Editorial

Translating Genetic Information into Clinical Disease Risk in Rheumatoid Arthritis

The increasing availability of genetic information on human subjects has raised expectations that in the near future genetic knowledge will make a central contribution to the management of rheumatic disease. Genetic markers as measures of disease risk and predictors of outcome have an inherent appeal: they are tangible and fixed, and relate directly to biological mechanisms of disease. Given that twin and family studies clearly show that a greater part of variation in disease prevalence for most common rheumatic disease is explained by population-level genetic variation, it seems reasonable to anticipate that genetic data will have potential to reveal a level of insight that could inform the management of disease in clinical practice.

Hopes that genetics will transform clinical management are high for rheumatoid arthritis (RA). Around 60% of the variation in population occurrence of this disease has been attributed to genes[1]. The contribution of the HLA system to this risk has been recognized since the 1970s[2-4], and it is now acknowledged that individual HLA variants encoding the shared epitope (SE)[5] confer specific levels of risk in relation to the severity and rate of progression of the disease[6,7]. The last decade has seen the completion of several genome-wide linkage scans as well as numerous candidate gene association studies that have extended our knowledge of the genetic contribution to RA susceptibility beyond HLA. Other confirmed associations now include genes encoding molecules involved in cell signaling (PTPN22[8], CTLA4[9]), and posttranslational modification (PADI4[10]). Further discoveries are anticipated from the whole-genome association studies currently in progress.

Given the wealth of genetic data in RA, it is reasonable to question objectively how genetic information might be translated into predicting the course of disease in individual patients. RA is clinically heterogeneous, and the value of genetic testing to an individual is dependent to a degree on their existing risk as evidenced by their demographic char-

acteristics and by the pattern and course of their disease. Thus, the measure of greatest interest is the extent to which additional knowledge from genetic testing can augment the ability to predict outcome based on an individual's preexisting risk factor profile.

Predicting disease outcome in RA presents significant challenges. RA has a relatively low incidence; the disease is often difficult to define at its outset; it has a lifelong relapsing and remitting course, which is influenced by individual-specific treatment decisions that also vary over time. Population studies may not reflect the patterns of disease managed in rheumatology practice, and risk factor profiles may differ between outcome studies. An adequate appraisal of the risk associated with particular genetic polymorphisms can only be derived from representative patient samples in a longitudinal setting with subjects followed over many years. These studies are rare.

The study by Janssens, et al[11] in this issue of The Journal is one of the few to examine the clinical utility of individual genetic testing for the SE in determining risk in RA among typical clinic attenders in a longitudinal setting. It provides results that are directly applicable to clinical practice. The study was conducted among a cohort of 154 Caucasian women with established RA who had been under the care of rheumatologists in Northern California, USA. The cohort had an average disease duration at inception of 10 years and had been followed for 12 years.

Their analysis demonstrates that knowledge of an individual's SE status is only helpful in restricted circumstances. As an individual's score increases on conventional clinical risk factors for erosive disease [including age at disease onset, family income, number of painful joints, Health Assessment Questionnaire, global pain rating, and seropositivity for rheumatoid factor (RF)], the likelihood that the SE is present also increases. The consequence is that, in individuals with clinical risk factors that predict a high prior

See Value of the HLA-DRB1 shared epitope, *page 2383*

likelihood of erosive disease, SE testing is only truly informative if it is negative. In these circumstances the risk of developing future erosions is halved. A positive SE test (especially if RF is positive) provides little additional prognostic information. Thus in subjects with RF-positive disease that is running an aggressive course, knowledge of SE positivity provides little additional predictive information on the risk of future erosions beyond clinical observations alone.

The size of these changes in magnitude of risk of erosive disease that are associated with SE positivity or negativity makes it unlikely that SE testing can be justified as part of the routine management of RA. Given the sustained interest that has surrounded the association between the SE and disease, this might seem disappointing. However, the fate of the SE as a predictive genetic test in RA is consistent with similar findings in a range of complex diseases, and questions have been raised as to whether genetic testing can ever have a place in the clinical management of these conditions[12].

The limited value of single common risk variants in complex disease can be appreciated from the results of analytical studies[13]. The positive predictive value of a genetic test is a function of the frequency of the genetic risk variant, the relative risk of disease in the presence of that variant, and the overall risk of disease in the population. For complex disease with a prevalence of 5% or less, the positive predictive value for most genetic variants is low unless the genotype frequency is less than 1% or the relative risk of disease associated with the genotype is 20 or more. These figures indicate that it is unlikely that single common risk variants will alone be useful in predicting disease clinical practice.

The performance of genetic testing improves if multiple genes are considered. With modest effects of genetic risk variants (equivalent to odds ratios in the region of 1.5 to 3.0) and modest interactions, the positive predictive value for individuals with 2 or more variants can increase 10-fold[14,15]. Simulations involving 20 genetic loci, each with odds ratios of 1.5 to 1.7, enable a hypothetical test to predict disease with an accuracy comparable to that of total cholesterol level in the prediction of coronary heart disease[16]. When hundreds of loci are considered, each associated with an odds ratio of 1.5 or less, the discriminative ability of a genetic test can be improved substantially if (1) at least a few genetic markers are included with higher odds ratios (1.5—3.0); (2) the proportion of variance in disease explained by genetic factors is high (> 30%); and (3) the disease prevalence is low (< 10%)[17]. These 3 conditions in which the simulated test has superior discriminating accuracy all might apply to RA.

Whether these levels of prediction can be achieved in practice is less certain. The allelic spectrum that defines the genetic model of complex diseases such as RA remains unknown, and the precise mode of gene action and interaction is to an extent speculative. Models frequently assume independent genetic effects and multiplicative interactions, assumptions that tend to inflate the risk estimates. It is also questionable whether the large number of associations with low odds ratios required for many of these models could ever be identified given the large sample sizes required. Even with the advent of genome-wide association scanning, identification of the rare variants and gene-gene and gene-environment interactions presents a formidable challenge.

The phenotypic heterogeneity of RA also presents difficulties in using genetic information for prediction. Many of the reported genetic associations with the disease are confined to subgroups. For example, the strength of the HLA association with disease varies according to gender. Of the more recently reported associations with RA, PTPN22 is predominantly associated with RF-positive and anti-cyclic citrullinated peptide-positive disease, the male sex[9,18], an early age at onset[9,19], and with the coexistence of other autoimmune diseases[20]. In developing their predictive model of RA, Jenssens, *et al*[11] were unable fully to take into account subgroup effects for reasons of statistical power. Cluster methods that assume an underlying phenotypic heterogeneity might provide a more realistic approach to modeling genetic effects in this setting.

Our capacity to predict the course of RA is necessarily limited by the inherently unpredictable biological processes that drive this disease and by the vagaries of random environmental exposure. The value of predictive testing itself is also limited by our ability to prevent the onset and modify the disease course in individuals even when their risk of future disease is known. With developments in genetic and clinical medicine, it is conceivable that in time genetic information will make an important contribution that helps translate risk of future disease in RA into information that can influence clinical practice at an individual level. However, it is clear that we are still at a very early stage. For the present, the predictive value of genetic information in RA remains limited. Rheumatologists will need to continue to rely on their own clinical observations and skill to inform their decisions in managing this complex disease.

ALEX J. MacGREGOR, MD, FRCP,
Institute of Health,
University of East Anglia,
Norwich;
SOPHIA E. STEER, PhD, MRCP,
King's College Hospital,
Denmark Hill, London, UK

Address reprint requests to Prof. A.J. MacGregor, School of Medicine, University of East Anglia, Norwich, NR4 7TJ, UK.
E-mail: a.macgregor@uea.ac.uk

REFERENCES

1. MacGregor AJ, Snieder H, Rigby AS, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000;43:30-7.
2. Stastny P, Fink CW. HLA-Dw4 in adult and juvenile rheumatoid arthritis. Transplant Proc 1977;9:1863-6.

3. Panayi GS, Wooley P, Batchelor JR. Genetic basis of rheumatoid disease: HLA antigens, disease manifestations, and toxic reactions to drugs. BMJ 1978;2:1326-8.

4. Stastny P. Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. N Engl J Med 1978;298:869-71.

5. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 1987;30:1205-13.

6. Weyand CM, McCarthy TG, Goronzy JJ. Correlation between disease phenotype and genetic heterogeneity in rheumatoid arthritis. J Clin Invest 1995;95:2120-6.

7. Gorman JD, Lum RF, Chen JJ, Suarez-Almazor ME, Thomson G, Criswell LA. Impact of shared epitope genotype and ethnicity on erosive disease: a meta-analysis of 3,240 rheumatoid arthritis patients. Arthritis Rheum 2004;50:400-12.

8. Begovich AB, Carlton VE, Honigberg LA, et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004;75:330-7.

9. Plenge RM, Padyukov L, Remmers EF, et al. Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. Am J Hum Genet 2005;77:1044-60.

10. Suzuki A, Yamada R, Chang X, et al. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 2003; 34:395-402.

11. Janssens AC, Steyerberg EW, Jiang Y, Habbema JD, van Duijn CM, Criswell LA. Value of the HLA-DRB1 shared epitope for predicting radiographic damage in rheumatoid arthritis depends on the individual patient risk profile. J Rheumatol 2006;33:2383-9.

12. Terwilliger JD, Weiss KM. Confounding, ascertainment bias, and the blind quest for a genetic 'fountain of youth'. Ann Med 2003;35:532-44.

13. Holtzman NA, Marteau TM. Will genetics revolutionize medicine? N Engl J Med 2000;343:141-4.

14. Yang Q, Khoury MJ, Botto L, Friedman JM, Flanders WD. Improving the prediction of complex diseases by testing for multiple disease-susceptibility genes. Am J Hum Genet 2003;72:636-49.

15. Khoury MJ, Yang Q, Gwinn M, Little J, Dana FW. An epidemiologic assessment of genomic profiling for measuring susceptibility to common diseases and targeting interventions. Genet Med 2004;6:38-47.

16. Yang Q, Khoury MJ, Botto L, Friedman JM, Flanders WD. Revisiting the clinical validity of multiplex genetic testing in complex diseases: Reply to Janssens et al. Am J Hum Genet 2004;74:589-91.

17. Janssens AC, Aulchenko YS, Elefante S, Borsboom GJ, Steyerberg EW, Van Duijn CM. Predictive testing for complex diseases using multiple genes: fact or fiction? Genet Med 2006;8:395-400.

18. Pierer M, Kaltenhauser S, Arnold S, et al. Association of PTPN22 1858 single-nucleotide polymorphism with rheumatoid arthritis in a German cohort: higher frequency of the risk allele in male compared to female patients. Arthritis Res Ther 2006;8:R75.

19. Steer S, Lad B, Grumley JA, Kingsley GH, Fisher SA. Association of R602W in a protein tyrosine phosphatase gene with a high risk of rheumatoid arthritis in a British population: evidence for an early onset/disease severity effect. Arthritis Rheum 2005;52:358-60.

20. Criswell LA, Pfeiffer KA, Lum RF, et al. Analysis of families in the Multiple Autoimmune Disease Genetics Consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. Am J Hum Genet 2005;76:561-71.