferentially involved with vessels of a certain size. Large-vessel vasculitides have not yet been associated with virus infections. Medium-vessel vasculitides can be associated with HBV infection [7], e.g. for c-PAN, but not for Kawasaki disease, whose aetiology remains unknown, although infection seems probable. Small-vessel vasculitides can be associated with HCV infection, as in cryoglobulinaemia. Conversely, viruses have not been found in antineutrophil cytoplasmic antibody-associated vasculitis [14]. HIV vasculitis can involve vessels of any size, but, usually, small-sized ones [3].

Different mechanisms of virus-associated vasculitides have been described; the role of immune complexes is accepted in cryoglobulinaemia and c-PAN, whereas various pathogenetic mechanisms have been discussed for HIV-associated disease, e.g. direct action of HIV on endothelial cells, an excess of CD8+ cells and/or immune complexes [8]. Indeed, all these mechanisms could be involved in the vasculitic process, but at different stages of the disease: excess CD8+ lymphocytes when vasculitis occurs during stage A and a pathogenic agent (virus, bacterium or parasite) when it occurs during AIDS [15].

The viral infection does not influence the clinical presentation of vasculitis. HBV-related PAN is obviously the most characteristic form of c-PAN and HCV-related...
cryoglobulinaemia has the same clinical manifestations as cryoglobulinaemia without concomitant HCV infection. Despite the similarity of symptoms when a virus is found, outcome is affected by the virus and the ability of specific treatments to clear it, as demonstrated for HBV PAN [7] and HCV cryoglobulinaemia [16]: the spontaneous outcome of HBV infection is recovery with HBe antigen (Ag) to anti-HBe antibody (Ab) seroconversion and, less frequently, from HBsAg to anti-HBsAb. HBV PAN is usually a recent post-infection event, and although some patients exhibit an accelerated evolution towards chronicity, chronic hepatitis is rare. In HCV-cryoglobulinaemia, regardless of the virus genotype, the spontaneous outcome is to disease chronicity, with HCV persistence, leading to chronic hepatitis and relapses of cryoglobulinaemia flares [16]. Nevertheless, follow-up of patients with other virus-associated vasculitides showed different patterns. For instance, in HIV vasculitis, the disease is limited to one flare [15], despite the persistence of the virus. In HBV PAN, it was also possible to observe that the patient can recover from the vasculitis despite persistent virus replication [7]. Thus, for some vasculitides, the relationship between virus evolution and clinical outcome has not yet been established.

The presence of a virus should modify the therapeutic strategy usually adopted for systemic vasculitides, which combines steroids and cyclophosphamide. This combination is effective on the clinical symptoms [17] of vasculitis, but stimulates virus replication. The paradox of this approach is that the clinical symptoms of vasculitis are rapidly attenuated, whereas, over the long term, virus persistence is facilitated and thus its consequences: more relapses and specific manifestations due to the virus, i.e. liver cirrhosis and cancer for HBV or HCV infection.

The therapeutic strategy that we developed is based on combining treatments adapted to the aetiology and the pathogenesis of vasculitis. Of course, successful treatment reflects the effectiveness of antiviral agents, which explains the poor results obtained in HCV cryoglobulinaemia [16, 18].

Conversely, excellent results have been obtained in HBV PAN [7]. Based on the efficacy of antiviral agents in chronic hepatitis and of plasma exchanges in PAN, we combined both therapies to treat HBV PAN. The rationale of the therapeutic sequence is as follows: initial corticosteroids to control rapidly the most severe life-threatening manifestations of PAN which are common during the first weeks of the disease, and their abrupt cessation to enhance immunological clearance of HBV-infected hepatocytes and favour HBeAg to anti-HBeAb seroconversion; plasma exchanges to control the course of PAN.

Forty-one patients with HBV PAN [10] were treated with antiviral therapy (35 with vidarabine and six with interferon-z2b). The 10 yr survival rate was 83%. HBeAg to HBeAb seroconversion was obtained in 21/41 (51.2%) patients and total virus clearance (HBeAg to HBeAb and HBsAg to HBsAb seroconversions) was documented in 10 (24.4%) cases.

No treatment is able to cure mixed cryoglobulinaemia and no therapeutic strategy is clearly defined. Because steroids and immunosuppressants, as for HBV PAN, generate paradoxical responses, a strategy combining antiviral drugs and plasma exchanges is recommended [18]. No argument supports treating asymptomatic patients for whom monitoring could be sufficient. For patients with moderate symptoms (e.g. arthralgias, purpura, peripheral sensory neuropathy), interferon-z alone or in combination with ribavirin should be tested. The probability of obtaining and sustaining a long-term disappearance of the virus and cryoglobulins remains low: <25% of the patients, despite good initial responses. For patients who have severe clinical symptoms of mixed cryoglobulinaemia, plasma exchanges could be useful. The treatment is symptomatic and is able to cure leg ulcers and reverse other manifestations. Nevertheless, they cannot be stopped for most patients because a rebound phenomenon may occur. They should, therefore, be used in conjunction with antiviral drugs (the interferon-z-ribavirin combination is the most effective) or other treatments. For severe vasculitis that does not respond to antiviral agents and plasma exchanges, steroids and/or immunosuppressants are indicated. A similar therapeutic approach can be proposed for other virus-associated vasculitides, like HIV vasculitis, and is usually effective against the symptoms of vasculitis, despite the persistence of HIV [15].

L. Guillevin
Hôpital Avicenne, Department of Internal Medicine, Université Paris Nord, 125, rue de Stalingrad, 93000 Bobigny, France

References


