

Setting Research Priorities for Arthritis: The Environmental Perspective

JOHN W. FRANK

ABSTRACT. Recent enthusiasm for genetic advances in prevention is out of keeping with the etiology of most common diseases of the industrialized world, including the major inflammatory arthritides. These conditions have a genetically “complex” causation, involving many genes, and strong influences of the environment, acting on our individual genetic endowments over the entire life course. Lines of evidence that this is so are reviewed - especially migrant epidemiological cohort studies, which are stronger etiological evidence than “twins reared apart” studies, since they tend to involve massive cultural and environmental change, while “holding genetic factors constant.” More such studies would better inform preventive strategies for the inflammatory arthritides, which lag behind cardiovascular disease in understanding causation, and therefore primary prevention. Finally, factors are briefly reviewed that affect risks, benefits, and costs of single-locus genetic tests to predict lifelong risk of chronic diseases with complex and multifactorial determination. Both negative and positive predictive values of such tests for predicting lifetime disease occurrence are generally unacceptable for use in the general population. Expert genetic counseling is therefore important before such testing, to ensure that an appropriate family and personal history justifies these expensive tests, the “labeling” effects of which can last a lifetime. (J Rheumatol 2005;32 Suppl 72:58-61)

Key Indexing Terms:

ARTHRITIS

PREVENTION AND CONTROL

GENETIC SCREENING

EPIDEMIOLOGY

It is a challenge to discuss the environmental approach to setting research priorities for any disease, but one useful departure point is to consider the prospects for primary prevention. True primary prevention for the chronic common diseases in our society is a challenge that transcends arthritis. But there are important clues, concerning the relative roles of environment versus genetics in prevention, from a number of other chronic diseases (those with a 50-year head start in etiological research). However it is encouraging that we are closer to a complete breakthrough in terms of major advances in treatments. I will focus primarily on etiology, since that is what epidemiologists, particularly population-based ones, are known for. This topic has recently been reviewed¹.

The enthusiasm implicit in media reports of newly discovered “disease genes” is the notion that genetic advances in the Human Genome Project will imminently lead to improvements in population-level health, from which everybody will benefit. Of course, it is true that we have had some significant breakthroughs in understanding how genes influence disease. It is significant that most

of these breakthroughs have to do with pathogenesis rather than the first steps in disease causation per se. However, these insights are reaching farther and farther upstream, which is very exciting, particularly for someone who has a strong interest in prevention. But the expectations of genetic breakthroughs set out for the public are unachievable, unless the environment is factored in.

In particular, the metaphor of DNA as the blueprint for our health, and indeed for our lives, is widely used. James Watson actually calls DNA “the book of life.” The difficulty is that that’s not exactly the way things work. There are serious shortcomings in the blueprint analogy. If one buys a house built with an architect’s blueprint, one may later look at the house to see how it is functioning, particularly some years after purchase. This later functioning is what we are talking about as a metaphor for human health. If the house does not correspond precisely to the blueprint, in terms of its structure and functionality, one has a legal case with the architect.

In contrast, the degree of concordance between identical twins for many common diseases is generally only around 40–60%; it is rarely higher for common diseases, because the diseases are multifactorial. So the notion of blueprint is just too deterministic. Perhaps “jazz score” is a better analogy, recognizing that a jazz person may think that’s pretty ridiculous — who uses a score in jazz? But the idea is that some jazz players do use a score and a small epidemiological study compares the genome sequence to a jazz score². The environment “plays upon the person” who is the musician in a way that constantly modifies what one hears when one listens, in the sense that human health and functioning is like the music that

From the Institute of Population and Public Health, CIHR; University of Toronto; and Institute for Work and Health, Toronto Canada.

J. Frank, MD, CCFP, MSc, FRCPC, Scientific Director, Institute of Population and Public Health; Professor, Public Health Sciences, University of Toronto; Senior Scientist, Institute for Work and Health.

Address reprint requests to Dr. J. Frank, Institute of Population and Public Health, Banting Building, 100 College Street, Ste. 207, Toronto, ON M5T 1L5.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2005. All rights reserved.

results from this interaction.

One reason for confusion is that many people with genetic expertise, including some clinical geneticists, have spent their lives studying Mendelian inheritance, which means studying diseases and traits caused by one gene. And there are still crucial roles for them in our healthcare system, because for those diseases we are better at identifying them early and, in some cases, modifying their course. But they constitute a relatively small burden of illness in the population, compared to multifactorial diseases. A simplistic graph from Chris Murray and Alan Lopez's 1997 *Lancet* article³ shows the pattern of the burden of illness in a typical industrialized country (Figure 1), with cardiovascular disease at the top and cancers, neuropsychiatric conditions, and respiratory and digestive conditions following. These diagnostic labels of course are heterogeneous, but none of them is largely Mendelian, in terms of the genetic contribution to their causation. The causal pathways of these conditions may involve dozens, hundreds, or thousands of genes acting in concert, in an intimate connection with the environment, as we live through it over our life course.

The quintessential example in rheumatology is reactive arthritis, in which there is dynamic interaction of genetic factors and environmental factors such as arthritogenic microorganisms. All the evidence for rheumatoid arthritis (RA) indicates that it is also a multifactorial disease. RA and most inflammatory arthritides are much more like heart disease, cancer, and other common conditions that are not Mendelian. So in thinking about multifactoriality, one must consider that there are broad fields of determinants of occurrence of such a condition – genetic, environmental, or related to diet and lifestyle, and social structure.

As an aside, interestingly, there is little written about social class gradients in inflammatory joint diseases. In developing research priorities, this could be a very important theme, since such research often provides very important clues to the determinants of health. Are there social class gradients for inflammatory joint diseases? Are they in the usual direction? These studies are usually not about genetics, but rather differential, cumulative environmental exposures over a lifetime.

A second theme, mentioned below, concerns what happens if one tries to use genetic tests for complex, multifactorial conditions where one has identified only a single locus or 2 loci, and one is going to use that test to predict lifetime occurrence of a disease. There are some problems with that approach⁴.

One revealing area of research, migrant studies, are the population equivalent of “twins reared apart” study designs. In developing an etiological research program for arthritis, it might be very informative to include migrants from areas with relatively high prevalence of arthritis migrating to areas with much lower levels, and vice versa. The best migrant studies hold constant the genetic component of causation while the environment changes. For

example, studies from San Francisco on cancer, heart disease, and stroke followed Japanese families who migrated via Hawaii over 100 years or more⁵. The studies examined rates of occurrence of various chronic conditions over the immigration period and afterward.

In the case of breast cancer, they demonstrated a 6-fold gradient in disease risk as people moved. Since it is essential in these studies to hold genetics constant, people must continue to intramarry within their ethnic group, which was the case for the Japanese migrants. They intramarried for one or 2 generations in the new country. This study demonstrated that this cancer and other common Western cancers are profoundly, broadly “environmental.” It doesn't simply mean nutrition, chemicals, or physical agents, since it may also involve biological agents (e.g., the hepatitis B virus in the case of hepatocellular carcinoma). Causation in these conditions therefore certainly has to do with more than genetics, even though their detailed pathogenesis necessarily involves many genetic mechanisms. Genetics helps determine which individuals express the disease, but is not responsible for much of the actual causal path at the population level.

There are other analogies between well-studied chronic diseases' causation and arthritis prevention, and one should be searching for them. In order to do really efficient primary prevention, one must find a continuously distributed marker of risk. Cancer epidemiology is being held back by an inability to identify such a factor. There are plenty of contenders, but they have relatively weak explanatory power for any given cancer.

Cardiovascular disease researchers were fortunate to discover early at least 4 or 5 continuously distributed risk markers for later disease, which allowed them to build the Framingham equation as of the early 1960s, with only a decade of followup data from that town. The prevention legacy from that early work has been perhaps most succinctly captured by Geoffrey Rose, in a must-read book by the most important epidemiologist of our time⁶. Rose said that there is likely to be a population distribution of a risk factor [like low density lipoprotein (LDL) cholesterol for heart disease], so one should be thinking what it might be for any kind of arthritis. Is there a biomarker that confers risk for arthritis that could be like that? Is it an autoantibody? (Note that sometimes the risk relationship runs in the reverse direction, where in fact it is protective to have a high level of the factor. So graphs of folate levels, for example, or income versus heart disease risk, run in the opposite direction, although the curve is about the same shape, as for LDL cholesterol or blood pressure.)

Rose further pointed out that there are 2 ways a society develops lots of a chronic disease, based on this sort of risk factor distribution. One is, that the whole population distribution (“bell curve”) of the risk factor is moved over, which is what has happened to modern Western societies. Our distribution of LDL cholesterol, body mass index, blood pressure, sedentary living, daily intake

of calorie-dense food – all these whole curves are shifted over, to the right by “upstream forces” – cultural forces; the *vector* of the modern chronic disease epidemic is culture. Once the whole distribution has been shifted over, one sees a high incidence population, rendering inadequate and inefficient clinical approaches that identify individuals at high risk one at a time, and put them on expensive lifelong management like statins, or professionally supervised exercise and diet. Such intensive clinical treatments can only deal with the tip of the curve, because only there will aggressive interventions be cost-effectively justified. In the middle of the distribution is a huge burden of illness attributable to people with “locally low” but “globally high” levels of the risk factor, in whom only upstream population-level interventions, aimed at the vector “culture,” can make a difference.

Thus if it turns out that certain autoantibodies linked to inflammatory arthritis are really elevated in our population compared to others, and that there are some we think could be lowered, the key question becomes, “What would be an ‘upstream’ intervention to shift the whole distribution?” A high-risk approach always is a little bit doomed because if one identifies sufficiently high-risk individuals, one ethically must try to treat them. But there is always a grey zone where one can no longer justify what is going to be done to them as individuals, because their individual risk doesn’t justify the risks and costs of clinical risk reduction. Then one faces a dilemma: the challenge of figuring out which patients – if elevated arthritis risk from those autoantibodies is what one anticipates – will be the ones to treat. How would one do it, and how will they be identified and, most importantly, where are those antibodies coming from to start with?

Such an etiological analysis must recognize that one can not find effects due to the environment where there is no variance in the environment. One can not identify a nutritional exposure to be a determinant of heart disease, for example, where there’s no variance in a society that all eats the same way, as might have been the case in the US and Canada, for example 50 years ago. That is why we have needed international cardiovascular cohort studies, in which diet varied greatly. One can only learn about the environmental factor when there is variance. We think of PKU as a genetic disease, but if we lived in a society on another planet where the PKU gene was common but where very, very few people had a diet with phenylalanine in it, PKU would be called an environmental disease. So study context is critical in diseases that result from strong interactions between genes and environment — which is most chronic diseases.

Geoffery Rose summarized his arguments this way⁶. To do primary prevention, learn the lessons from diseases that have gone before. And cardiovascular disease “got there” in the 1950s and 1960s. Cardiovascular primary prevention has not been unsuccessful and although we still don’t understand much of the 60%+ decline in coronary disease mortality since then, some of it certainly has

been occasioned by scientific knowledge and the policy actions that fell from that, best seen in tobacco control. In sum, to change a whole population’s level of a chronic disease that is multifactorial, one must look at those upstream forces that shift the distribution curve of risk factors.

Walter Willett, the noted nutritional epidemiologist, has recently written a critique of “geneticism”⁷. What is geneticism? It’s a term coined by social scientists that describes the scientific approach of focusing on exclusively genetic causes of complex multifactorial diseases, without investing in a “balanced portfolio” of research with both genetic and environmental risk factor/exposure measurements, to elucidate etiologic pathways that explicitly acknowledge interaction. Willett’s argument, extended to a disease such as RA, might lead to the rebuttal that one can’t do this yet for RA because we don’t know how much of RA is caused by each of the particular risk factors that would have been identified by good cohort studies if they had been done. Thus one way to begin would be with focussed cohort studies examining the health consequences of having candidate autoantibody levels, and suspected cofactors. It would be very important to get a whole list of such cofactors and measure them. When one does that, one gets results like we have for other chronic diseases. One finds out that there is, if one takes all the risk factors, a certain proportion of the occurrence of the disease that one can explain with existing risk factor knowledge. So for diseases such as colon cancer, stroke, coronary disease and type 2 diabetes, there’s a very large (70%–90%) proportionate reduction we could achieve in those diseases just from our current knowledge of the value of diet, exercise, smoking, and similar factors that we can, in theory, modify if we could get a handle on culture, the vector.

Willett’s point, in his landmark 2002 Science article⁷: Who needs genes to prevent these diseases – if there are hundreds of genes, will each one have a small effect? That doesn’t mean we shouldn’t research genetic components of etiology: There will be subpopulations, such as some ethnic groups, who have importantly elevated risks because they carry, at a high frequency, risk-conferring genes that are uncommon in the whole population.

Now for my second theme: What happens when one tries to use, in the general population, lifelong predictive tests, genetic tests for chronic disease occurrence, such as *BRCA1* or *BRCA2* for breast cancer? These tests are now widely used in Ontario, but only through an appropriate screening system with proper genetic counseling before testing; this is necessary because the tests are costly, present some complications for one’s life, and are not 100% accurate. They have sensitivity and specificity problems. But much of the public doesn’t “get it” yet, since they would apparently still like to have a result of a genetic test, even with no family history of breast cancer. But they shouldn’t do that for important reasons other than resource wastage. The risk-cost-benefit balance in testing

low risk persons from the general population is unfavorable. For example, we define the clinical efficiency of preventive treatments by the “number needed to treat;” and similarly the “number needed to screen” that describes the clinical efficiency of testing programs refers to the number needed to get each clear beneficiary.

As well, there is a false negative problem. Suppose there were a single locus that explained 5% of the genetic contributions to the occurrence of RA in our population (this is about the proportion of breast cancer explained by the *BRCA1* and *BRCA2*). One can anticipate public pressure for genetic screening for the locus. However the gene cannot possibly explain most of the disease occurrence overall. So when people have a negative test some will mistakenly think they are “guaranteed” not to develop the disease and may ignore early signs of (or forego effective preventive measures for) the disease later. This is referred to as the “false reassurance effect,” and may be harmful, for example, if one misses out on early and promising treatments like the new biologicals for RA. One does not want people with early symptoms to ignore them because they had a test years before that they interpreted as indicating that they couldn’t get a disease. In fact the test isn’t designed to rule out lifelong propensity.

There is also an inherent test specificity problem. In genetic terms, this means low gene penetrance (those with markers may not necessarily develop disease). Studies done soon after a disease gene is discovered often examine families self-identified by the presentation of very dense case clusters, so they overestimate the penetrance: the life-long probability of the disease, given that one has the marker. But with each subsequent year, further publications start appearing and down goes the penetrance, because later and better studies come from more representative populations that are unselected, and not confounded by other things like gene-to-gene and undetected environmental interactions.

This is particularly illustrated by hemochromatosis. There was initially high enthusiasm for the predictive power of screening for a single mutation apparently underlying much of this disease. But when properly studied in an unbiased population sample with no “referral filter” bias or “spectrum” bias, one finds out that the mutation doesn’t perform as a very good disease predictor⁸.

In the case of *BRCA1*, the positive predictive value is so low that one must screen 2500 women to prevent one case⁹. To achieve this, one must also put 4 women who will not benefit on lifelong chemoprophylaxis or do prophylactic mastectomy and oophorectomy. We have all learned from many other lessons in screening that one must have high pretest-likelihood (i.e., high risk to start with) in the population screened. There is one difference, however, in that genetic tests, unlike usual screening tests for the current presence of cancer (e.g. mammography), may take a lifetime to manifest their consequences

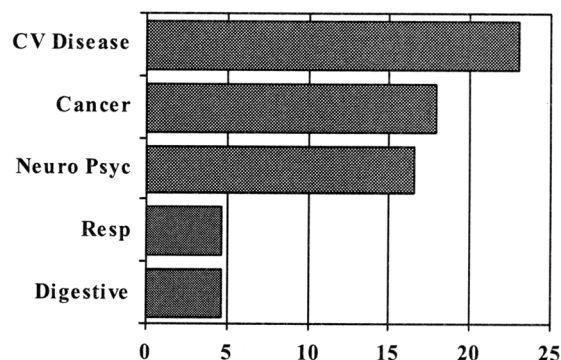


Figure 1. The top 5 disease in industrialized countries ranked by percentages of total disease burden. with permission from Murray CJ, Lopez AD. Lancet 1997; 350:144

because people are permanently labeled, in terms of their risk status. In other words, one’s genetic test result is there forever, and it may have implications for one’s family as well.

To sum up, when a field starts to invest in research for primary prevention, there are many lessons to be learned from other diseases that have had 40 to 50 years of trying to achieve primary prevention by applying knowledge from research. On the other hand, those in the arthritis field are at the dawn of a new era, and I wouldn’t hold you back.

REFERENCES

1. Frank JW, Lomas G, Baird P, Lock M. Genes and environment in human health: the need for a balanced approach. In: Heymann J, Hertzman C, Barer M, Evans R, editors. 2004. Healthier societies: A guidebook. Oxford: Oxford University Press; 2004 (in press).
2. Porta M. The genome sequence is a jazz score. *Int J Epidemiol* 2003;32:29-31.
3. Murray CJ, Lopez AD. Global mortality, disability and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 1997;350:144.
4. Khoury MJ, Yong Q, Gwinn M, Little J, Flanders WD. An epidemiologic assessment of genomic profiling for measuring susceptibility to common diseases and targeting interventions. *Genetics Med* 2004;6:38-47.
5. Marmot M, Syme SL, Kagan A, Kato H, Cohen JB, Belsky J. 1975. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: prevalence of coronary and hypertensive heart disease and associated risk factors. *Am J Epidemiol* 1975;102:514-25.
6. Rose, G. The strategy of prevention. Oxford: Oxford University Press; 1992.
7. Willett WC. Balancing life-style and genomics research for disease prevention. *Science* 2002;296:695-8.
8. Bleutler, E, VJ Felitti, JA Koziol, NJ Ho, T Gelbart. 2000. Penetrance of 845G→A (C282Y) HFE hereditary haemochromatosis mutation in the USA. *Lancet* 359:211-18.
9. Vineis P, Schulte A, McMichael. Misconceptions about the use of genetic tests in populations. *Lancet* 2001;357:709-12.