PAPER 95-15

THE CHALLENGES OF PRODUCING HEALTH IN MODERN ECONOMIES

Greg L. Stoddart

"NOT FOR CITATION WITHOUT PERMISSION"
THE CHALLENGE OF PRODUCING HEALTH IN MODERN ECONOMIES

Greg L. Stoddart

Professor, Centre for Health Economics and Policy Analysis, and Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada

and

Fellow, Canadian Institute for Advanced Research


ABSTRACT

Although health economics may be considered to be the application of economics as a way of thinking about questions regarding the production, maintenance and distribution of health, health economists have historically paid more attention to some production inputs than to others. Health economics has been mainly, even if not exclusively, health care economics. Relatively little attention has been paid by economists to the role of other (non-health-care) determinants of health, or to the nature of the production process underlying the health of populations. This has recently begun to change, however, as the attention of both researchers and policy-makers has shifted from the downstream question, “How do people get well once they are ill?”, to the upstream question, “How do people get ill?”

A recent example of the interdisciplinary collaboration of health economists in questions related to the determinants of health is found in work from the Population Health Program of the Canadian Institute for Advanced Research. This group has been engaged for the past seven years in extracting from the international and cross-disciplinary literature a new understanding of “heterogeneities” in health status -- why some groups within populations are healthier than others. Many of the apparent underlying ‘causal’ factors -- genetic predisposition, culture, stress, control over one’s professional and personal environments, nurturing and enrichment in early childhood -- are not topics on which health economists generally have much to contribute. But others -- income; employment; economic growth; the distribution of income and wealth -- are rather closer to home. The key message from this work is the complex (and as yet far from completely understood) interplay between these myriad economic and other influences, and the equally complex biological interplay among immune, endocrine and cardiovascular systems (at least) in explaining how these influences get translated at the cellular level into ill health or premature death.

This then raises several sets of challenges -- conceptual, methodologic, empirical, practical, and political -- for researchers attempting to push out the frontiers of understanding of the relative importance of various determinants, for policy-makers attempting to act on new knowledge about determinants, and for health economists interested in participating in the collaborative cross-disciplinary ef-
forts required for both research and policy development. Perhaps the biggest single challenge, for everyone concerned about the production of health in modern economies, is to find the proper balance between investment in the quality of the social fabric necessary to facilitate and sustain economic progress, and investment in the economic growth itself essential for providing the capacity to sustain the quality of social environments over long periods.
Good health has been a universal goal of both individuals and societies since ancient times. For just as long, it has been recognized that the achievement and maintenance of health involves delicate interactions among biological, behavioural and environmental determinants. Yet for most of this century, health policy has been dominated, some might say captured, by policies concerned with health care, and more specifically by those concerned with the largest components of health care - physicians, hospitals, pharmaceuticals and increasingly, advanced diagnostic and therapeutic technologies.

The health care industry is one of the largest clusters of economic activity in developed economies (Schieber and Poullier, 1989; OECD 1989). Its major achievements and still outstanding promises have commanded the gratitude, faith and fascination of the public, the continual attention of policy-makers, and until very recently in most countries (and still in some, the United States being the most striking example) a steadily increasing share of national product. Although neither policy-makers nor the public are so naive as to believe that factors other than health care are not important determinants of health, the record of policy effort, public support, and resource allocation implies a conceptual model of determinants (Figure 1) in which health care plays the central role in reducing the burden of disease in the population while the role of “other factors”, though not ignored, is minimized (Evans and Stoddart, 1990).

In this world, individuals become ill or injured for a variety of external and unspecified reasons and present themselves to the health care system, which responds by assessing their needs and providing the services necessary to relieve symptoms and pain, provide information and reassurance, cure or control disease, and prevent future disease or deterioration. It is essentially a “repair shop” model of the health care system. If needs - as perceived by either individuals or health care providers - are infinite, then there is no self-equilibrating mechanism to control the flow of resources to the health care sector. Although obviously too simple, the framework of Figure 1 makes the role of access to health care dramatically clear, as demonstrated by the preoccupation of health policy-makers in developed countries for most of this century with the design, implementation and operation of systems of sickness care insurance, regardless of the private-public mix in those systems. The framework also makes abundantly clear why the attention of those same policy-makers has inevitably shifted from concerns about access to concerns about expenditure control (OECD, 1994a).
Like health policy-makers, health economists too have focused their attention mainly on health care. Although health economics as a field may be considered to be the application of economics as a discipline, a way of thinking, to the topic of the production and distribution of health (Williams, 1979; Culyer, 1981), health economists have historically paid more attention to some production inputs than others. Health economics has been, for the most part, health care economics. Figure 2 presents a schematic of the principal topics studied in health economics prepared by Culyer and colleagues at the University of York, England, and presented to a 1986 meeting of Canadian health economists (Culyer, 1987). Academic health economists in particular have historically devoted substantial effort to the topics in boxes C, D, and E dealing with the demand for and supply of health care and their market interactions, and more recently to boxes A, F, G and H dealing with the measurement of health, the cost-effectiveness, cost-benefit or cost-utility of specific services or programs, and system-level planning and evaluation. Considerably less effort has been devoted to box B, the role of influences on health other than health care, and the nature of the production function for the health of populations.

Recently this has begun to change, however, as the attention of both researchers (from many disciplines) and policy-makers has shifted from the downstream question, “Why and how do people get well once they are ill?” to the upstream question, “Why and how do people get ill?” Both questions are integral to the study of the health of populations, but they are very different questions and they need to be separated more often than they are. They also lead in very different directions, the former toward the economics and politics of health care and the latter toward the economics and politics of health. The need to consider more explicitly the “other factors” from Figure 1 requires economists in particular to expand their exploration of the production function for health.

For several years, a group of Canadian health economists has been attempting to do this, as part of the Population Health Program of the Canadian Institute for Advanced Research, a non-profit research institute funded by both private and public contributions. The central questions addressed by the Population Health Program may be stated in a deceivingly simple form, “Why are some people healthy and others not ..........and what does this imply for health policy?” The research agenda includes study of the determinants of health other than health care, the
effectiveness and efficiency of health care systems, the measurement of health, and the processes of health policy-making. Members of the Population Health Program are linked by the Institute in an interdisciplinary and international research network with expertise in such fields as genetics, medicine, epidemiology, anthropology, sociology, political science, economics and policy analysis. The Population Health Program is also linked to two other programs of the Institute, one in Economic Growth and the other in Human Development, which affords interaction with researchers from an even wider range of disciplines.

A primary focus of the Population Health Program since its inception in 1987 has been the gathering and synthesis of research on systematic differences in the health of groups, both within and across national populations, and the factors which are associated with and may account for these differences. The results of this work have recently appeared in three books, one directly from the Program (Evans, Barer and Marmor, 1994) and two from a conference convened by the Program which involved other investigators interested in the determinants of health in general and the relationship between health and economic prosperity in particular. (Canadian Institute for Advanced Research, 1994; Daedalus, 1994).

In this paper, I briefly review some of the main themes and findings of this continuing interdisciplinary synthesis and discuss several types of challenges that they present for improving our understanding of the determinants of health and for policy actions based on our current state of understanding.

Efficiency and the Production of Health

I begin from the view that economics as a way of thinking provides a unique perspective for framing policy issues concerning the production of health. That is not to say that an economic perspective is - or should be - the only, or even the most important, perspective, but simply to say that I think it is a very useful one.

I mean to use “an economic way of thinking” in quite a general sense, that is to say a concern for alternative ways of allocating scarce resources to competing uses of value to those making or being called upon to make the allocations. An economic perspective therefore recognizes choices among alternative ways of producing health and seeks to identify, and where possible to measure, the relative costs and benefits of those choices, and the different distributions of both costs and benefits which accompany them.
One example of a taxonomy which emerges from an economic perspective, and which may be useful in framing questions about the production of health through health care or other inputs is shown in Figure 3, drawn from Lavis and Stoddart (1994). In this taxonomy, developed to illustrate the distinct ways in which it is in principle possible for a society to devote too many resources to health care, the rows and columns are defined by concepts of efficiency and the nature of the output, respectively. Technical efficiency (maximizing output for a given combination of inputs), cost-effectiveness (maximizing output for a given cost), and allocative efficiency (producing and distributing the most highly valued mix of outputs) are the concepts of efficiency. Health care, health and well-being (a higher-order concept to which health is but one contributor) are the outputs.

In practice, the taxonomy is less straightforward than it appears. Assessments of allocative efficiency are fraught with value judgement, even in welfare economics, and distributional issues are paramount (Culyer, 1989; Reinhardt, 1990). Regarding outputs, as one proceeds from narrow to broad definitions of health, the dividing line between health and well-being can be quickly blurred (Evans and Stoddart, 1990). However, for my present purposes it is possible to set aside these issues (though we shall not escape them completely as will be seen later).

As a preface to the research synthesis on determinants of health and the discussion of the challenges it poses for researchers and policy-makers, the taxonomy sensitizes us to the existence of choices in the production of health. There are different investment strategies available to societies - cleaner air, safer cities, better income support, more health care, and improved education, to name but a few of the more obvious examples - each with a different profile of costs and consequences. If the cost-effective production of health (cell 3) is a policy goal, then the breadth and complexity of this potential investment portfolio must be understood and appreciated. And it is indeed both broad and complex.
Conceptual Frameworks and Anomalous Findings

Conceptual frameworks matter. How a problem or task is framed is important. It affects judgements about which types of evidence are considered to be relevant, and therefore given legitimacy and weight, and which are discarded or ignored. It also shapes and limits the sets of admissible questions and debates for that subject. The grossly oversimplified framework of Figure 1 is an example. As the foundation for the production, assembly and synthesis of evidence on the determinants of health, it is clearly inadequate. As a guide for policy, it is exceedingly narrow and seriously incomplete.

In Canada, the recognition of this appeared in the form of a 1974 federal government White Paper, A New Perspective on the Health of Canadians, popularly known as the Lalonde Report (Canada, 1974). Widely cited and acclaimed internationally, the report introduced a broad conceptual framework (Figure 4) consisting of four health “fields” for categorizing the determinants of health. The addition of the fields Lifestyle, Environment and Human Biology to that of Health Care was intended to make explicit some of the “other factors” in Figure 1, thereby raising their public and policy profiles and providing a counter-weight to the view that continuing expansion of the health care system offered the best avenue for further improvement in the health of Canadians.

In the words of the report, lifestyle referred to “the aggregation of decisions by individuals which affect their health and over which they more or less have control” (p.32). Environment was defined to be “all those matters related to health which are external to the human body and over which the individual has little or no control” (p.32), and Human Biology “all of those aspects of health, both physical and mental, which are developed within the human body as a consequence of the basic biology of man and the organic make-up of the individual “(p.31, emphasis in original).

The report also examined the major causes of death in Canada in 1971 by age and gender groups and the distribution of health expenditures from 1963-1973 in relation to the four fields. It concluded that although most of the burden of ill-health emanated from the fields of Lifestyle, Environment and Human Biology, most of the expenditures for health went through the health care system, an imbalance which it suggested should be addressed.
The report was received very positively, but the breadth of policy response did not match the breadth of the framework (Marmor, Barer and Evans, 1994). With the exception of aggressive anti-smoking policies limiting advertising and marketing by tobacco companies and restricting smoking in public areas, the thrust of most of the responses focused on the health education of individuals, to enable them to choose healthier lifestyles.

The focus on individuals, and the “choices” under their “control” also extended into research, spawning an era of detailed analysis of individual risk factors, and often led back to the health care system, where physicians and other health care professionals assumed roles as the primary agents for assessment and modification of those risk factors (Evans and Stoddart, 1990). The heavy emphasis on lifestyles, combined with the view that they originated with and were controlled by individuals, obscured the significance of processes that might be operating at the level of groups or populations (Buck, 1985). In its most naive form, this view suggested that people made unhealthy lifestyle choices largely because of ignorance. The possibility that even informed people were conditioned by their surroundings into unhealthy choices, or lacked the resources to change them, or found them to be their most effective coping mechanisms, was, though not ignored entirely, pushed into the deep background.

On the contrary, it is precisely the focus on groups or populations rather than individuals that is one of the distinguishing characteristics of the work of the Population Health Program

“Systematic differences between the average health status of people in different regions, occupations, time periods, educational levels, and social classes, for example, contain important information about the determinants of health that is often not apparent, or ignored, when one looks only at individuals.”

(Evans, Barer and Marmor, 1994, preface, p.xiii)

These differences are sometimes referred to as “health inequalities”, although a potentially less normatively charged word for them, “heterogeneities”, has been suggested by Hertzman, Frank and Evans (1994). The publications generated by the Program have assembled and highlighted findings on heterogeneities in the
health status of groups and the clues about determinants of health which the heterogeneities provide. Another distinguishing characteristic has been the way in which the work juxtaposes observations and studies from a wide range of disciplines:

“The questions raised by this concern take one from the society wide to the subcellular level and back again, from economic and social policy to molecular biology. The individual person represents the point of contact between cellular and social aggregates. Determinants of health must ultimately show their effects on particular individuals, but their origins may be well ‘above’ or ‘below’ the individual level - mass unemployment, say, or genetic predisposition.”

(Evans, Barer and Marmor, 1994, preface, p.xi)

The picture that emerges does not constitute (nor was it intended to) a complete synthesis of the literature on determinants. Furthermore, the meaning of the findings assembled is not fully understood. They do, however, appear to all be pieces from the same huge jigsaw puzzle. Moreover, many of them are difficult or impossible to represent within the still relatively simple conceptual framework of Figure 4. They do not seem to fit, and seem to require a much broader, but also a much more subtle and complex, framework for coherence.

Several of the most interesting of these anomalous findings are identified below, in a brief review of some of the main themes of the population health approach to determinants. Although attention is drawn to single studies, it should be noted at the outset that they are often representative of, or doorways into, rich bodies of literature on their topics.
a. The Role, Effectiveness and Limits of Health Care

Although timely and effective health care may be an important determinant of how and why individuals recover from injury or disease, it is not in general a key determinant of how and why they become sick or injured to begin with. This does not diminish the many accomplishments of modern health care, nor the dedicated efforts of its practitioners. Indeed, in individual cases, the availability and quality of health care may be - often is - decisive. But it does limit the scope of their contribution to population health.

It is also difficult to explain observed differences in the health of populations, both over time and across comparable national or sub-national jurisdictions, in terms of differential access to or utilization of health care. Examples of the latter are discussed within a larger context in sub-sections b. and c. below. The best known example of the former is the work of McKeown (1976a, 1976b), who examined reductions in mortality in the U.K. during the past two centuries from the principal infectious diseases, showing that the dramatic declines in mortality largely occurred prior to the development of effective medical therapy (Figure 5). Although the specific causal pathways are unknown and the subject of debate (Reeves, 1985; Szreter, 1988; Mackenbach, 1994), McKeown attributed the secular decline to the increased prosperity and rise in the general standard of living resulting from modern economic growth. He attached particular importance to improvements in nutrition as one pathway, an importance which recent work by Fogel (1994) and others reinforces (see Frank and Mustard, 1994).

While debate about the contribution of health care continues (McKinlay, McKinlay and Beaglehole, 1989; Bunker, Frazier and Mosteller, 1994), both “older” observations like McKeown’s and “newer” ones by investigators such as Marmot (1992) caution against over-emphasizing the role of health care. Marmot points out that over half of the recent large decline in coronary heart disease mortality in several developed countries seems to be due to lower rates of sudden unexpected death in previously well persons and therefore not attributable to medical care.

The patterns of provision of health care, and their effectiveness and appropriateness, have also been subjected to increasing scrutiny in the past two decades, with often sobering results. Investigations of the style and intensity of medical practice for similar populations within relatively small geographical areas fre-
quently and consistently uncover wide variations that are apparently unrelated to either patients’ needs or clinical outcomes (Roos and Roos, 1994; Ham 1988). Furthermore, it is not uncommon in the clinical epidemiology and health services research literatures to find rates of ineffective or inappropriate provision of specific services of 20%, or higher (Brook and Vaiana, 1989; Dixon, 1990; Lomas, 1990; Lavis and Anderson, 1993).

b. National Wealth and Health

It is not necessary to go back to the early 19th century to find examples of the association between prosperity and the health of populations. Striking examples are currently unfolding, in quite different directions, in Eastern Europe and Japan.

In Eastern European countries such as Poland, Hungary and Bulgaria, steady improvements in life expectancy since World War II have ceased and life expectancies have begun to decline (in contrast to Western Europe where improvements continue) during the past two decades as the economies of these countries have faltered (Jozan, 1990; Hertzman and Ayers, 1993). Although the pattern of specific characteristics varies from country to country, poor national economic performance has been accompanied by a deterioration in the quality of the social environment for families and workers, pollution of the physical environment, and a general demoralization of the population.

Russia provides an even more dramatic example. Life expectancy for men has fallen from 65 years in 1987 to 59 years in 1993 as economic and social systems have been in turmoil. Death rates are increasing for both children and adults as increases occur in the rates of infectious disease as well as accidents, poisonings, suicides and violence (Hockstader, 1994).

In Japan, meanwhile, there has been a remarkable growth in life expectancy over the post World War II period such that the Japanese have caught and surpassed the rest of the developed world and are now widening the gap. Life expectancy at birth for Japanese males was 63.6 years in 1955; by 1991 it was 76.1 years. For Japanese women, the comparable figures are 67.8 and 82.1 years (Marmot and Smith, 1989; OECD, 1994a).
The gains in life expectancy are occurring across the age spectrum, with the world’s lowest infant mortality rate and one of the lowest perinatal mortality rates, and extensions of life for the elderly. The size and significance of the changes in Japanese life expectancy in approximately three decades can be placed in perspective by the observation that for the United Kingdom to achieve the same change in life expectancy at birth would now require the elimination of all deaths from cardiovascular disease and most of those from cancer (Marmot and Smith, 1989; OECD 1994a; Miyaji and Lock, 1994).

That something significant has happened in Japan seems clear, but the underlying factors are much less clear. The quality of urban physical environments is generally considered to be poor, and the popular impression is that Japanese life is crowded and stressful. Offsetting this may be unique features associated with diet or social structure, although it is not apparent that these have shifted markedly over the past thirty years (Marmot and Smith, 1989; Evans and Stoddart, 1990). Moreover, the gains in life expectancy do not appear to be attributable to the health care system, which is not given high ratings internationally, and absorbs one of the lowest shares of national resources among developed countries (6.9% of GDP in 1992, compared to “leaders” Canada, at 10.1%, and the United States, at 14.0%) (OECD, 1994a). There is, however, a focus on maternal and child health, and motherhood and family bonds remain central tenets of Japan’s cultural and social ideology, despite the obligations and restraints they impose on Japanese women (Miyagi and Lock, 1994).

An obvious phenomenon paralleling the Japanese life expectancy gain has been the strong performance of the Japanese economy, which has displayed the fastest rate of growth among OECD countries for significant parts of the post World War II period. A less well known but potentially more significant observation is that Japan has one of the narrowest income distributions in the developed world, and that a substantial part of the narrowing has occurred in the past twenty years (Marmot and Smith, 1989; Wilkinson, 1994a,b).

Wilkinson (1994a,b) presents both cross-section and longitudinal data which clearly illustrate that among developed countries it is not the absolute level of income but the degree of equality in the distribution of income which is most closely associated with life expectancy. In the past twenty years, for example, the fastest
increases in life expectancy have occurred in countries where income differences in society have narrowed. The idea that past some threshold of economic development the health of a population appears to be influenced more by the distribution of income than by its overall level or growth leads Wilkinson to conclude that social factors have replaced material factors as the predominant determinants of health, and that the psychosocial effects associated with relative income in the social hierarchy are somehow critical to health.

c. Socio-economic Gradients in Health

There exists an extensive body of literature documenting significant differences across social classes in morbidity and mortality (Hertzman, Frank and Evans 1994; Pincus and Callahan, 1995; Bunker, Gomby and Kehrer, 1989; Feinstein, 1993). Higher socio-economic status, however measured (often by income, but also by education or occupation), correlates positively with good health. What is most significant, however, is that this relationship exhibits a gradient: groups at every rung of the socio-economic ladder are healthier than those at the rung immediately below them, throughout the entire range of the socio-economic hierarchy. This suggests that whatever processes are at work to determine health are operating at all levels of society, and that the pathway to poor health is not just one of material deprivation arising from poverty.

Moreover, these socio-economic gradients are common to all developed nations, persist over time, and appear to be unaffected by changes in the leading causes of death in societies. They also do not appear to be very sensitive to the provision of health care. They therefore constitute a major finding in the heterogeneity literature, and suggest that income and social status may be markers for more fundamental processes that occur in all societies.

Two powerful examples of socio-economic gradients are the analysis of British data on mortality rates by social class in the Black report Inequalities in Health (Black, Morris, Smith and Townsend, 1982; OPCS, 1972, 1978) and the analysis of health differences by rank among British civil servants by Marmot (1986; additional references in Evans, Barer and Marmor, 1994).

The data on mortality rates by social class display a clear and persistent health gradient from upper to lower classes, extending through most of this cen-
tury, and becoming more accentuated recently, despite significantly improved accessibility to health care for the lower classes through the introduction of the British National Health Service after World War II. As Hertzman, Frank and Evans (1994) point out, “The observation that the mortality gradient did not respond to changes in medical care is the obverse of McKeown’s findings, that the dramatic declines in deaths......occurred in the absence of any effective medical therapy” (p.69, emphasis in original).

Marmot’s Whitehall Study has followed more than 10,000 British civil servants, within a well-defined hierarchy of income and rank, for two decades. The data (Figure 6) show a definite gradient in age-standardized mortality for men aged 40-64 years. Mortality was lowest in the highest grade (administrative) of the civil service, and rose steadily in the successively lower “professional/executive”, “clerical”, and “other” grades. The gradient was not explicable by conventional individual risk factors and lifestyle choices. For example, cholesterol, blood pressure and smoking, the primary risk factors for coronary heart disease, accounted for only about one-quarter of the gradient in that disease.

The Whitehall results are particularly important because they show that something powerfully associated with hierarchy itself influences health. Although differentiated by rank, none of the study subjects would be considered “poor” in an absolute or material sense. All had decent access to the basics of life, including good health care. This strong association of hierarchy with health recalls Wilkinson (1994a,b), who also reported that countries which narrowed their income differences across social classes over time were more likely to narrow their mortality differences across social classes as well.

The gradients literature has important implications for our views of both disease and individual behaviour. It suggests that a disease-specific approach to understanding heterogeneity in health (and perhaps health care intervention) is inadequate. If one disease is conquered or disappears, the general socioeconomic gradient in vulnerability will express itself through other diseases. And even if entire populations led “healthier lifestyles” (which might continue the steady secular improvement in life expectancy in most countries) it appears that socio-economic gradients in health would remain (as they have during previous improvements).
The gradients literature is also a signal that something quite complex is taking place inside production functions for health. Whatever the specific pathways might be, they appear to have less to do with income itself, for example, than with the social situation which income provides or marks. Gradients direct attention to patterns of interpersonal relations in social environments and to a variety of “non-economic” variables and phenomena including, for example, self-esteem, stress and the skills and resources to cope with it, control over one’s life, and social support (Marmot and Mustard, 1994; Marmot, 1994; Hertzman, Frank and Evans, 1994).

d. The Social Environment

The structure of social environments - how people work and play, how they care about, support and condition others, and how they nurture and teach successive generations to develop competencies for living - is an important determinant of health. Social environments can moderate or accentuate the relationship between status and mortality through psycho-social effects (eg. depression) as well as material effects (eg. nutrition). High quality social environments can reduce vulnerability to ill health at all stages of the life cycle (Hertzman, Frank and Evans, 1994).

The work environment is one critical component of the social environment. Karasek and Theorell (1990) have shown the importance of the psycho-social nature of work, for example. In a study of 1,600 randomly selected Swedish working men they found that symptoms of coronary heart disease were significantly more common among workers whose jobs were characterized by a high degree of psychological demands and a low degree of control, latitude for decision-making, and discretion over use of skills. This links to the Whitehall Study in which civil servants in the lower grades reported feeling that they had less personal control over their work than those in the higher grades (Marmot and Theorell, 1988).

Employment itself, or more accurately the lack of it, is strongly associated with health. Iverson et al. (1987) found Danish mortality rates up to 50% higher among the unemployed than the employed, although interestingly the relative death rate among the unemployed was lower in areas of the country with higher overall unemployment rates. A World Health Organization review of international research on unemployment and health concluded that:
“high levels of unemployment and economic instability cause a significant increase in the levels of mental ill health and also have adverse effects on the physical health, not only of the unemployed but also of their families and the community in general.”

(Westcott et al., 1985, p.3)

Other studies report that those unemployed or insecurely employed feel more stressed, engage in more health-detrimental behaviour (e.g. weight gain, smoking), experience more disability days, and contact the health care system more frequently than their employed or more securely employed peers (Ontario, 1991).

The nurturing of infants and young children, and the structure of environments affecting early childhood experiences, are other critical components of the overall social environment. High socio-economic status can be a powerful buffer against very different kinds of threats to successful human development (Hertzman, 1994a,b). For example, in the Kauai Longitudinal Study (Werner and Smith, 1982, 1989; Werner 1989) children who suffered moderate or severe perinatal stress (which can result in long term damage) but who were raised in stable or high socio-economic status family environments suffered little or no developmental disadvantage as measured at 20 months compared to their counterparts raised in unstable or low status environments. And in a study of an early childhood enrichment program for children in high-risk inner-city families in the U.S., those receiving the intervention showed significantly better outcomes than the control group at age nineteen on a range of measures such as mental retardation, teenage pregnancy, arrest, receipt of welfare, and employment (Schweinhart et al., 1985). Childhood clearly contains sensitive periods of vulnerability and development, and the effects of experiences during those periods can and do have lifelong consequences (Hertzman, 1994a,b; Hertzman, Frank and Evans 1994).

The importance of social environment can be demonstrated throughout the life cycle. The existence of social support networks for adults has been studied extensively. House, Landis and Umberson (1988) synthesized a number of these international studies (Figure 7) showing that the relative risk of mortality declines as the degree of social support increases. The strength of this relationship often rivals that of other well-established individual and lifestyle risk factors, although as Corin (1994) points out, the specific nature of the social network and the
cultural content are critical and have the potential to reverse the buffering effects of social contact.

Against this setting of the importance of the social environment for health, it is not difficult to grasp the significance of growing concerns regarding heightened economic insecurity, fissures in social cohesion, stagnant employment growth, increasing income inequality, and arising number of economically vulnerable family structures documented in the recent OECD study _New Orientations for Social Policy_ (1994b). Their long term health consequences are potentially quite severe.

e. The Role of Genetics

The heavy emphasis on the role of both “macroenvironmental” and microenvironmental phenomena here is not meant to exclude the potential significance of individual factors, such as personality or genetics. Indeed, in specific contexts and for specific individuals they may be critical (Werner and Smith 1982, 1989). In the case of genetics, however, although it is likely underrated as an historical determinant of health, it is unlikely that it is a primary determinant of heterogeneity in the health of modern populations (Mustard and Frank, 1991; Baird, 1994; Wilkinson, 1994a). As Rose (1985) points out, there is an important distinction to be made between the causes for sick individuals and for sick populations.

In addition to the observation of significant changes in the health of entire populations (eg. Japan) over time spans too short to encapsulate changes in the genetic endowment, there is also evidence from studies of migrating populations that people who change countries exhibit within one or two generations the disease patterns of the populations in their new environments (Marmot et al., 1975; Marmot and Syme, 1976). And studies of the health of identical twins, although often showing a high concordance of diseases and disorders, do not show complete concordance, thereby indicating the presence of other determinants (Baird, 1994).

Although genetic inheritance obviously plays a role in many diseases, the role of genetics is far more complex than the notion of a fixed genetic endowment leading directly to disease that is the basis for the “Human Biology” box in Figure 4. While genetic endowment may be fixed at birth, its ultimate effects are not.
Genes set the limits to the responses of individuals to their environments and experiences, but endowment is not the same as expression. Increasingly, genetic knowledge emphasizes the role of genetic “predispositions” or susceptibilities; whether or not they result in disease or disorders depends critically on the environmental influences and challenges experienced by individuals over their lifetimes. Baird (1994, p.143, emphasis in original) notes, “The important point is that genes determine who may get sick within a class, but environmental factors determine the frequency of sickness among susceptibles.” Genetic and environmental factors therefore represent complementary and interacting rather than competing influences as explanations of health differences among groups.

f. Physical Environments

Structures or conditions which can be physically altered, whether natural or made by humans, constitutes another important category of determinants of health. This includes a wide array of factors, ranging from the quality of water, soil and air (both indoor and outdoor) to workplace hazards and other dimensions of everyday life such as the design of vehicles, roads and communities for personal safety and security. Motor vehicle accidents and workplace injuries and illness in particular are two of the leading causes of death and disability in modern societies.

As with genetics, frequently it is the interaction of influences in this category with aspects of the social environment (or with genetic factors themselves) that produces adverse health outcomes, although there are relatively “direct” influences as well.

Hertzman (1990) reports on the dramatic example of industrial pollution in Poland, where five areas of the country have been designated as areas of ecological disaster and show evidence of accompanying poor health status of their residents. The region of Katowice, for example, although relatively prosperous by economic standards, exhibited both one of the highest infant mortality rates in the country and extensive heavy metal contamination of soil, in part due to practices of industrial waste disposal. Moreover, local health outcomes appeared to be disproportionate to the extent of the contamination, suggesting at least the possibility that the local residents’ perception of environmental quality and the difficulties of rectifying it might also be associated with poorer health. Mustard and
Frank (1991) and Eyles et al. (1993) draw attention to the potential significance of this for existing and planned industrial waste sites in North America.

g. Biologic Pathways

Death, coronary heart disease, and other adverse health outcomes in the literatures scanned above are ultimately biologic events, and it is therefore difficult to make the causal connections between them and the social, cultural, economic and physical environments in which they occur, even acknowledging the genetic component involved. Some of the biologic pathways seem clear; malnutrition caused by material deprivation is one example. But most are not.

One of the distinguishing features of the work of the Population Health Program and the groups of investigators it has at various times brought together is a concern for the potential biologic pathways that might constitute the linkage. Such pathways are only just beginning to be understood, at this stage with considerable insights from animal rather than human populations, but the emerging biological picture is at least as complex and diverse as that of the environmental influences themselves. There appear to be numerous potential pathways, some acting over relatively short periods, others acting progressively and cumulatively, and still others acting over long periods but characterized by significant periods of latency.

It is not possible here (or for this author) to give but a cursory sketch of the emerging picture and its speculative possibilities. More detailed discussion may be found in Evans, Hodge and Pless (1994), Frank and Mustard (1994), Hertzman, Evans and Frank (1994), Hertzman (1994a,b) and Cynader (1994a,b).

The main message is that the pathways begin with sensory stimuli from the external environment which are interpreted by the brain and directly transmitted to other biological systems, such as (but probably not limited to) the immune, endocrine or cardiovascular system. The sights, sounds, smells, tastes and, perhaps most importantly, the feelings (such as fear, loss of control or hope, or other forms of stress) produced by one’s environment and experiences induce physiologic responses within the body. In other words, perceptions, feelings, and stimulation are biologic events as well as social ones. A passage from Wilkinson (1994a) about the role of hierarchy seems especially pertinent in this regard:
“Talking to people who suffer relative poverty, what is striking is the sense of stress and hopelessness, the high rates of depression, desperation and anxiety. Dominating people’s whole consciousness, feelings such as these have a much more devastating impact on the quality of people’s lives than the presence or absence of various consumer durables does in itself. What is almost inescapable in modern society are the powerful psychosocial connotations of material differences.” (p.198)

Cynader, for example, points out that during sensitive periods of brain development, the uses to which particular regions of the brain are put determines how well it functions for the rest of life. He also notes that in addition to causing problems in body systems controlled by the brain, chronic stress has deleterious effects on the brain itself. Frank and Mustard emphasize that the quality of nourishment and nurturance given to infants and young children strongly influences the development of the brain; for example, the development of the visual cortex is very sensitive to appropriate stimulation. The social environments of families can therefore have important consequence for the development of children and their later functioning as adults. In particular, failure to develop basic competencies and coping skills early, (although not entirely impossible to achieve later, though more difficult) can play an important role in negative health and life trajectories.

Work on the interaction of the nervous system with the immune and endocrine systems also offers clues to potentially important biologic pathways. A reduction in immune status associated with bereavement, for example, is a well known phenomenon and may accompany other forms of stress as well. It may be the case, then, that pathways whereby factors such as unemployment, lack of control over work environments, or inadequate social support lead to adverse health consequences involve a diminished host-defence response to stress, or the inability to “turn off” undesirable responses to stress.

Identification and understanding of biologic pathways is only at a very early stage. But the potential linkages are there and advancing research in the basic sciences is rapidly making several specific ones more concrete. As Evans (1994a, p.34, emphasis in original) concludes “.......the pathways are not clear. But it is clear that such pathways exist. The connections between health and social environment, broadly defined, are very real and may be very powerful.”
h. Animal Studies

As mentioned at the outset, another distinguishing feature of the work of the Population Health Program is the breadth of evidence it seeks to bring to bear on the subject of why some people are healthy and others are not. This quest has led to an awareness of another very different family of literature which describes and analyzes the processes of health and development in the non-human animal world. Here too, the social environment appears to be a powerful determinant of health, and there may be important clues about the biologic pathways linking the two. As in the case of human populations, status hierarchies are a key element of the social environment of certain animals; often their association with health effects offers striking parallels to results in human populations (Evans 1994b).

Again, limitations on space and the trans-disciplinary ability of this economist-author confine the treatment of the subject here. Nevertheless, it seems important to attempt to provide some idea of the flavour of the evidence available.

Biological effects from the psycho-social environment have been demonstrated in a number of studies. For example, in one set of experiments rabbits fed a cholesterol-rich diet but which had music played to them had a 60% lower incidence of atherosclerosis than those given standard laboratory treatment (Nerem, Levesque and Cornhill, 1980). In another, a moderate cholesterol diet was found to produce four times the degree of occlusion of the coronary arteries in low versus high status monkeys (Hamm et al., 1983). This recalls Marmot’s finding in the Whitehall Study that the increased risk of heart disease among the lower grades of the civil service could not be adequately accounted for by conventional risk factors.

In another study, this time of monkeys randomly assigned to stable or unstable living conditions (housing companions were changed monthly), Cohen et al. (1992) found significantly depressed immune responses among those assigned to the unstable conditions. In numerous other studies of different animals, it has consistently been found that experimentally generated, unavoidable stress over prolonged periods can cause significant morbidity and mortality (Dantzer and Kelley, 1989).
Sapolsky’s studies (1990,1993) of free-ranging baboons in Kenya are of special interest because there is a well defined status hierarchy among baboon troops and the primary sources of stress are psychological, since food is plentiful. He found systematic physiologic differences in endocrine system function between dominant and subordinate males. Subordinates did not appearable to “turn off” their stress responses after the stressful situation passed, but were constantly in a state of low-level readiness, thereby subjecting their bodies to increased physiologic strain; the hormonal patterns accompanying such strain are associated with a number of diseases. The observation that the animals who coped best with the stresses of everyday life and were able to shut off their stress response appropriately had much lower blood cortisol levels recalls Marmot’s finding that while both higher and lower grade civil servants had elevated blood pressures while at work, that of the higher grade individuals dropped much more when they went home.

Sapolsky’s observations of disruption in hierarchy suggested that the optimal hormonal profile found in dominants followed from rank, and not vice-versa. Social rank protects. But there was also much subtlety in the results; this protection depended also on the stability of the society, the specific individual experiences producing the stress, and the personality of the animal. Not all of the dominant animals displayed the optimal hormonal profile, and those who did also displayed a well-defined set of personality traits and behaviours.

A final example comes from the work of Suomi (1991), who investigated the effects of the quality of parenting and nurturing on infant monkeys bred to be genetically resistant to or susceptible to stress. The resistant groups fared well regardless of the quality of parenting; however, in the susceptible group the quality of parenting made a significant difference in the infants’ physiological and behavioural responses to stress. Genetic vulnerability appeared to be socially buffered; predisposition is not the same as expression.

Transference from animal to human populations is of course far from automatic. These and other studies of animal populations, often involving study designs that would be impossible in human populations, are very interesting, and potentially very significant, for they demonstrate the biological respectability of hypotheses about the social determinants of health.
i. Subtlety, Complexity and Interactions in the Determinants of Health

The picture, incomplete though it is, that emerges from the wide range of observations and research findings highlighted above emphasizes that the underlying causal structure of the determinants of health is much more complex and subtle than many popular, almost unidimensional views imply. The answer to the question “What matters most” is both “nothing” and “everything”, or more accurately, “it depends”. No one factor is paramount, and simplistic interpretations to the contrary are likely to be both seriously incomplete and misleading as guides to policy-making, whether public or private.

Moreover, the phrase “determinants of health” itself may create a false impression of a cluster, even if it is a large one, of single factors acting alone. Perhaps “patterns of determinants” would be a more appropriate language to employ to indicate that it is the interactions among factors which seem most important. Yet even this adjustment of terminology does not capture the full range of complexity, because the key patterns, although they exhibit some regularity, appear to show important variations or exceptions from one context - different times, places, populations and sub-populations, stages of life cycle - to another. Patterns of factors (or policies) that appear to “make a difference” in one context may not, or may do so with a different relativemagnitude of impact, in another. There are always exceptions and discrepant observations.

Several examples of this occur in the literatures scanned above. Unemployment is bad for an individual’s health, but it is not as bad if many people are unemployed in the community (Iverson et al., 1987). Good nurturing in stable environments is very important, but some “vulnerable but invincible” children thrive, despite its absence (Werner and Smith, 1982, 1989). The psycho-social structure of workplaces, with more control and opportunity for a sense of belonging and personal contribution, appears influential, and Japan may be an example of this; but Japanese women show the same large and rapid gains in life expectancy as Japanese men (Marmot and Mustard, 1994). Some of the widespread social gradient in health is undoubtedly due to individuals who are “fitter” - either through inherited or learned traits - rising to the top, but there is a sufficient literature to the contrary (Evans, Barer and Marmor, 1994) as to render this only a partial explanation. And what accounts for that portion of success in both health and life that come from differential learning? The list goes on, and is long.
In particular, competition between the simple “behavioural” view that individuals choose unhealthy lifestyles and the simple “biological” view that some individuals are blessed with “strong” genes while others are not does not seem useful. (Nor, incidentally, does the sometimes popular and equally simple view that “more health care” will solve the problem.) It is the interaction of behaviour and biology which appears important. But more important it seems are the environments, both large and small, in which the interaction takes place. The social, cultural, economic and physical circumstances under which people live, work, and play appear to condition and modify the effects of individual behaviour and biology in critical ways that are only beginning to be appreciated.

Evans (1994c) uses an engineering analogy - the ability of a beam to support a load - to draw together some of the most important categories of determinants and illustrate their interaction. This analogy may be developed further by thinking of the engineering application of a bridge, and why bridges don’t fall down. The ability of a bridge to withstand the internal strain from external stress depends upon both the amount of stress and the strength and resilience of the bridge itself. The former depends upon the intensity, frequency and duration of the stress. The latter depends upon the characteristics of the materials themselves from which the bridge is constructed; the way in which these materials are treated, tempered or reinforced before or during construction; the design of the bridge and in particular the number and type of supports it has; and the way in which the bridge, once built, is maintained and repaired.

The same factors may apply to humans. External stresses to the organism vary in intensity, frequency, and duration, and although some strain is good, even necessary (as it is for bridges to work), excessive amounts can be overwhelming. How individuals respond to stress and strain depends not just on the amount of it, however, but also on their nature (genetic endowment, personality, and other characteristics of their individual “material”), and on how their “material” is handled in early life through nurturing and nourishment, how it is supported by families, friends, and social institutions throughout the life cycle, and how well it is maintained and repaired when necessary by health care.

The potential significance of hierarchy can also be illustrated in this context. Those higher up the socio-economic ladder may experience less stress (or less
“harmful” stress) or may possess more resources, both financial and personal, inherited or acquired, with which to avoid or cope with the stress they do encounter. One way or another their position buffers them from the threats of daily life compared to those at progressively lower ranges. The typical components of socio-economic status - income, education, and occupation - are therefore indicators of exposure to stress or ability to respond to stress, or both.

j. Broader Frameworks ...... and a Tightrope

In order to accommodate the array of evidence above, a broader and richer conceptual framework than that of Figure 4 is required; one such framework is introduced in Figure 8, drawn from Evans and Stoddart (1990). As mentioned earlier, the recognition of the need for such a broad framework is hardly new, although an understanding of just how broad a range of evidence needs to be accommodated has only occurred more recently. The public health, community development, health promotion and sustainable development literatures implicitly or explicitly employ the same components, or at least large segments of them, and the World Health Organization and its officials and associates have actively promoted a broad view of the determinants of health (Marmor, Barer and Evans, 1994; WHO 1984; WHO et al., 1986; Epp, 1986). Gunning-Schepers and Hagen (1987), Hancock (1986) and Hurowitz (1993) provide other examples of conceptual frameworks for the task.

Nor is the expanded framework of Figure 8 wholly adequate for such a daunting task as the representation of the subtlety, complexity, and interactions discovered above. The separation and presentation of categories as distinct boxes masks the blurring of boundaries in practice. Like Figure 4 before it, it is too neat, too tidy for reality. Perhaps a better graphic would depict the centre of each box as a different colour, bright and pure at its origin, then fading in intensity as it progressively diffuses in all directions to the perimeter, thereby creating a melange of interesting hues among the boxes in the process. The boxes themselves are best viewed as not homogeneous entities, but a rich mixture of influences. Perhaps they should be considered to be doorways, opening into additional worlds with their own complex relationships and structures.

The “causal” arrows, too, remain an oversimplification of the rich structure of causal pathways and sub-pathways involved, although they are intended to
portray some of the channels of principal interest. Perhaps a cross-section of a thick wire cable, exposing hundreds of individual, finer, tightly-bound filaments is a better image. Finally, the framework of Figure 8 does not communicate the important temporal aspects of latency and the embedding of early influences, or the accumulation of experiences in life trajectories. It appears static, although the processes within it are definitely dynamic.

Nevertheless, it extends the conceptual fields of Figure 4 in several important ways. First, the insertion of an element for Individual Response allows for the operation of an individual “host-defence” mechanism which encompasses the interaction of behaviour and biology which was absent from the earlier framework, but which seems to be critical for the pathways (whatever their exact nature) from social environment to disease and ill health. The Individual Response box includes some factors or processes in the earlier boxes of both Lifestyle and Human Biology and allows, for example, the possibility for lifestyle choices previously considered to be under the “control” of the individual to be represented as socially conditioned. The widely observed social gradient in smoking behaviour finds a place here, as do the observations of the importance of concepts such as competencies and coping skills, and feelings of self-esteem, hope and control. Again, though the box may be too abstract and imprecise for the task asked of it, the representation of this nexus at least serves as a reminder of its importance.

Second, the unbundling of the earlier labels of Environment and Human Biology into Social Environment, Physical Environment and Genetic Endowment allows for important changes of emphasis. The key role played by the myriad factors comprising the social environment, including that of socio-economic hierarchy, becomes a focus of attention in itself, and the role of genetic predispositions which may or may not find expression in diseases or disorders depending upon interactions with the social and physical environments is highlighted, in addition to those genetic factors expressed at birth.

Third, although it is not the subject of this paper and will not receive detailed comment (see Evans and Stoddart 1990), distinctions are drawn between Disease, Health and Function, and Well-Being as concepts of health. The same “disease”, as recognized and responded to by professional experts in the health care system, may have very different consequences for “health and function” (
including economic productivity) for different individuals from their own viewpoint (or those of their families) depending upon their individual social, cultural and economic circumstances. These health-related differences may in turn affect their sense of “well-being” or life satisfaction, and their ability to achieve their aspirations.

Fourth, the extended framework places the activities (and resource demands) of the health care system in a more balanced perspective. Health Care remains, but it is shifted from centre-stage as the primary determinant of the health of populations. The health care system is thereby seen as but one of many social institutions involved in the production of health - one investment vehicle and strategy amidst a larger social policy portfolio. This shift well illustrates the earlier point that frameworks count. The extended framework opens up different questions and concerns, and highlights the existence of both options and tradeoffs in health policy.

For example, health care obviously contributes directly to the well-being and economic progress of society by reducing disease, improving health and function, and contributing to the quality and productivity of a nation’s human capital. Less obviously, however, it simultaneously draws large amounts of national resources away from other activities which improve health. These activities may be ones outside the health care sector which nevertheless have health as their explicit objective (such as industrial safety programmes or pollution control) but they may also be ones which, although not undertaken for health purposes, nevertheless have important health consequences (eg. education, day care, pensions) (Lavis and Stoddart, 1994; Evans, 1994a). These opportunity costs simply do not appear in the narrower determinants frameworks, although they figure prominently in cell 3 (at least) of the efficiency framework of Figure 3.

Finally, the pervasive macro-influences (both positive and negative) of prosperity and economic growth on the social and physical environments, health and well-being can be accommodated and are highlighted in the extended framework. And, again at this most general level, there are tradeoffs and feedback loops, this time between investments in economic growth itself and improvements in social environments. A recurrent theme of the commentators on the determinants of health is the critical importance of the quality of the social environment. Yet the
capacity of a society to create and sustain a high quality social environment is not unlimited. It is bounded at some point, even with borrowing, by a national “ability (as well as willingness) to pay”. Failure to make investments in the capacities of economies to grow and adapt limits their ability to maintain the institutions, with or without health objectives, that produce health (Frank and Mustard, 1994; Mustard 1994). Nations, like individuals, must live within their means. It is possible to over-invest in the social environment.

Conversely, under-investment in the social environment also handicaps a society. Failure to nurture succeeding generations adequately, to develop competencies and coping skills, and to support workers and families both under “normal” circumstances and during periods of vulnerability or of difficult social transition (such as the current one to information and innovation based economies) not only undermines the capacity for future prosperity but also results in the diversion of increasing amounts of resources to support marginalized groups. In the context of health and well-being it is a “macroeconomic high-wire act” of enormous responsibility and significance, along a tightrope which extends into the infinite future, even if there may be occasional rest platforms along the way.

The dexterity and balance required is succinctly stated by Frank and Mustard (1994, p.14), for Canada although it might equally well be applied to many other developed countries:

“Canada faces the challenge of trying to sustain its social systems, including health care, social support, and education with diminished resources while simultaneously trying to rebuild the economy.”

The ways in which the overall wealth of a society may affect the health of its citizens and reduce social gradients in health go beyond the direct effects and the level of prosperity per se, as Evans (1994b) notes:

“First of all, affluent societies maintain institutions to protect their members against external threats, supplementing the resources of individuals and families. Such collective buffering mechanisms are costly, and inevitably require some degree of redistribution of resources.... A second potential channel of influence has to do with the perception of progress. Economic growth, even if it lifts all the boats
together, gives each individual the sense of progress and of hope. Tomorrow will be better than today. This seems to be an important component of psychological well-being, which also translates into physical well-being.” (p.22, emphasis added)

The added emphasis on the distribution of the fruits of growth and prosperity is worth noting. It is a theme emphasized also by Wilkinson (1994a,b) who questions whether growth that accentuates socio-economic differences in society is desirable, and by Sen (1993) and Caldwell (1986) who stress the importance of a society’s commitment to allocate resources to key sectors such as maternal and child health, nutrition and education - a lesson for the developed countries from countries with higher levels of health than their wealth would lead one to expect.

New Understandings Create New Challenges

The work of investigators attempting to synthesize, interpret and link diverse evidence on the determinants of health such as that reported in this paper has begun to provide new understandings about how health is produced in modern economies. These new understandings in turn create new challenges - conceptual, methodologic, empirical, practical and political - for researchers attempting to push out the frontiers of knowledge about the relative importance of various determinants, for policy-makers who wish to act on emerging knowledge about determinants, and for health economists who may be interested in participating in the collaborative, cross-disciplinary efforts required for both research and policy development.

In this section I provide a personal view of some of the more interesting and important challenges. I do not discuss them in detail, although I look forward to the opportunity to do so at the workshop.

a. Policy and Policy-Making

At the most general level, the challenge is to change the “mindset” or mental image that most people, including public decision-makers, have of health policy as either health care policy or at best, policy with health as its explicit objective. Whether it is impractical to think of health policy as any policy with health consequences is a reasonable question; however it seems important to do so conceptually, else regression to narrower frameworks almost unconsciously occurs. As
Marmor, Barer and Evans (1994, p.217) note, although everyone when asked acknowledges that health is determined “broadly”, health policy debates in general and medical care spending debates in particular proceed “as if these broader determinants were irrelevant to decisions about how much should be spent for narrowly defined medical services”. In Canada, the education of legislators and bureaucrats about the emerging “determinants synthesis” has resulted recently in the “adoption” of a broad determinants framework not unlike Figure 8 by the federal government and two provincial governments. Whether this adoption influences policy to any significant degree remains to be seen and is worth monitoring.

At a more operational level, the framework and synthesis challenge the existing institutional structures and processes for policy-making and implementation. How is strategic thinking about public policy relevant to health to be accomplished within governments? By inter-ministerial committees? Many observers and participants view them as ineffective vehicles. By a requirement that a “health impact analysis” be done for all new policy initiatives? Perhaps, but does a credible methodology for such exist (or if not, can one be created), and can it be made administratively feasible? The challenge is to devise new structures that will integrate social and economic policy and extend beyond a single electoral cycle. Can incentives for inter-sectoral cooperation be created, even within public decision-making bodies, for example? In many, perhaps most public bureaucracies, large parts of the knowledge base employed for the determinants synthesis exist but are distributed across numerous departments and agencies each of whom applies the knowledge “vertically” to its policy sector. Creating “horizontality” in the sharing and application of knowledge will not be easy.

It must also be recognized that governments are not the only policy-makers. The policies of private employers - for example, designing healthy workplaces or offering flexible working hours or job sharing to accommodate child care or elder care responsibilities - are especially important. New policy partnerships and coalitions will need to be developed.

Regarding specific policies, the challenges are formidable. Recall that the answers above to the question “What matters most?” in relation to determinants were “nothing”, “everything” and “it depends”. The subtlety, complexity and in-
The interactive nature of the determinants, and the importance of context assure that there are no simple answers, and no single magic bullet. This is clearly recognized in the case of children’s health by a recent OECD study:

“But the multiple problems cannot be addressed by a single solution such as income support. Children may need educational enrichment at home and at school, adequate nourishment and safe living arrangements, as well as a generally supportive environment; a parent or parents may need assistance in improving their earnings capacity and in finding a good job; and both may benefit from access to adequate health care, housing, and appropriate child-care services.”

(OECD, 1994b, p.18)

More problematic, however, is that there is not likely any optimal policy mix that is “scientifically” knowable in advance, to then be filtered through and shaped by public values and preferences. There is not currently, and will likely never be any policy maker’s “fantasy equation”, with health as a dependent variable, a function of a long list of well specified (and behaved) independent variables, each with their own coefficient showing relative magnitude and significance. One wishes there were. But the obstacles to such a policy-maker’s guide seem almost overwhelming. At present we but vaguely understand the relative magnitude of the coefficients on the independent variables that would inform specific policies rather than broad directions, even if we are beginning to see the variables themselves more clearly. The interaction terms are much less clearly specified, however, and it seems fair to say that we do not know the overall functional form. Moreover, the entire specification is suspiciously context-specific, and needs to be dynamic.

It is at this stage possible to indicate some potentially important targets for health policy (as per the new definition) from the determinants literatures - for example, providing people more control, hope, support, and dignity, and compressing socio-economic differences - but with the possible exception of the last of these (at least in some jurisdictions), they are not the traditional types of targets with which policy-makers are familiar. It is also possible to emphasize the need for policies and programs targeted at the level of whole populations or groups rather than individuals. Syme (1994) and Marmot and Winkelstein (1975) make this
point in the context of coronary heart disease by noting that after forty years of intensive and expensive study of both major and minor individual risk factors, it is now known that all of these factors taken together account for only about 40% of the disease.

Specific policies, however, come not from frameworks but from evidence about which interventions work well for different targets and in different groups. At this level of detail, our knowledge of effective interventions appears to lag behind our knowledge of determinants, although some may find my view too pessimistic.

Specific policy choices also depend on the values and preferences of the citizens of democratic states. Priorities will differ across jurisdictions. In Ontario, one of the Canadian provinces which has adopted the determinants framework, the Premier’s Council on Health Strategy (Ontario, 1991) has proposed healthy child development, adult adjustment and labour market adjustment, and the physical environment as its priority areas for policy development. The Council identified several specific initiatives in each case, for example: incentives for business and schools to provide child care facilities; comprehensive prenatal care targeted at high-risk groups; reform of income support programs to give priority to recently separated, divorced or widowed single parents, and to pre-employment training supports, especially for women experiencing family disruptions; legislation to ensure minimum standards for advance notice of business closures and layoffs; tougher enforcement of workplace safety regulations; and comprehensive waste management approaches which provide incentives to business and consumers to prevent or minimize waste.

Regarding health care systems, the main challenge for policy-makers may well be how to remain firm in their resistance to the constant political pressure for expansion inexorably applied by those with a powerful economic interest in health care, including both health care professionals and others who stand to benefit. One important strategy will be to insist on clinical effectiveness as the sine qua non for both current and future spending for health care. Another strategy, for those whose goal truly is cost-containment rather than cost-shifting, will be not to relinquish control over health care expenditures through increasing privatization of health care financing.
Yet other strategies complementary to a “determinants-based” health policy are discussed by Roos and Roos (1994) and Lomas and Contandriopoulos (1994); they reflect a common theme of the last twenty years of health services research and evaluation (and acceptance of it by “official” reports and inquiries in almost all developed countries) that more and better management should precede more money for health care delivery. In the context of Figure 3 earlier, there is substantial scope for efficiency improvements in cells 1 and 2 even before arriving at cell 3.

With respect to both containment of health care spending specifically and more general resource re-allocation to act on the messages from the determinants literatures, policy-makers face a significant challenge in creating a constituency for change among the public. Health care is personal, concrete, and immediate. Broader determinants seem anonymous, abstract and distant. Their role in keeping people healthy may be recognized by the public but, once ill, the public understandably wants health care - now. Reinforcing this tendency, the normal information channels used by the public are dominated by health care rather than health messages. The “voice on the other side”, though not inaudible, is weak. The political challenge of “selling” policy change is considerable.

b. Research and Research Transfer

There is no shortage of specific, sometimes very specific, candidate issues and questions for study, as the “future research” section of any of the studies in any of the literatures above will make quite clear. What I wish to do here, however, is identify several somewhat more general challenges arising from the themes and frameworks in the determinants literatures and offer some closing comments on a few challenges for health economists especially.

More longitudinal studies of populations or sub-groups are required; ideally these would extend across more than one generation (Hertzman, Frank and Evans (1994). Although they are both expensive and difficult to administer, longitudinal studies have provided some of the most important findings in the determinants literature (Marmot, 1986; Werner and Smith, 1982, 1989; Schweinhart et al., 1985), and are essential to the study of both effects with long latencies and those which progress cumulatively.
Even though the complete “fantasy equation” for determinants may be a mirage, important empirical and methodologic tasks remain. Can the defining characteristics of “supportive communities” be identified and their interaction with individual-level variables investigated? Although researchers have examined interactions among determinants, the focus has been on interactions among variables at the level of the individual, or at the level of the community, but seldom between the community and the individual levels. There may be fruitful applications of multi-level modelling possible in the determinants of health field.

Further work on estimation of the effect of disadvantage or failure at previous stages on conditional probabilities of “success” in progressive life transitions, perhaps through Markov models, may be another important exercise which would add to knowledge of “downward” or “upward” life trajectories such as those described by Hertzman (1994a,b) which may extend through eight decades of life. Another potentially fruitful research avenue, given the emerging focus on stress, might attempt to specify models of “net” stress (i.e. overall or “gross” stress adjusted for resources available to cope with it), perhaps by social class and over both short and longer periods of time. This would be a major challenge, but would bring empirical substance to the integrating metaphor of bridges and why they don’t fall down.

The measurement of health, and more importantly the ability to link such measurement to the outcomes of policies intended to improve health, are critical aspects of population-based health (in contrast to health care) information systems that require improvement. Wolfson (1991, 1994a,b) discusses the contents and organization of the data systems required to mount and monitor health policy broadly defined, including methodological difficulties in arriving at an aggregate population health status index, and strategies for the integration of administrative and self-reported data and for sophisticated simulation modelling and detailed analyses of large volumes of microdata. In an extension of this line of thinking to the political task of creating constituencies for change, Renaud (1994) speculates about the creation of a “national social deficit index”, in parallel to and in competition with indices of fiscal deficit, to highlight the quality of the social environment.
The generation of another type of information, on the costs and benefits of different investment alternatives for health across different sectors, affords a direct opportunity for health economist researchers. There is at present both a lack of a substantial literature on inter-sectoral economic evaluation for health and a significant set of methodologic problems in conducting such evaluations. The economist’s traditional cost-benefit “tool kit” may not be adequate for the task; moreover, the “style” of presentation of inter-sectoral evaluations may have to be altered if they are to be “user-friendly” and therefore more likely to be used by decision-makers (Drummond and Stoddart, 1995).

Two examples may provide helpful illustrations of at least some of the issues involved. The importance of comparing health care with health producing strategies outside the health-care system was demonstrated by Joyce et al. (1988). They evaluated six strategies to reduce infant mortality: neonatal intensive care; prenatal care; abortion; teenage family planning; a supplemental food program for women, infants and children; and community health centres with maternal and infant care projects. Neonatal care was found to be one of the least cost-effective strategies, while the early initiation of prenatal care was found to be the most cost-effective strategy, and supplemental food programs were among the more cost-effective alternatives.

Labonte (1990) demonstrated that the opportunity costs of health care spending decisions are sometimes more vivid if expressed in real rather than nominal terms. He pointed out that the $350 million budget increase for hospitals in the Canadian province of Ontario in the 1990-1991 fiscal year could have funded 70,000 more publicly-subsidized housing units for low-income families or 547,000 more subsidized day care places for children. Joyce et al., and Labonte are good, but rare, examples of inter-sectoral evaluations for health (cell 3 again) at the level of specific services and entire systems, respectively.

There are also more fundamental challenges for health economists. I shall propose several here for discussion at the workshop. They include: moving beyond being primarily health care economists; focusing on the macroenvironmental production function for health (e.g. Auster et al., 1969) in addition to microeconomic and household production functions, but trying to better understand the “inside” of the broader production function (for example exploring socio-economic status
instead of controlling for it) while re-examining the concept of “choice” in the narrower ones (e.g. Birch and Stoddart, 1991); giving greater emphasis to distributional and equity issues in what are ostensibly efficiency analyses (witness the confusion of cost-control and cost-shifting in the user charge literature); engaging in a new research program on the economics of caring, i.e. how societies can show that they care about their members (a frequently cited symbolic role of health care) without committing large amounts of highly specialized resources; and seeking to disentangle and assess competing claims about the wealth-creating versus wealth-draining effects of health care systems (Lavis and Stoddart, 1994; Mustard, 1994; Evans 1994a).

The last of these may prove to be the most important for the macroeconomic high-wire act mentioned previously, because knowledge of the determinants of health and the determinants of economic prosperity are complementary. But the two need better integration. As their name implies, “health economists” could be ideally suited for this task, especially if they work in close conjunction with researchers from other disciplines and with policy-makers.
FIGURE 1

A SCHEMATIC OF HEALTH ECONOMICS

F. MICRO-ECONOMIC APPRAISAL
Cost-effectiveness, Cost-benefit and Cost-utility analysis of alternative ways of delivering care (e.g. mode, place, timing or amount) at all phases (detection, diagnosis, treatment, after-care, etc.)

E. MARKET ANALYSIS
Money prices; time prices; waiting lists & non-price rationing systems as equilibrating mechanisms and their differential effects in markets for physician and hospital services.

B. WHAT INFLUENCES HEALTH? (OTHER THAN HEALTH CARE)
Genetics; occupational hazards; consumption patterns; education; income; capital, human/physical; family background, etc.

A. WHAT IS HEALTH? WHAT IS ITS VALUE?
Perceived attributes of health; health status indexes; value of life; utility scaling of health.

C. DEMAND FOR HEALTH CARE
Influences of A + B on health care seeking behaviour; barriers to access (price, time, psychological, formal); agency relationship; need, altruism; insurance, demand for and effects of demand for care.

D. SUPPLY OF HEALTH CARE
Costs of production; alternative production techniques; input substitution; markets for inputs (manpower, equipment, drugs, etc.), remuneration methods & incentives; for-profit and non-profit organizations; HMOs, etc.

G. PLANNING, BUDGETING, REGULATION, & MONITORING MECHANISMS
Evaluation of effectiveness of instruments available for optimising the system; interplay of budgeting, manpower allocations, regulation, and the incentive structures they generate.

H. EVALUATION AT WHOLE SYSTEM LEVEL
Equity & allocative efficiency criteria brought to bear on E + F; inter-regional & international comparisons of performance; financing methods.
**FIGURE 3**

**FIGURE 3. A Taxonomy of Ways in Which Too Many Resources can be Devoted to Health Care**

<table>
<thead>
<tr>
<th>Is production efficient?</th>
<th>Technically inefficient</th>
<th>Cost-ineffective</th>
<th>Allocatively inefficient</th>
<th>What is being Produced?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Health Care</td>
<td>Health</td>
<td>Well-Being</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Individual</td>
<td>Population</td>
<td>Individual</td>
<td>Population</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Health care that is not effective.
2. Effective health care that is more costly than it need be.
3. Health that is more costly than it need be.
4. Health care that is valued at less than its cost.
5. Health that is valued at less than its cost.
6. Well-being that is more costly than it need be.
FIGURE 4

Life-style

Environment

Human Biology

Disease

Health Care

Other Factors
FIGURE 5

Respiratory tuberculosis: mean annual death rates (standardized to 1901 population), England and Wales, 1840–1970.

Source: McKeown (1979)
Whitehall study: all-cause mortality among total population by year of follow-up.

Source: Marmot (1986:23)
FIGURE 7

Level of social integration and age-adjusted mortality for males in five prospective studies. *RR*, the relative risk ratio of mortality at the lowest versus highest level of social integration.
FIGURE 8

Source: Evans and Stoddart, "Producing Health, Consuming Health Care," fig. 5, 1359.
REFERENCES


