

HEALTH POLICY INSTITUTE

LIES, DAMNED LIES, AND HEALTH CARE ZOMBIES: DISCREDITED IDEAS THAT WILL NOT DIE

**Morris L. Barer, PhD, Robert G. Evans, PhD,
Clyde Hertzman, MD, and Mira Johri, PhD**

HPI Discussion Paper #10

March 1998

HPRU 98:5D

THE UNIVERSITY OF TEXAS - HOUSTON
HEALTH SCIENCE CENTER



**UNIVERSITY OF TEXAS-HOUSTON
HEALTH SCIENCE CENTER**

HEALTH POLICY INSTITUTE

The Health Policy Institute is a component of the School of Public Health at The University of Texas-Houston Health Science Center, which is a part of The University of Texas System. The Institute was established to serve the Houston area, State of Texas, and U.S. through the provision of timely and relevant information on pressing health policy issues, and by conducting research and educational programs which contribute to the study of health policy. It sponsors conferences, seminars, visiting lecturers, and a discussion paper series to inform and facilitate the health policy debate. These activities provide direct linkages between academic research and policy making communities. The Institute is also affiliated with the U.T. Houston Program on Humanities and Technology in Health Care, which addresses ethical and technological issues in health care and science.

The Health Policy Institute has collaborative relationships with other policy institutes in the United States and Canada, with state legislative committees and executive agencies, particularly health and human services agencies, and local health coalitions. It also benefits from the breadth and depth of expertise of faculty members and health professionals within the U.T. School of Public Health, U.T. Houston Health Science Center, Texas Medical Center, and other academic institutions in Texas, especially those in Houston. The Institute also provides support for the National Medicine/Public Health Initiative, which has as its mission the "joining of Medicine and Public Health in a search to explicate problems and produce innovative solutions to deal with the health needs of the people of the United States."

Through publication of this Discussion Paper Series and other works authored by associated faculty and investigators, the Health Policy Institute makes information available to academics, government decision makers, other health policy institutes, and individuals concerned with issues affecting health.

This report is produced by the Health Policy Institute. However, the interpretations and comments are those of the author, and neither the Institute, the University of Texas, nor any supporting organization necessarily endorse the views and statements published here. For more information about the Institute or to obtain additional copies of this or other publications, please contact: Health Policy Institute, R.O. Box 20186, RAS E901, Houston, Texas 77225. Tel.: 713/500-9485. Fax: 713/500-9493.

To every complex problem there is a simple answer: neat, plausible, and wrong.
-H.L. Mencken

Why are health care costs so high, and so hard to control? Everybody knows the answer, or at least a large part of it. “Spiraling costs” have been driven by the aging of the population. As the baby boomers turn geriatric, the pressure can only get worse. And everybody is wrong. All the evidence points in a very different direction. Mencken would be amused.

The aging population story is, moreover, only one of a number of common misunderstandings, widely held and erroneous beliefs, about health care and its costs. Others include:

- Spiraling costs are driven by “frivolous” patient demands; these can be dampened through the judicious application of user fees.
- Costs would be even higher if insurance coverage were universal; no one is against covering the uninsured, if only it weren't so expensive. (This belief is exclusive to the United States.)
- You get what you pay for; higher costs buy more and/or better health care. Cost control inevitably implies “rationing” and imperils health.
- More and/or better health care is the only, or at least the major, way to improve health.

Each of these is based on a mix of truths, half-truths, and fallacies. They offer interpretations of the forces driving cost expansion in the health care sector that are “neat, plausible, and wrong” – and dangerously misleading for health care policy.

In fact, on each of these points, the evidence is already in. Health services researchers have studied these relationships for decades. Their findings, and **their** “common understandings,” are largely or wholly at variance with the propositions above. Yet the popular beliefs live on. So why is the evidence so impotent, and why are the

fallacies so strangely difficult to kill? However deeply buried in refutation, they have a zombie-like tendency to re-emerge.

The primary purpose of this paper is to explore the sources of this resilience. We will briefly review the current evidence on each of the ideas listed above, but all of this is (or should be) well-known to the research community. We will then explore two important classes of reasons for the resilience of discredited ideas in health policy debates.

First, these false ideas (or “zombies”) carry with them implicit policy recommendations bearing on some aspect of health care financing. In all cases these amount to “Send money, quickly.”¹ As American policy analyst Ted Marmor points out, “Nothing that is regular is stupid.” Particular individuals and groups stand to benefit, economically or otherwise, from the promotion of these ideas. Although often hidden in public debate, these vested interests are easily identifiable, and part of our task in what follows will be to expose them.

Second, the resilience of these ideas also depends crucially on the extent to which they resonate in the popular imagination. Although this is a complex phenomenon, we will show that it is closely linked to their fit with our prior understandings and experiences, both inside and outside the formal health care system.

Zombie # 1: The Health Care System is Drowning in a Wave of Grey — Send Money!”

The societies in which we live are aging. And, as everyone knows, growing older leads to a greater average need for health care services. As the population ages, therefore, health care costs must increase if the health care system is to respond to increasing needs. And the worst is yet to come! Since the greying of the population in the United States and Canada will accelerate over the next few decades, keeping pace with growing needs will require that the rate of increase in health care spending rise as well. If not, the result may be more avoidable illness, suffering and death.

¹ The apparent exception, user fees to control costs, turns out—as will emerge below—to be in reality another way of raising money.

The demographic parts of this story are true—the populations of the industrialized countries are becoming older, and older people do, on average, need and use more health care. The rhetoric of crisis, which portrays the health care system as facing imminent collapse due to exogenous changes in the demand for health care, is however completely at odds with current evidence. Although, over long periods of time, the effects of demographic trends can be (and likely will be) quite substantial, these effects move like glaciers, not avalanches. They have not strained our health care systems in the past, and need not pose a future threat to the overall capacities of those systems.

This point has been demonstrated through both Canadian and international empirical research. This has repeatedly shown that population aging *per se* accounts for very little of the increase in health care costs seen over the last three decades; and that, despite accelerated aging, it will not be the key driving force over the next three.² But the anticipation of disaster, and the claim that the aging of the population threatens to bankrupt our healthcare system, persists. Why?

We see two principal reasons for the persistence of this myth. First, the crisis definition has been strategically useful--from the perspective of those who stand to gain--for promoting increased spending in the health care sector. Health care systems do not simply become bankrupt. The rhetoric of demographic inevitability is a thinly disguised claim for more resources. And the claim is non-negotiable because the alleged trend towards increasing needs is portrayed as both exogenous and inevitable. Yet the reality is that the forces driving health care costs have little to do with demographics and are largely within the health care system itself.

For example, while use of health care services by seniors in British Columbia has increased dramatically, the crucial factor driving this increase has been changing medical practice with respect to the care of seniors. Put simply, more is being done for, and to, each elderly person. This suggests that the critical issue facing health care managers and policy-makers is the very difficult one of appropriate care of seniors. Part of the

² We offer an extended analysis of the “aging crisis” in Canada, making use of empirical findings from British Columbia and elsewhere, in Morris L. Barer, Robert G. Evans, and Clyde Hertzman, (1995) “Avalanche or Glacier?: Health Care and the Demographic Rhetoric,” *Canadian Journal on Aging / La Revue canadienne du vieillissement*, Vol. 14, No. 2; pp. 193-224.

seductiveness of the “demographic inevitability” story lies in its ability, under the guise of necessity, to help us to evade these complex problems.

In a similar vein, Northcott argues that blaming the population aging trend for our current fiscal difficulties serves at least two purposes.³ First, the rhetoric of crisis provides a straightforward answer to the intricate problems of public policy. It succeeds in focusing attention on a simple, visible “cause” –that of population aging–and thereby diverts attention from more significant causes, which may be more difficult and politically dangerous to address. Second, blaming demographics mobilizes support for sacrifices (in the form of either increased taxation or program cuts) to be borne by the public generally, and seniors in particular. It is then a short hop from blaming the aging trend for problems of public policy, to blaming the aged themselves.⁴

But these reasons, although undoubtedly significant, are insufficient to account for the widespread influence of this idea in public debates. The myth of demographic inevitability has a peculiarly strong intuitive appeal. We believe that this is due in large part to several forms of apparent *observational* support for this myth. We term these “compression of time frame,” “selection of sub-populations,” and “clinical experience.”⁵

Compression of Time Frame:

Populations, like people, age slowly. Estimates of the impact of aging on per capita total health care costs (in real terms, net of inflation) in Canada, for the *whole* population, generally place it at about 1 per cent per capita per year, averaged across all types of health care services.⁶ If sustained over a long period of time, this process will have a very large cumulative impact on real per capita health care costs. But the way in which these data are framed can communicate conclusions that are more, or less, alarmist.

³ H.C. Northcott, (1994) “Public Perceptions of the Population Aging ‘Crisis’” *Canadian Public Policy-Analyse de Politiques*, Vol. 20, No.1; pp. 66-77.

⁴ On this point see Robert G. Evans, (1985) “Illusions of Necessity: Evading Responsibility for Choice in Health Care” *Journal of Health Politics, Policy and Law*, Vol. 10, No.3; pp. 439-467.

⁵ See once again Barer, Evans and Hertzman, “Avalanche or Glacier?” (1995).

⁶ See e.g. Woods Gordon Management Consultants (1984), *Investigation of the Impact of Demographic Change on the Health Care System in Canada -Final Report* (Prepared for the Task Force on the Allocation of Health Care Resources (Joan Watson, chairman)), Toronto: Woods Gordon. The point was first made by J.A. Boulet and G. Grenier (1978) “Health Expenditures in Canada and the Impact of Demographic Changes on Future Government Health Insurance Program Expenditures,” *Economic Council of Canada Discussion Paper #123*. Ottawa: ECC (October).

Figure 1(a) depicts a hypothetical increase of per capita costs at 1 % per year over a 40-year period. By choosing a “scaling” which simultaneously compresses the presentation of time-frame on the horizontal axis and enlarges the scale on the vertical axis, and by isolating this trend from any other data, we can graphically convey a sense of crisis. Figure 1(b) presents exactly the same projection, but uses a different scale. By situating the trend line within a confidence band projecting hypothetical high and low rates of economic growth, it defuses the impression of impending crisis which the first diagram created. This approach gives greater emphasis to the passage of time, and underscores the fact that other relevant variables change as well.

The latter context enables us to see that even a sustained trend of low economic growth would enable us to support an expansion of health care services adequate to satisfy the needs associated with the aging of the population, without having to increase the share of resources devoted to this purpose. Time provides the capacity to respond. Or at least it would, if current utilization patterns were maintained. That, as we will show, is the intriguing part of the rising costs story.

Selection of Sub-Populations:

The previous estimates of the impact of population aging in Canada, which place it in the neighborhood of 1 % per capita per year, were calculated for all services provided to all people. Disaggregating these data reveals what seems intuitively obvious; this trend will have a differentially *greater* impact on those programs and services used more frequently by elderly patients. In some sectors of health care provision, the impact may be very large. The relatively rapid proportional increase in those over 85 years of age, for example, is having (and will continue to have) a major impact on the long-term care sector in Canada. However, because this group is only a small subsection of the population, the stress it places on the system as a whole is limited. This broader, society-wide perspective helps to explain why inquiries into the impact of population aging on Canada's Medicare program yield much less alarming results than similar examinations

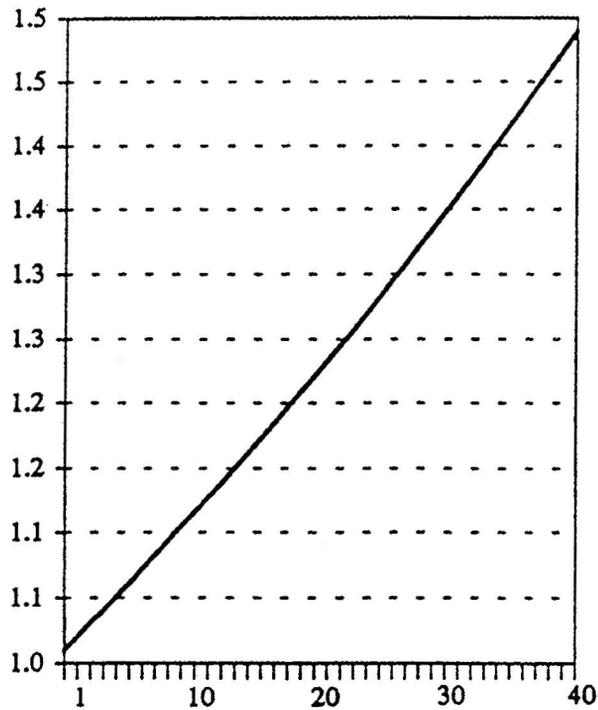


Figure 1a Health care costs per capita: A: Growth at 1% per annum, 40 years

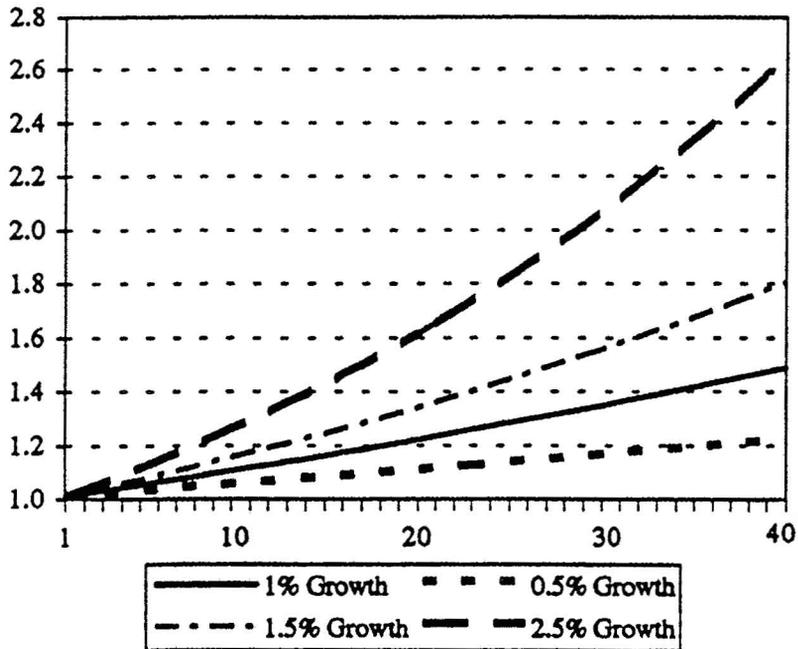


Figure 1b Health care costs per capita: B: Growth at 1% per annum, 40 years

SOURCE: Barer, Evans and Hertzman (1995), "Avalanche or Glacier?"

of the US Medicare system.

The way in which state-funded health care is partitioned administratively in the United States leads to a particular problem of perception. The American Medicare system provides coverage almost exclusively to persons over 65. Within this group, both average age and average utilization are increasing faster than in the population as a whole. Studies focusing only on this sub-population yield more worrying estimates of growth in costs and utilization. Once again, partial data presented in isolation are alarmist, and tend to feed the rhetoric of crisis.

And the rest of the story—the changes in patterns of utilization for other age groups—is very interesting. We will illustrate this point below, using data from our studies in British Columbia.

Clinical Experience:

Perhaps the most compelling source of empirical corroboration for the idea that our health care programs are being irresistibly submerged in an encroaching grey tide comes from the daily experience of health care workers, who perceive their workload to have become dominated by elderly patients. And they are largely right. An appreciation of how the experiences of clinicians may feed the story of demographic inevitability can be gained through considering the evolution of the Canadian hospital sector over the past 30 years. During this time, the mix of patients in Canadian hospitals has changed dramatically. As Figures 2a and 2b illustrate, in 1961, those 65 and over made up about 7.6% of the population, accounted for a share of hospital separations somewhat greater than 10%, and about 30% of patient days. But the caseload of the Canadian hospital system in 1961 included a substantial representation of patients from each stage of the life cycle.

By the 1991/92 fiscal year, about 60% of inpatient days in Canadian hospitals were taken up by the care of seniors. Pediatric use was down to a mere 5%. Persons over 65 made up 30% of the patients served, and also had much longer average lengths of stay than their younger counterparts. Fully forty percent of all patient days were dedicated to those over 75 years of age. Hospitals, or at least their inpatient component, had by the

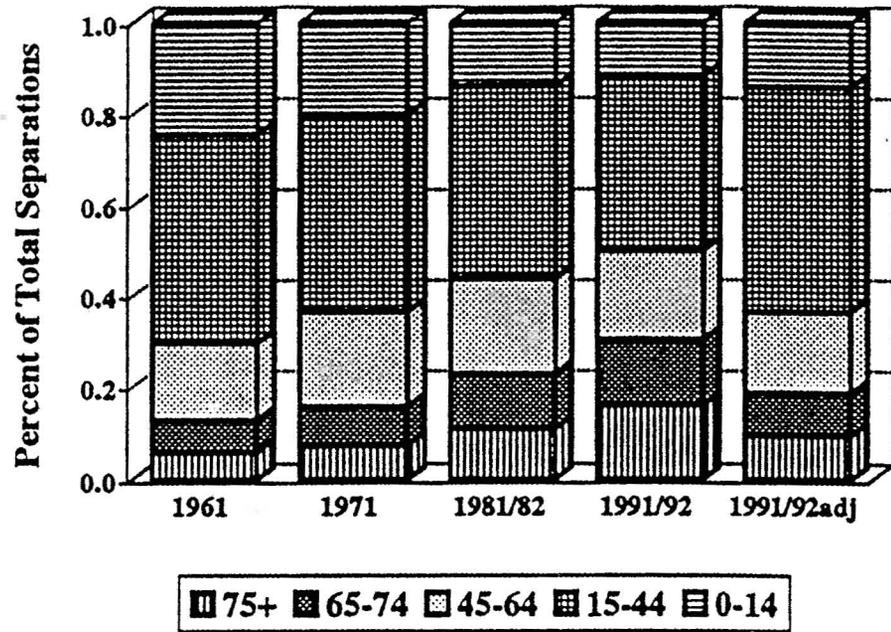


Figure 2a Canada: hospital separations by age, general and allied special hospitals

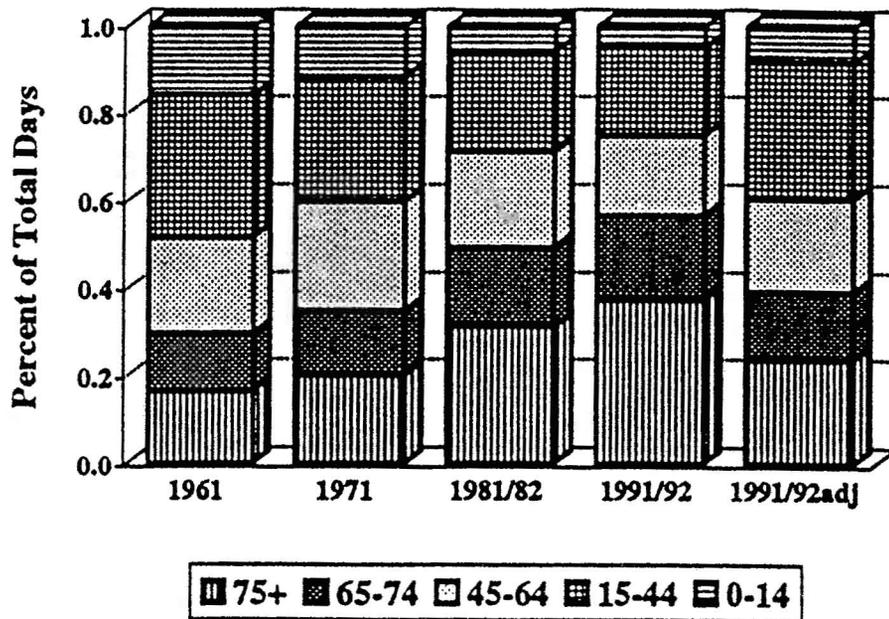


Figure 2b Canada: hospital inpatient days by age, general and allied special hospitals

SOURCE: Barer, Evans and Hertzman (1995), "Avalanche or Glacier?"

early 1990s come to be primarily facilities for care of the elderly.

Clinicians' experiences of this dramatic reorientation of the hospital sector, conjoined with the fact of rising costs, appeared to provide a powerful confirmation of the rhetoric of demographic inevitability. Yet a closer analysis of the data yields quite a different story. The demographic composition of the Canadian population did change in this 30-year time period. In 1991/92, people over the age of 65 made up 11.7% of the population (a significant increase over the 1961 share of 7.6%), and 4.7% were over 75. However, although some of the difference in separations and patient days clearly is explained by this shift, it is not the most significant part.

The last bar in each of Figures 2(a) and 2(b) is an "adjusted" bar for the 1991/92 year. It represents what the mix of patients in 1991/92 would have been, if the composition of the population had changed as it did, but population-specific rates of hospital use had remained unchanged since 1961. It thus shows how much of the change in patient mix would have been generated through demographic restructuring alone. These adjustments show that the increasing proportion of elderly people in the population would have accounted for just under 20% of all separations in 1991/92, and 40% of the days, rather than the 30% of separations and 60% of days which were in fact used. Roughly speaking, then, the increasing numbers (and ages) of the elderly accounted for about one third of their increased share of hospital use over this thirty year period.

Much more significant, however, were the major changes that occurred in the rates at which people in different age groups use the hospital system (Table 1). Per capita use by people under the age of 45 decreased by approximately one-third in separations and over one half in patient days over the period 1961 to 1991/92. Use (as measured in hospital separations) by those 45-64 fell by about 7.5% during this same time period. However, separations per capita for those over 65 rose between 30 and 40%. Substantial decreases in length of stay realized in the 1980s were not enough to offset these increases. Per capita use rates have thus "twisted around" the age distribution, falling sharply at the young end but rising at the old, and this, *not* demographic change, has been the primary factor behind the differences that health care providers observe in the hospital patient

Table 1

Hospital inpatient utilization, general and allied special hospitals, rate per 1000 population by decade

<i>Year</i>	<i>Total</i>	<i>0-14</i>	<i>15-44</i>	<i>45-64</i>	<i>65-74</i>	<i>75 +</i>
<i>Separations per 1000 Population</i>						
1961	149.8	107.7	167.8	143.7	211.0	330.4
1971	165.9	112.3	164.7	184.5	276.1	404.2
% change	10.73%	4.31%	-1.89%	28.35%	30.83%	22.34%
1981/82	146.8	90.4	127.5	160.9	284.4	449.1
% change	-11.50%	-19.48%	-22.54%	-12.78%	3.01%	11.10%
1991/92	134.8	73.7	109.1	133.2	276.1	461.5
% change	-8.16%	-18.51%	-14.48%	-17.20%	-2.92%	2.76%
% change since 1961	-10.00%	-31.56%	-35.01%	-7.30%	30.84%	39.68%
<i>Patient days per 1000 population</i>						
1961	1632.9	764.3	1292.3	2084.2	4389.4	9780.1
1971	1907.6	752.3	1251.4	2564.3	5659.6	12450.3
% change	16.83%	-1.57%	-3.16%	23.04%	28.94%	27.30%
1981/82	1805.7	479.2	841.8	2043.6	5346.7	15404.7
% change	-5.34%	-36.31%	-32.74%	-20.31%	-5.53%	23.73%
1991/92	1529.6	354.3	651.8	1413.9	4213.7	12044.0
% change	-15.29%	-26.06%	-22.56%	-30.81%	-21.19%	-21.82%
% change since 1961	-6.33%	-53.64%	-49.56%	-32.16%	-4.00%	23.15%

SOURCE: Nair (1991, pp. 189-197). Updated to 1990/91 with data from "Hospital Morbidity, 1989-90". Health Reports Supplement no. 1, 1992 Vol. 4, no. 2, SC Cat. #82-003S1, p. 8, and "Hospital Morbidity, 1990-91" SC Cat. #82-216, p. 8. Updated to 1991/92 with unpublished data from Statistics Canada.

population. These shifts in per capita use rates accounted for roughly two thirds of the change in Canadian national data over this thirty-year period.

And these data are likely conservative. Although they include long-term care units in acute care hospitals, they exclude free-standing long-term care hospitals, as well as various intermediate care facilities. The patient populations of such institutions expanded greatly over this period, and they are almost entirely elderly, mostly very elderly. More comprehensive data drawn from the province of British Columbia, which include extended care facilities, and a finer degree of disaggregation among subdivisions of the elderly population, demonstrate even more clearly the extent of these major changes in age-specific hospital use patterns. As shown in Table 2 and Figure 3, use of hospital services in 1985 as compared to 1969 skyrocketed among the elderly, particularly in those over the age of 85, while falling precipitously for the very young, adolescent, and middle-aged groups.⁷ Since the elderly represent a small proportion of the total population, very large increases in their utilization rates were almost precisely offset by smaller decreases for the rest of the population.

The BC hospital utilization analyses also provide data necessary to identify the factors driving the growing concentration of resources on the care of the elderly. Figure 3 shows that increases in utilization by seniors resulted from both increased hospital separation rates per capita, and from large increases in average length of stay. In order to appreciate more clearly what is happening, it is helpful to subdivide these data into acute and extended care sectors.

The elderly in BC occupied an increasing share of a declining acute care sector, in which they accounted for 28% of acute care days in 1969, and 41 % in 1985/86. Separation rates for the elderly went up sharply in this setting, in marked contrast to the declines in these rates seen in the rest of the population. This trend was, however, offset by declining average lengths of stay, such that patient-day use per capita by the elderly actually fell. In fact, the BC data indicate that virtually all of the increase in elderly acute

⁷ R.G. Evans, M.L. Barer, C. Hertzman, G.M. Anderson, I.R. Pulcins, and J. Lomas (1989), "The Long Good-Bye: The Great Transformation of the British Columbia Hospital System" *Health Services Research*, Vol. 24, No. 4; pp.435-59.

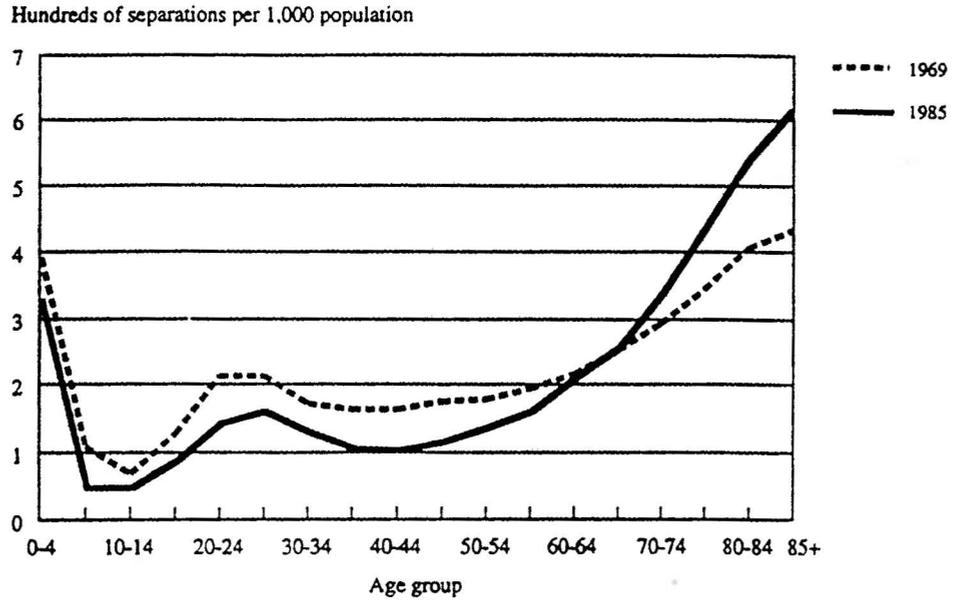
Table 2

Proportion of change in B.C. hospital days accounted for by increases in population, changes in age distribution of the population, and changes in rates of use, 1969-1985

Total hospital days, 1985-86		5,906,271
Total hospital days, 1969		4,050,504
Increase in hospital days		1,855,767
Increase in population	38.6% →	1,564,000
Increase in use due to changes in population age structure	8.0% →	452,000
Change in use associated with age-specific changes in rates of utilization:	-2.6% →	-158,000
0-4 yrs	-41.3%	
5-14 yrs	-48.7%	
15-24 yrs	-37.8%	
25-44 yrs	-41.5%	
45-64 yrs	-26.3%	
65-74 yrs	-9.0%	
75-84 yrs	54.6%	
85+ yrs	224.8%	

SOURCE: Barer, Evans and Hertzman (1995), "Avalanche or Glacier?"

Total Separations



Total Days of Separations

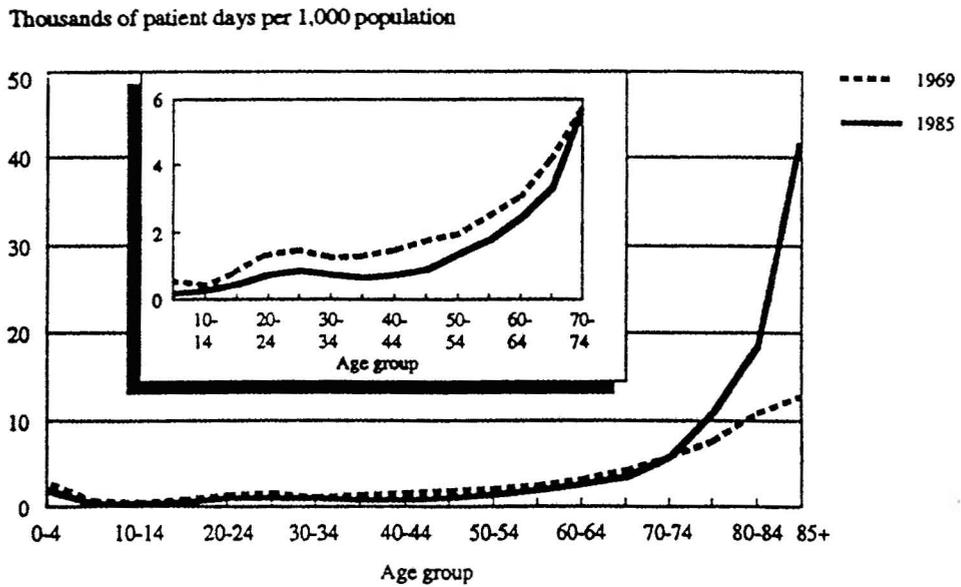


Figure 3 Age-use curves, total days of separations and total separations per 1,000 population, British Columbia, 1969 and 1985

SOURCE: Barer, Evans and Hertzman (1995), "Avalanche or Glacier?"

care patients is accounted for by rapidly increasing surgery rates in persons over 65, at a time when such rates were falling for the rest of the population. Moreover, surgical cases tend to have shorter average lengths of stay.⁸ This observation from the BC data is strongly echoed at the national level. From 1971 to 1985/86, surgical rates for the over 64 population in Canada rose by 45.9%, while falling in every other age group. The total population rate fell by 3.1 %. By 1985/86 people 65 years of age or over were undergoing surgery at roughly double the rate for the younger population; in 1971 their rates had been only 10% higher.⁹ Non-surgical separations fell in every age group, by amounts ranging from 2.8% for the elderly, to over 40% for people aged 15 to 44, with an average decline for the population as a whole of 20.6%.¹⁰

By contrast, in the extended care setting, both separation rates and lengths of stay rose dramatically, resulting in the startling increases in the patient days shown in Table 3. The increase in days of care in extended and rehabilitation beds was concentrated in a very small number of diagnostic categories. Of the total increase in patient days between 1969 and 1985/86, one third were apportioned to patients with diagnoses of Alzheimer's senile dementia, senility, and other degenerations of the brain. One quarter of the increase during this period was attributable to the chronic sequelae of stroke, which constituted 40% of the increase in the 1970s, but made no further contribution to increases after 1980. An additional 10% of the increase was accounted for by heart disease. During the 1970s, the primary condition responsible for the increase was chronic ischemic heart disease; however, during the 1980s heart failure took over as the major cardiovascular source of increase. The other major sources of increase were confined to one time period only: 15% of the increase during the 1970s was related to arthritis, while

⁸ G.M. Anderson, I.R. Pulcins, M.L. Barer, R.G. Evans and C. Hertzman (1990), "Acute Care Hospitalization Under Canadian National Health Insurance: The British Columbia Experience from 1969-1988," *Inquiry*, Vol. 27; pp. 352-358.

⁹ R. Riley and M. Kanigan (1989) "Surgical Procedures and Treatments, 1985-86" *Health Reports* Vol. 1, No.2 (SC 82-003), Ottawa: Statistics Canada, pp. 253-54.

¹⁰ M. Kanigan and P. Mix (1989) "Surgical Procedures and Treatments-Historical Trends and Recent Data Characteristics" *Health Reports* Vol. 1, No.1 (SC 82-003) Ottawa: Statistics Canada, pp. 1-96.

Table 3

Distribution of days in British Columbia hospitals, 1969 and 1985/86, by age of patient and type of bed*

1969 Age	Days per 1000		% of All Days	
	Acute	Extended	Acute	Extended
0-14	1186.7	41.4	17.5	0.6
15-64	1479.4	74.8	46.4	2.3
65-74	4078.4	817.5	11.3	2.3
75-84	6421.8	2430.2	10.3	3.9
85+	8416.2	4332.8	3.6	1.9
Total	1750.4	215.6	89.0	11.0
1985/86				
Age	Days per 1000		% of All Days	
0-14	657.7	92.4	6.6	0.9
15-64	809.0	201.6	26.4	6.6
65-74	2819.6	1635.5	10.0	5.8
75-84	5402.9	8278.9	9.2	14.1
85+	7041.9	33049.4	4.1	16.9
Total	1164.1	904.2	56.3	43.7

*We classified patients as residing in one of four types of bed: acute care bed in acute care hospital; extended care bed in acute care hospital; rehabilitation care bed; or extended care bed in extended care hospital. All but the first are included in our "extended" category in this table.

SOURCE: Barer, Evans and Hertzman (1995), "Avalanche or Glacier?"

14% of that during the 1980s was related to hemiplegia and quadriplegia.¹¹

The narrative of “demographic inevitability” would lead one to suppose that the increase in hospital utilization represents a response to increased needs. However, we were unable to find any evidence that the increased use of hospital resources to treat dementia and related conditions was a response to an increase in prevalence in the provincial population. The underlying incidence of stroke and heart disease actually *fell* during this period. Alternatively, one might argue that the increase is due to the fact that patient needs for these conditions were previously unmet, or rather, that advances in medical technology have improved our ability to treat these conditions. This explanation is certainly plausible in some cases.¹² However, for a condition such as senile dementia, for example, it is doubtful whether more extensive institutionalization of patients represents improved therapy and better outcomes. Our finding of longer stays ending in death for this class of conditions would seem to support the conventional view that little can be done to affect outcomes in these cases. In brief, the problems and outcomes themselves appear to be unchanged over this period; it is the manner in which Canadians have chosen to deal with them that has altered dramatically.¹³

Empirical data drawn from Canada as a whole, and from BC, thus enable us to

¹¹ C. Hertzman, I.R. Pulcins, M.L. Barer, R.G. Evans, G.M. Anderson and J. Lomas (1990) “Flat on your Back, or Back to your Flat? Sources of Increased Hospital Services Utilization Among the Elderly in British Columbia,” *Social Science and Medicine* Vol. 30, No.7; pp. 819-828.

¹² We have explored the issue of the appropriateness of the relative increases in acute and extended care use by the elderly in some detail in *Avalanche or Glacier?*

¹³ M.L. Barer, R.G. Evans, C. Hertzman and J. Lomas (1987) “Aging and Health Care Utilization: New Evidence on Old Fallacies,” *Social Science and Medicine* Vol. 24, No. 10; pp. 851-62. It has been argued by A. Walker (1987) in “Meeting the Needs of Canada’s Elderly with Limited Health Resources: Some Observations Based on British Experience,” (Economic Council of Canada, *Aging With Limited Resources. Proceedings of a Colloquium on Health Care, May 1986*, Ottawa, Minister of Supply and Services Canada) that this transformation came about as a result of changes in patterns of informal support available to those affected, which placed increased pressure on the institutional sector to absorb more of these individuals. But any such pressure was assisted by an evolving social movement that converted what had previously been considered the inevitable cognitive declines of aging into a series of discrete clinical diagnoses. P. Fox (1989), in his “From Senility to Alzheimer’s Disease: the Rise of the Alzheimer’s Disease Movement,” (*The Milbank Quarterly* Vol. 67, No.1; pp. 58-102), has argued that this resulted in the clinical institutionalization of Alzheimer’s disease, and the increasing awareness of other dementia-related conditions as disease entities worthy of clinical and research interest. The inevitability of the human condition was thereby converted into new opportunities for research and new “needs” for care. To date there is no evidence that patients have benefited as a result of this change, although informal caregivers and the research community may have.

establish three factors underlying clinicians' observations:

- The proportion of hospital care used by seniors has, indeed, increased; but
- Only a small fraction of the increase has been due to demographic changes *per se*. A much larger proportion has been due to changing rates of age-specific utilization among seniors; but
- The connections between these changes in utilization patterns, and underlying population needs, are at best unclear.
- Increased use by seniors has been accompanied by large decreases in use among the rest of the population. The “grey” proportion of the patient population has thus increased dramatically, but not as a result of demographic changes.

Our analysis substantiates clinicians' experience, but shows that its causes are quite different from what is usually supposed.

Parallel findings on the impact of aging have been made in the physician services and pharmaceuticals sectors. Although there were relatively large increases in medical care for all groups, little of the increase in physician services can be attributed to changes in the age structure of the population of British Columbia. A “twisting” of the age-use curve similar to that seen in the hospital sector was observed, with the highest rates of increase in use per capita seen among the elderly. The increases were, moreover, concentrated in the area of specialist (particularly medical specialist) services for seniors.

Studies of pharmaceutical utilization in the province of BC also found the overall impact of population aging *per se* on increased drug costs to be small. As with hospital services, the reallocation of British Columbia's health care resources toward the elderly occurred not principally because there were more of them, but because, at any given age, each was serviced on average more intensively. These results strongly suggest that the changes in utilization patterns are motivated by changing clinical standards of investigation and intervention.

The trends described above are not unique to Canada. Indeed, they appear to represent the common experience of the industrialized world, independent of the details of health care system organization. Similar observations emerge from work in

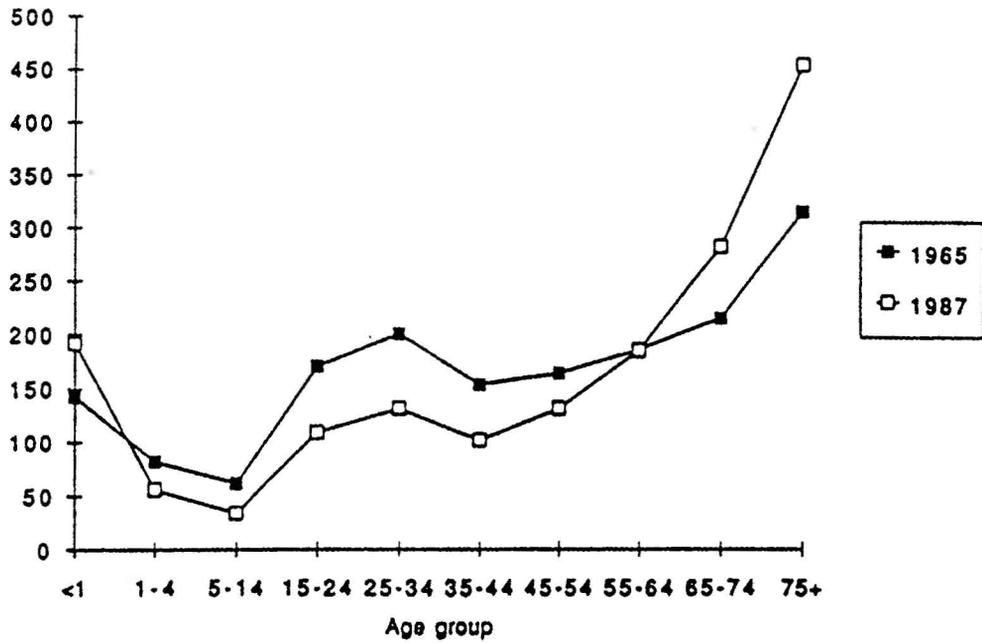


Figure 4 Hospitalizations per 1,000 Population, by Age Group, United States, 1965 and 1987

SOURCE: National Center for Health Statistics. 1965, 1987.

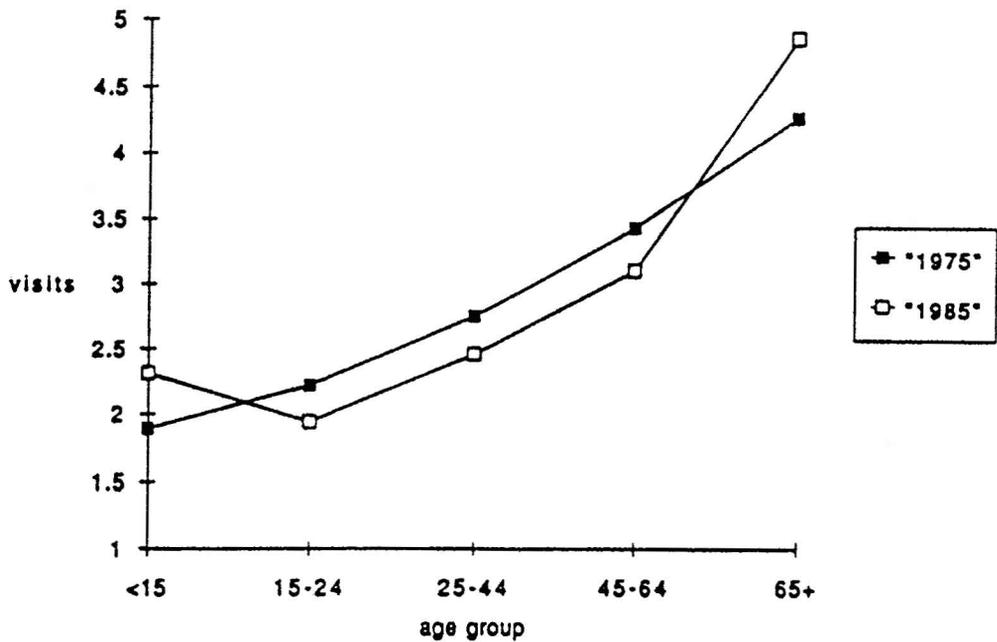


Figure 5 Physician Visits per Person, United States, 1975 and 1985

Source. Nelson and McLemore 1988.

Australia¹⁴, Sweden¹⁵, and the United States¹⁶. Figures 4 and 5 show a “twist” in the age-use curve for the US hospital and physicians sectors analogous to that seen in our research from BC.

An analysis of the relationship between demographic structure and health care costs for the countries of the GECD led Getzen to conclude:

...aging is not a significant cause of rising health care costs...spending is a result of political and professional choices, rather than the outcome of objective trends...beyond our control.¹⁷

The perceived “aging crisis” in health care delivery and finance is, then, almost entirely a consequence of the changes that have taken place, for whatever reasons, in patterns of care for the elderly. An examination of the appropriateness of these changes is of paramount importance. It is also very difficult. And in spite of the consensus on this issue in the scientific literature, the pattern of fallacious explanation—the idea that the trend is inevitable, because it is due to demography—persists.

Zombie # 2: “The problem with the health care system is too many patients making too many frivolous demands....increase user fees!”

“Runaway costs” of health care are a consequence of ever-greater demands placed on the system by individuals, and user fees are essential to combating this trend. Such fees help to control costs by encouraging consumers to think twice about using care, and thereby to reduce unnecessary use—read “abuse.”

¹⁴ M.L. Barer, M. Nicoll, M. Diesendorf and R. Harvey (1990), “From Medibank to Medicare: Trends in Australian Medical Care Costs and Use, 1976-1986,” *Community Health Studies* (Australia), Vol. 14, No.1; pp. 8-18.

¹⁵ R.G. Evans (1991), “Reflections on the Revolution in Sweden,” in *International Review of the Swedish Health Care System* [Den Svenska Sjukvarden] (chaired by A.J. Culyer, with J.M. Graf von der Schulenberg, W.P.M.M. van de Ven, and B.A. Weisbrod) Occasional Paper No. 34, SNS Sweden, Stockholm: SNS, 1991; pp. 188-57.

¹⁶ M.L. Barer, C. Hertzman, R. Miller & M.V. Pascali (1992), “On Being Old and Sick: The Burden of Health Care for the Elderly in Canada and the United States,” *Journal of Health Politics, Policy and Law*. Vol. 17, No. 4; pp. 763-82. See also Mendelson, D.N. and W.B. Schwartz (1993) “Effects of Aging and Population Growth on Health Costs” *Health Affairs*, Vol. 12, No.1; pp. 119.25.

¹⁷ T.E. Getzen (1992), “Population Aging and the Growth of Health Expenditures” *Journals of Gerontology (Social Science)*, Vol. 47, No.3; S98-S104.

This story is, on the surface at least, based on an understanding of the determinants of health care use quite different from that underlying the “inevitable demographic crisis.” The latter assumes implicitly that use and costs are determined by patient needs, professionally interpreted, rather than by “consumer” demands that may be unrelated to need (*i. e.* frivolous).

The idea that user charges are an essential part of an overall expenditure control strategy, however, is widespread in only in one OECD nation--the United States.¹⁸ While user fees have advocates in every country, most of those advocates have interests only in thwarting expenditure control strategies. They represent provider groups, who anticipate that such charges will serve to *increase* overall costs. The relative cost experience of the United States tends to bear out their expectation-it is also the country that has been least successful in limiting the growth of health care spending.¹⁹

Yet in the United States the idea that consumer co-payments make good economic (and perhaps moral) sense steadfastly resists permanent burial. Why? The interest in user charges bears the familiar hallmarks of a zombie. First, in spite of its popularity, it is intellectually dead, and second, its overwhelming appeal is a product both of its public resonance, and of the efforts of powerful interest groups to keep it on the agenda. This issue demonstrates the role of vested interests with particular clarity. There are good and logical reasons why claims for user fees hang around, but they bear little relation to those commonly offered.

¹⁸ Public payment systems in all OECD countries require patient cost-sharing for some services, and / or by some sub-populations. Outside the US, however, such charges play a relatively minor role; and nowhere can they be said to be part of a cost control strategy. See Greg L. Stoddart, Morris L. Barer and Robert G. Evans (1994) “User Charges, Snares and Delusions: Another Look at the Literature,” Toronto: Ontario Premier’s Council on Health, Well-being and Social Justice, pp. 28-30.

¹⁹ National and international evidence reveals that user fees are strongly correlated with higher and less controllable costs. Furthermore, this process appears to be both causal, and mutually reinforcing. Introduction and expansion of user fees are possible avenues of response to fiscal crisis, an attempt to shift costs one cannot control. However, the resulting diversification of funding sources can lead to an increasing loss of control, and hence, more rapid cost escalation.

Why User Fees?²⁰

The argument that user fees will control health care costs by deterring unnecessary use of health care services begins with the assumption that the rising cost of health care is the consequence of ever-growing patient demand. Once this definition of the problem has been accepted, user fees offer themselves as an apparently commonsense solution. As most of us are only too aware, raising the price of even our favorite goods encourages us to reduce consumption. And if this use was excessive to begin with, this can only be a good thing.

Formal economic theory can also be recruited—it is often interpreted (incorrectly) as predicting that increasing the price faced by a consumer/patient must reduce that person's use of care. When summed over all individuals this will produce a drop in total utilization, with the result that the costs of health care would then fall, or at least be less than otherwise. Experimental evidence has also been assembled, apparently supporting this hypothesis. Probably the most widely known is the US-based RAND health insurance experiment, which has shown that people will respond to financial incentives by reducing their use of care.

The apparent coherence of this straight-forward account dissolves rapidly under more detailed scrutiny. If the goal is simply to reduce health care expenditures, patient cost-sharing turns out to be a peculiar choice of instrument. By far the largest proportion of health care spending--hospital admissions, specialist referral, diagnostic tests and other procedures, and prescription drugs--depends upon the decisions of physicians and other providers. Only a small share of health care expenditures is accounted for by patient-initiated visits.

If, however, the goal of a user fee policy is to deter patients from purchasing unnecessary or inappropriate care, then the policy would have to discriminate, within the subset of patient-initiated use, between “needed” and “frivolous” services. In fact, the

²⁰ We have examined a number of these ideas elsewhere. See especially Greg L. Stoddart, Morris L. Barer, Robert G. Evans, and Vandna Bhatia, “Why Not User Charges? The Real Issues,” (1993); and Morris L. Barer, Vandna Bhatia, Greg L. Stoddart, and Robert G. Evans, “The Remarkable Tenacity of User Charges,” (1994) Toronto: Ontario Premier's Council on Health, Well-being and Social Justice, June 1994.

vast majority of patient-initiated first visits have been found to be medically appropriate. Of those which could be classed as unnecessary, only a small subset are likely to represent “abuse”—care which is not only medically inappropriate, but which could reasonably have been known in advance to be so by the average (non-clinically trained) patient. Rough estimates for medical expenditures in Canada on patient-initiated abuse are in the range of 1 to 2% of total health care spending.²¹ Policies aimed at modifying patient demands are chasing the wrong target.

What Can We Learn From the Evidence?

It is important to bear this point in mind in assessing the results of the RAND study. The aim of the RAND Health Insurance Experiment (HIE) was to determine the effect of patient cost-sharing on the use of health services. It undertook to answer this question by conducting a large-scale social experiment.²²

Between November 1974 and January 1982, 3,958 persons ages 14-61 from some 2,005 families in six American cities took part in the trial. Of this number, 70% were enrolled for three years, and 30% for five years. Families were randomly assigned to fee-for-service insurance plans which provided a comprehensive set of benefits, but varied in two principal dimensions: the co-insurance rate and the Maximum Dollar Expenditure ("MDE"), or upper limit on family out-of-pocket annual expenses. The co-insurance rates selected were 0 (free care), 25, 50 or 95%. The expenditure ceiling or MDE was set at 5, 10, or 15% of family income, up to a maximum of \$1,000. Beyond this point, health care was free to the family.

Although the enrollees for the experiment were selected through a process of random sampling, several groups were not eligible to participate. Significant exclusions from the study population were the elderly (62 years of age or older at the time of

²¹ Greg L Stoddart, Morris L. Barer, Robert G, Evans, and Vandna Bhatia (1993), “Why Not User Charges? The Real Issues,” Toronto: Ontario Premier's Council on Health, Well-being and Social Justice, June 1994.

²² The experimental approach generates data untainted by the problem of “adverse selection.” If people in poorer health purchase health plans with more comprehensive coverage than those in better health, this could create a false impression that this coverage leads to greater use.

enrollment), the wealthy (the 3% of the study population with 1973 income in excess of \$25,000), and some disabled (those eligible for the Medicare disability program).²³

The experiment generated several interesting findings. It demonstrated that participants randomly assigned to fee-for-service plans with high levels of co-insurance had significantly lower health care expenditures than those assigned to a free care fee-for-service plan. Per capita annual expenditures in the free care group were 23% higher than those for the 25% co-insurance group, and 46% higher than those for the 95% co-insurance group.

From these observations the investigators calculated a price elasticity of demand²⁴ of -0.2, implying that a 10% increase in user fee should lead to a 2% decrease in the quantity of care received. Demand for care is hence relatively insensitive (or inelastic) to changes in its price.²⁵ The price elasticities of demand were generally similar for both inpatient and outpatient care, except that cost-sharing did not affect inpatient pediatric stays.²⁶

Utilization reductions were almost entirely the result of individuals reducing the number of episodes of care-seeking, rather than the result of decreased utilization per episode of care. The study found that patients were less likely to seek treatment or be hospitalized for their illnesses. The average cost per hospitalization and per ambulatory episode was the same across insurance plans.²⁷ Importantly, however, the study also found that participants randomly assigned to HMOs, with no co-insurance, had significantly lower health care expenditures – similar to those in the high co-insurance plans.

²³ Kenneth B. Wells, Willard G. Manning, Jr., Naihua Duan, Joseph P. Newhouse, and John E. Ware, Jr. (1987), "Cost-Sharing and the Use of General Medical Physicians for Outpatient Mental Health Care" *HSR. Health Services Research*, Vol. 22, No. 1.

²⁴ "Elasticity" is a measure of the responsiveness of demand to changes in price. The investigators assumed *a priori* that all use represented patient "demand."

²⁵ Hence, a health care provider might have an interest in supporting what in the US is termed "balance billing" and in Canada, "extra billing," because the increase in prices would more than offset the reduction in utilization, such that the net effect would be to increase one's income. We shall return to this point below.

²⁶ W.G. Manning, J.P. Newhouse, N. Duan, et al. (1987) "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *American Economic Review*, Vol. 77; pp. 251-277.

²⁷ Emmet B. Keeler, John E. Rolph (1983), "How Cost Sharing Reduced Medical Spending of Participants in the Health Insurance Experiment," *JAMA*, Vol. 249, No. 16.

Was participant health status compromised by the various financing options? The evidence here is mixed. Adults in the free care group experienced better health outcomes than those in the co-insurance group in three areas—diastolic blood pressure²⁸, corrected vision²⁹, and the risk of dying for those at elevated risk³⁰, and were more likely to reduce smoking, maintain a low-salt diet and comply better with medication regimens.³¹ However, their experience was not better on several self-assessed health measures, including physical functioning, role function, social contacts, mental health or health perceptions, and free care was not associated with reduced weight or cholesterol levels.³²

It is worth reminding ourselves how these results have been used by those believing that they provide support for user fees. The popular rhetoric describes cost increases as caused by ever-greater patient demands, many of which may be “frivolous” or unnecessary. User fees will prompt patients selectively to reduce or eliminate these “frivolous” demands. Health care will be not only less costly, but also on average more effective, as the unnecessary services will be pruned away by the discipline of the market. The findings of the RAND co-insurance study, which show significant decreases in utilization with little indication that such decreases compromise health status, seem to support this account. Or do they?³³

In fact, the investigators found that user charges were about equally likely to deter patients from using both necessary and unnecessary services. Cost-sharing did **not** lead to patterns of care seeking that were more “appropriate” from a clinical perspective.³⁴ This finding held true even for hospital admissions. The rate of inappropriate hospital admissions was found to vary by geographic location. Areas with low admission rates

²⁸ R.H. Brook, J.E. Ware Jr., W.H. Rogers, et al. (1983), “Does Free Care Improve Adults’ Health?” *New England Journal of Medicine*, Vol. 309; pp. 1426-1434.

²⁹ N. Lurie, C.J. Kamberg, R.H. Brook, et al. (1989), “How Free Care Improved Vision in the Health Insurance Experiment,” *American Journal of Public Health*, Vol. 79; pp. 640-642.

³⁰ Brook, et al. (1983)

³¹ E.B. Keeler, R.H. Brook, G.A. Goldberg, et al. (1985), “How Free Care Reduced Hypertension in the Health Insurance Experiment,” *Journal of the American Medical Association*, Vol. 254; pp. 1926-1931.

³² Brook, et al. (1983); E.B. Keeler, E.M. Sloss, R.H. Brook, et al. (1987), “Effects of Cost Sharing on Physiologic Health, Health Practices, and Worry,” *Health Services Research*, Vol. 22; pp. 279-306.

³³ Robert G. Evans, Morris L. Barer, and Greg L. Stoddart (1995), ‘User Fees for Health Care: Why a Bad Idea Keeps Coming Back (Or, What’s Health Got to Do With It?)’ *Canadian Journal on Aging / La Revue canadienne du vieillissement* Vol. 14, No. 2; p. 375.

³⁴ K.N. Lohr, R.H. Brook, C.J. Kamberg, et al. (1986), “Effect of Cost-Sharing on Use of Medically Effective and Less Effective Care,” *Medical Care*, Vol. 24, No.9, Supplement.

were not, however, necessarily those with low proportions of inappropriate admissions, nor did cost-sharing selectively reduce inappropriate hospitalizations.³⁵ Contrary to the rhetoric, the RAND evidence demonstrated clearly that patients could not discriminate between necessary and unnecessary services. User fees were as likely to deter important health-preserving or enhancing care, as “frivolous” abuse.

Moreover, the deterrent effects of user charges bear more heavily upon those with lower incomes, as this group is more sensitive to increases in price. In fact, low-income persons reduced use of care that was judged by researchers to be highly effective more frequently than did their higher-income counterparts. This result was particularly evident among children in lower income families; those in the co-insurance group used only 56% as much care deemed to be highly effective as their counterparts in the free care group. For adults, the corresponding figure was 59%.³⁶ Low-income individuals at elevated risk benefited most from free care, and low-income people in poor health who received free care experienced the largest reduction in serious symptoms.³⁷ In conclusion, the greatest beneficiaries of free care were low-income persons with elevated health risks.

Although the RAND experiment showed that patients in the study modified their use of the health care system in response to financial incentives, it has a number of design limitations that compromise its external validity or generalizability. It is not representative of the overall US population, in that it is restricted by both geography and age. Since 1 % of individuals account for about 30% of all expenditures, and 10% account for over half, exclusions by age and disability status have the potential significantly to alter reported results.

The most serious limitation of the RAND study, however, was inherent in its

³⁵Albert L. Siu, Frank A. Sonnenberg, Willard G. Manning, George A. Goldberg, Ellyn S. Bloomfield, Joseph P. Newhouse, and Robert H. Brook (1986), “Inappropriate Use of Hospitals in a Randomized Trial of Health Insurance Plans,” *New England Journal of Medicine*, Vol. 315; pp. 1259-66.

³⁶Lohr, Brook, Kamberg, et al. (1986).

³⁷R.H. Brook, J.E. Ware Jr., W.H. Rogers, et al. (1983) “Does Free Care Improve Adults’ Health?” *New England Journal of Medicine* Vol. 309; pp. 1426-1434; and M.F. Shapiro, J.E. Ware and CD. Sherbourne (1986) “Effects of Cost Sharing on Seeking Care for Serious and Minor Symptoms,” *Annals of Internal Medicine* Vol. 104; pp. 246-251.

design. The study was based on a fundamental assumption that all observed utilization reflected patient "demand" --the independent decisions of patients. This allowed the researchers to presume that the only reactions to user fees would be those of patients--the reactions that their study was designed to measure. That design excluded potential adjustments by providers to changes in their patient loads.

Only a very small percentage of the population in anyone site participated in the HIE and, in consequence, most physicians had a negligible number of (high co-insurance) HIE patients in their practice. Individual hospitals also had only a small proportion of their overall utilization affected by the experiment. The effects of RAND experiment cost-sharing on health care providers' workloads and incomes were insignificant.

The real question is, "What would happen to overall costs if cost-sharing similar to that in the HIE were required of all patients?" The quick answer is that the HIE cannot tell us--because the populations that it excluded, the elderly and the disabled, are not only the heaviest users of care but also the groups whose problems are subject to the greatest clinical uncertainty. There is no basis for generalizing the elasticities from the HIE to this group.

But this is not the principal problem. If the HIE results were applicable to all age groups, health care providers would face a 15~30% decrease in both workloads and gross revenues. There is ample evidence from other sources that physicians do not passively accept such assaults on their practices. With more time available, they can and do change their patterns of practice, seeing a different mix of patients and providing a different mix of services. In particular, the patient groups excluded from the HIE, the elderly and the disabled, are precisely those patients whose conditions would be most amenable to more intensive investigation and follow-up.

How large the aggregate "countervailing" response might be, and what forms it might take, are not questions which the RAND study can answer.³⁸ But to assume that there would be **no** response, as the investigators implicitly do in generalizing from their findings, seems wholly without justification.

³⁸ Greg L. Stoddart, Morris L. Barer, Robert G. Evans (1994), "User Charges, Snares and Delusions: Another Look at the Literature," Toronto: Ontario Premier's Council on Health, Well-being and Social Justice, p.26

Lessons drawn from three natural population-based “experiments” shed some light on this issue. They underscore the fact that patterns of utilization of health care respond to incentives faced by providers, and that it is the attitudes, knowledge and behaviour of providers, rather than patients or so-called “consumers,” that are the key determinants of utilization and costs.³⁹

In 1977, after 25 years of a health insurance plan with first-dollar coverage, members of the United Mine Workers of America Health and Retirement Fund were required to begin making co-payments. Fahs analyzed patterns of utilization for the entire patient population of a large, multi-specialty group practice in western Pennsylvania, for the year before and two years after the introduction of cost-sharing.⁴⁰ More than 80% of the patient population for this practice was drawn from two groups: mineworkers and their families, for whom user fees were introduced, and steelworkers and their families, whose health benefits remained unchanged in the 1976-1979 study period.

Fahs found that "when the economic effects of cost sharing on physician service use are analyzed for all patients within a physician practice, the findings are remarkably different from those of an analysis limited to those patients directly affected by cost-sharing."⁴¹ Physicians responded strategically to the introduction of cost-sharing for their mineworker patients by increasing inpatient lengths of stay and prices for the remainder of their clientele. The physicians recommended longer intervals between follow-up visits for mineworker patients, but shorter intervals for other patients. Overall, expenditures per episode of inpatient care rose by 19% and total expenditures by 7%, for non-mineworker patients. Fahs concluded:

“Thus, the results suggest that increasing cost sharing among large groups of patients may be less effective as a tool to reduce total health care expenditures than has been implied by studies that omit the effect of cost sharing on physician practice patterns. It appears from this

³⁹ For further discussion see Greg L. Stoddart, Morris L. Barer, Robert G. Evans (1994), “User Charges, Snares and Delusions: Another Look at the Literature,” Toronto: Ontario Premier’s Council on Health, Well-being and Social Justice, p.26.

⁴⁰ M.C. Fahs, (1992) “Physician Response to the United Mine Workers’ Cost-Sharing Program: The Other Side of the Coin,” *Health Services Research*, Vol. 27; pp. 25-45.

⁴¹ Fahs (1992); p. 26.

analysis that compensatory actions will be taken by physicians following the reduction in benefits by a large insurance carrier.”⁴²

These findings are largely consistent with those from studies based on the entire populations of two Canadian provinces.

Universal, first-dollar coverage for the costs of physicians' services was introduced in the Canadian provinces at the end of the 1960s. In effect these public programs eliminated user fees for these services, for the entire population of the country, but on different dates in each province, spread over a period of about three years. In the province of Quebec, a research team studied the impact of this system-wide change by conducting pre- and post-introduction interviews in the city of Montreal with both a sample of physicians, and a sample of households.⁴³ The public program began operation in October of 1970; the interviews were conducted between October and May of 1969 and 1971.

The household surveys showed **no** increase in the average number of physician visits per person per year. The mix of visits changed, with office visits reported as increasing and telephone contacts and home visits falling sharply. The survey of physician practices confirmed this change in practice patterns.⁴⁴ But while the total visit rate stayed constant, the household survey showed a substantial change in use across the income distribution. Visits by lower-income people increased; those by higher income people declined. The most marked increase was among the low-income elderly; the most marked drop was among children in higher income families.

Such a pattern might be found if physicians were already working to full capacity. An increase in use, “quantity demanded,” triggered by “free” care could not be accommodated, and so some patients were crowded out as others received more care. But in fact the physician survey found that physicians' self-reported hours of work actually declined after the change, with a particularly marked decline among

⁴² Fahs (1992); p. 39.

⁴³ See P.E. Enterline et al. (1973), “Effects of Free Medical Care on Medical Practice—The Quebec Experience,” *The New England Journal of Medicine*, Vol. 288; pp. 1152-1155; and P.E. Enterline et al. (1973), “The Distribution of Medical Services Before and After ‘Free’ Medical Care—The Quebec Experience,” *The New England Journal of Medicine*, Vol. 289; pp. 1174-1178.

⁴⁴ This may be related to the fact that telephone contacts are not reimbursed in the public program, and home visits are probably under-reimbursed for the time required.

pediatricians. And although the numbers of office visits went up, total patient contacts reported by physicians actually declined (although responses to the survey of patients found no change).

The researchers speculate that physicians were able to maintain their incomes with a shorter workweek, after the change, presumably because all of their patient contacts were now reimbursed. And in fact physician income data for Quebec, based on taxation statistics, do show a large increase immediately after 1970.⁴⁵ Thus while it is always difficult to rule out a hypothetical but unobserved increase in “demand,” the observations seem more consistent with changes in physician behaviour. They appear to have reduced their hours of work and changed the mix of services offered in response to the incentives in the new reimbursement environment. (The authors note that physicians also increased their employment of office nurses, but reduced their use of bookkeepers.)

Whatever the mechanism, however, there was a clear finding that the elimination of user fees was associated with increased use by some, and reduced use by others. How, if at all, did this change relate to needs? In the household survey, the researchers asked about care-seeking for symptoms that they judged, on medical grounds, to be important. There was a very clear pattern of increase, after the removal of user charges, in the proportion of “important” symptoms for which care was sought. Before “free” care, that proportion was lowest in the lowest income class, and rose monotonically as income increased. Afterward, the proportion had increased in all income classes but the highest, and was effectively equal across classes! Out of pocket payment was associated with a clear income class gradient in use of “needed” care; when user fees were removed, that gradient disappeared.

Physicians were asked about the proportions of their patients who had either sought care unnecessarily, or failed to seek care as soon as they should have. (Obviously they could not know about patients who should have sought care but did not, unless future problems developed as a result.) Both changes were in the expected direction, but what is most striking is that both were very small. The proportion of patients seeking

⁴⁵ M.L. Barer, R.G. Evans and R.J. Labelle (1988), “Fee Controls as Cost Control: Tales from the Frozen North” *The Milbank Quarterly*, Vol. 66, No.1; pp. 1-64.

care “without reasonable cause,” as judged by physicians, rose from 1.1 to 1.9%; the proportion of inappropriately delayed visits fell from 2.6 to 1.8%. From the physicians’ perspective, the overall impact on appropriateness of care use seems to be hardly detectable. In addition, even in an environment of “free” care, the ceiling on the proportion of care which could reasonably be construed as patient “abuse,” was very low.

The other provincial experience involved the introduction rather than the removal of user fees, and took place in the context of an already-established universal publicly funded provincial plan. Saskatchewan introduced user fees for physician visits and hospital days in 1968, withdrawing them in 1971. That province's experience demonstrated both the differential impact of such fees on various income classes, and the potential for provider response. Analyses by Beck and Horne of the utilization of physician services before, during and after introduction of mandatory co-payments demonstrated an average co-payment effect of negative 7.7%.⁴⁶ However, this modest global decline was the result of a much larger (approximately 18%) decline in utilization among lower-income groups, coupled with some increases in utilization among higher income groups. The authors concluded that the co-payment effect for physician services “undoubtedly” contained a supply response by physicians to decreased demand by patients, with the consequence that the observed aggregate co-payment effect on utilization “likely understates the behavioral response of consumers to direct charges.”⁴⁷

Moreover, the introduction of user fees was accompanied by a significant increase in the schedule of fees paid by the provincial insurance program. As a result, any reduction in physician incomes arising from a fall in use, was more than offset by increased fees. Furthermore, even the modest initial decline in use dissipated over the period during which the user charges were in place. This would be consistent with a gradual response by suppliers, as the initial fee increase was absorbed.

These examples demonstrate the dangers of uncritical extrapolation from (even experimental) evidence from small samples of patients to conclusions about overall

⁴⁶ R.G. Beck and J.M. Home (1980), “Utilization of Publicly Insured Health Services in Saskatchewan Before, During and After Copayment,” *Medical Care*, Vol. 18; pp. 787-806. Subsequent analysis by Home (unpublished) showed that the hospital per diem user fee had no effect on overall use.

⁴⁷ Beck and Home (1980), pp. 797.

system-level effects. However, this invalid form of inference—one might call it the “RAND fallacy” –occurs repeatedly. Observed behaviour by individuals cannot necessarily be used as a basis for generalizations about “behaviour” at the aggregate system level. Even the architects of the HIE have tended to slide too easily from descriptive reporting of its results to implications for system-wide health care reform.

There is in fact no evidence that user charges can achieve the policy objectives for which they are most often advanced. And, as international evidence on health care spending demonstrates, co-payments are neither a necessary nor a sufficient condition for overall cost control in health care systems. So what do they accomplish? The answer to this question goes some way towards explaining their zombie-like qualities.

The real effects of user fees

The key point to bear in mind is that user fees redistribute the costs and benefits of health care in quite predictable ways. More of the costs fall on those who need and use care, while more of the care goes to those willing and able to pay. Since there is a well-documented inverse relation between income and illness, user fees will shift more of the costs onto those less able to pay.

It is unfortunate that the RAND experiment has been mis-used in the way described above, because it is quite helpful in understanding the responses of individual patients. Like the Quebec experience, it clearly showed that user fees will disproportionately deter care-seeking for necessary services. And like the Saskatchewan and Quebec experiences before, RAND showed that user fees redistribute care use from lower to upper income groups. Those with lower incomes get less care when faced with user fees; they also pay disproportionately more for it.

Even if health and wealth were not correlated, user fees still take a larger share of the incomes of those with lower incomes. The major alternative mode of financing care (even in the United States) is through various forms of taxation, which draw a larger share of contributions from people with higher incomes. The impact of private insurance is more ambiguous.

Taken in total, people pay (all types of) taxes in rough proportion to their incomes, and use health care in rough proportion to their health status or need for care. In general, sicker people use more health care, and richer people pay more taxes. Whether one is a gainer or loser from a user fees policy thus depends on where one is located in the distribution of both income-or, at least, tax liability-and health-or at least, use of health care services. In general, a shift to more user fee financing redistributes net income from lower to higher income people, and from sicker to healthier people.⁴⁸ The wealthy and healthy gain; the poor and sick lose.⁴⁹

These points are clearly in evidence in data from the 1987 National Medical Expenditure Survey in the United States. Figure 6, reproduced from Rasell, Bernstein and Tang (1993), provides a breakdown of family expenditures for health care by form of payment (out-of-pocket, insurance premiums, and taxes), and by income level.⁵⁰ Because the authors have some concerns about the reliability of the income data for the lowest 10% of families, they focus attention on the next highest group (decile #9, which denotes families in the bottom 20% of incomes, but not in the bottom 10%) in comparing low with high income families. They also subdivide the top 10% of families into the top 5% and the next 5%, groups 1a and 1b.

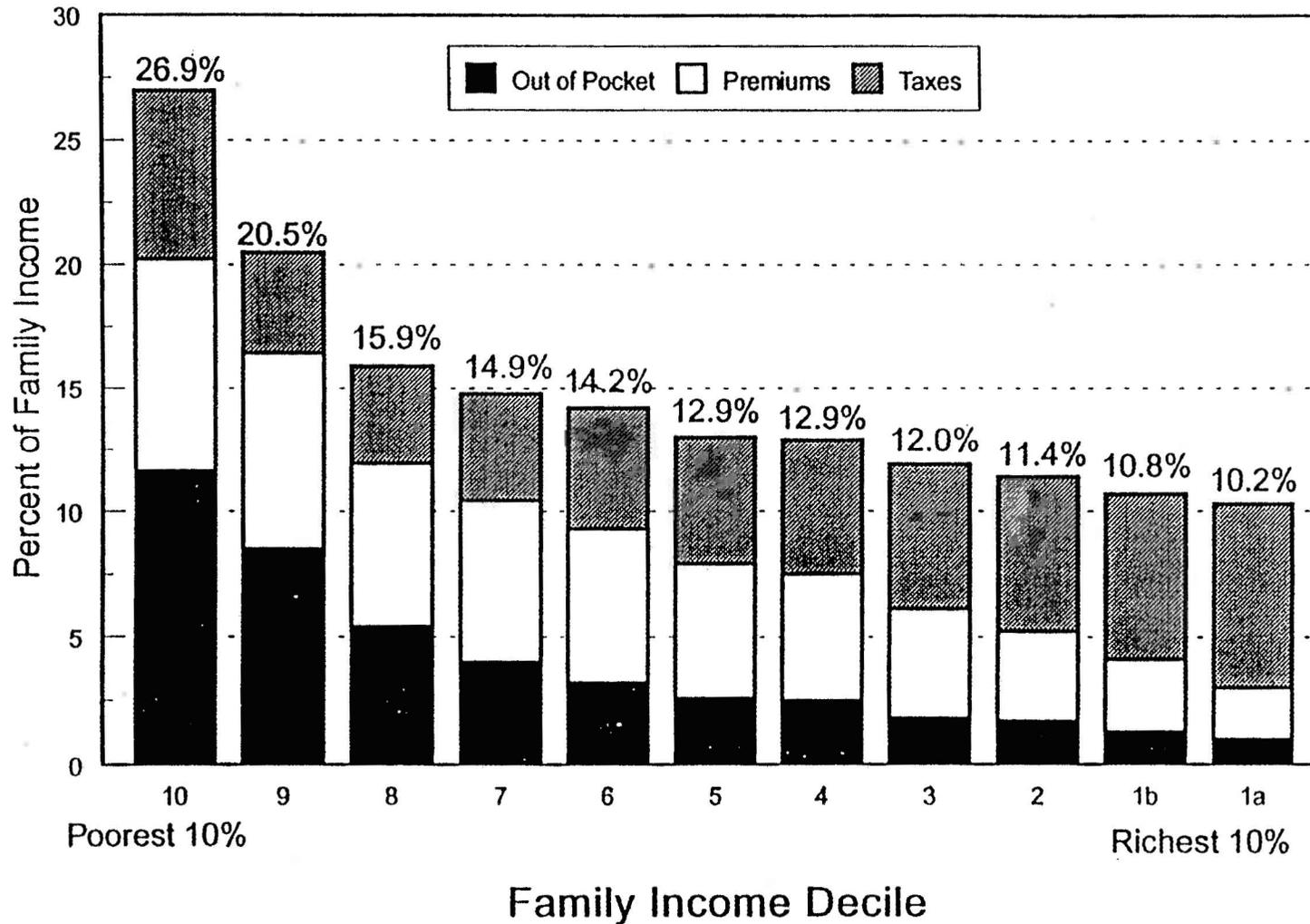
These data show a pronounced and consistent negative relationship between income level and percent of income spent on health care. The pattern is most pronounced for out-of-pocket charges; families in the second lowest decile spend 8.5% of their income on such charges as compared with 1 % for the top 5% of families. But the effect of insurance premiums is almost as regressive, taking 7.9% of income at the ninth decile, and only 2% in the top category, in spite of the fact that many low income people (and

⁴⁸ And the effects of the shift are compounded by the fact that health and income are positively correlated.

⁴⁹ A detailed account of the distributional consequences of substituting more user charges for tax finance is given in R.G. Evans, M.L. Barer & G.L. Stoddart (1994), "Charging Peter to Pay Paul: Accounting for the Financial Effects of User Charges," Toronto: Ontario Premier's Council on Health, Well-being and Social Justice. The authors show generally that an individual gains financially from introduction or increase of user fees if her share of total income exceeds her share of total health care utilization, and that the gain is proportionate to income. In other words, the substitution of user fees for tax finance not only transfers money from the less to the more healthy and wealthy, but if one is wealthy enough, one can be an above average user of care and still come out ahead.

⁵⁰ E. Rasell, J. Bernstein and K. Tang (1993), "The Impact of Health Care Financing on Family Budgets" Economic Policy Institute Briefing Paper (April). Washington, DC: EPI.

FIGURE 6
Family Expenditures for Health Care, 1987
 (by Family Income Decile)



Source: Rasell et al., n.d. (1993)

very few with high incomes) had no insurance coverage.⁵¹ Taxes for health care partially redress the balance. They take a larger share as income increases, but have a much less steep gradient than the other forms of health care spending.

In consequence, Americans at the top of the income distribution spend about half as large a share of their income on health care as those near the bottom, even though they actually spend much more absolutely, in dollar amounts. In fact, total spending rises steadily from \$1,756 per family in the ninth decile (\$960 in the tenth) to \$13,234 in the top 5% of families⁵², and most of this difference is a result of the higher taxes paid by those with higher incomes. As one moves up the income distribution, actual differences in out-of-pocket spending per family are relatively small. However, the more one relies on out-of-pocket payments to finance health care, the greater the relative burden on those at lower incomes; the more one relies on taxation (at least of income and consumption), the greater burden borne by those at the upper end of the income scale.

Similar findings emerge from the large-scale ECUITY project, a detailed empirical study of the pattern of distribution of health, health care, and financing burdens across income classes in the countries of the European Community, as well as the United States.⁵³ The authors conclude:

“The two countries with predominantly private financing systems – Switzerland and the US – have the most regressive structures overall. This is scarcely surprising in view of just how regressive private insurance and out-of-pocket payments are when used to finance such a large proportion of health care expenditures for such a large proportion of the population. The group of countries with the next most regressive systems are the countries operating the so-called social insurance model,...countries which...rely mainly on tax-finance...have the least regressive financing systems.” (p.44)

⁵¹ These data include premiums for the public Medicare program as well as for private insurance. For the average family in the survey, private insurance accounts for 95.7% of the total, \$1,396 as compared with \$63 for Medicare premiums. However, in the bottom two deciles, Medicare accounts for a significant (greater than 10%) portion of total insurance premiums; 15.9% in decile 9, and 23.6% in decile 10.

⁵² Rasell et al., 1993.

⁵³ E. van Doorslaer, A. Wagstaff and F. Rutten (Eds.) (1993), *Equity in the Finance and Delivery of Health Care: An International Perspective*, Oxford: Oxford University Press.

But while they find that “...out-of-pocket payments tend to be a highly regressive means of financing health care...” (p.42), the impact of private insurance is more nuanced. Private insurance that is purchased as supplementary cover in a more or less universal public system appears to be a “luxury” that is more commonly bought by people with higher incomes; such payments are thus progressively distributed.

But if, as in the United States, private insurance is purchased by a large proportion of the population because public coverage is restricted or non-existent, the distribution of its costs is highly regressive. Private coverage for deductibles or co-payments in the public system tends to be progressive or regressive depending upon the scale of such charges. In general the larger the user charges and the more people who must pay them, the more regressive are the costs of private insurance to cover them. As noted above, this is what Rasell *et al.* found.

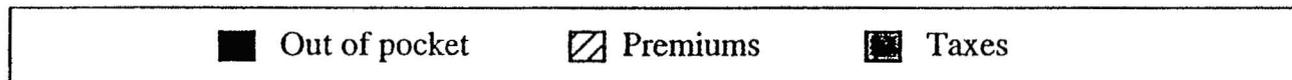
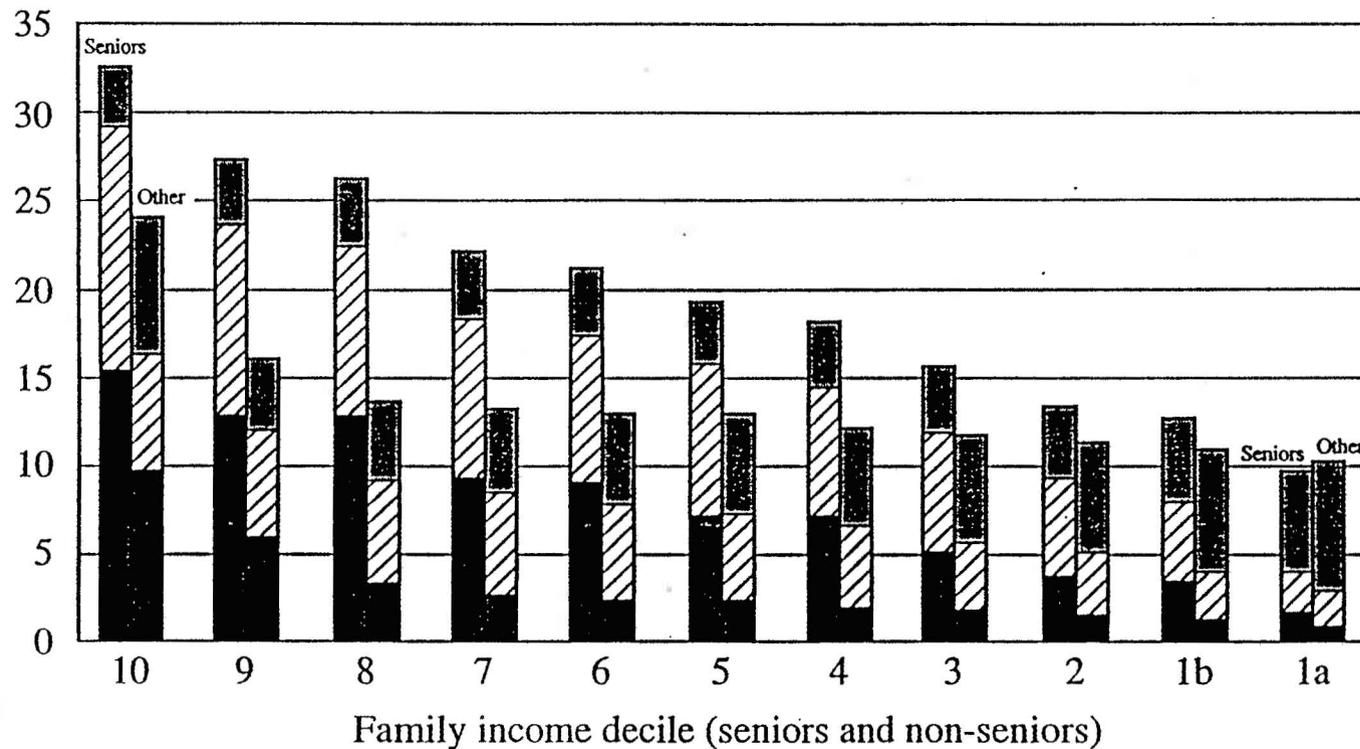
In Figure 7 the Rasell *et al.* data are further disaggregated by age of household, into categories of those over and under 65. The regressive nature of the funding structure for health care in the United States is even more marked among the elderly, despite the fact that the federal Medicare program offers virtually universal coverage to everyone over 65. The Medicare program requires people to pay substantial deductibles and co-insurance; as a result most elderly Americans purchase private "Medigap" insurance to cover some part of these costs. On average, elderly American families in 1987 paid \$1,407 for insurance premiums--almost exactly the same as the \$1,471 paid by non-elderly families--to fill in the holes opened in their "universal" public insurance program by user fees! Although private insurance expenditures rise with income, there is surprisingly little difference at each income level between the insurance outlays of elderly and non-elderly families. Yet, because the average income of elderly families is lower, these outlays represented a greater share of their income, 8.3%, as compared with 4.9% for non-elderly persons.

Despite the combination of Medicare and private insurance, elderly Americans still spent, on average \$1,239 or 8.3% of their incomes in out-of-pocket payments for

FIGURE 7

FAMILY EXPENDITURES FOR HEALTH CARE BY INCOME DECILE AND AGE OF HEAD

Percent of family income



Source: Rasell et al., n.d. (1993)

health care, as compared with \$712 or 3.2% for the non-elderly. Moreover, the relative burden is much greater for those in the lower income groups. Seniors in the second-lowest decile spent 12.7% out of pocket, compared with 1.5% in the highest income group. This heavy burden of private health care spending by the elderly has no counterpart in Canada.⁵⁴

In both countries, the combination of out of pocket payment and privately purchased health insurance take up an increasing proportion of consumer spending as the household (head) ages. But the percentage in Canada rises from 1.5% for the youngest Canadian families to 3.2% for the oldest; in the United States it starts at 2.8% and rises to 17.1 % (Figure 8)! The much greater reliance on patient cost sharing in the American system places a substantially heavier financial burden on those with the greatest needs.⁵⁵ In fact, the cost-shifting effects of user fees are so obvious, once one thinks about them, that there has been very little debate on the subject. Rather, the controversy and empirical analysis have focused on other (non-distributional) effects which user fees are alleged to have on the functioning of health care systems.

But it is not just patients and payers who have an economic stake in user fees. Although the conventional economic story emphasizes the putative effects of such charges in lowering overall costs, most of their advocates have quite different effects in mind.⁵⁶ Providers of care consistently advocate user fees because they believe, and explicitly state, that such fees are a way of **increasing** total costs by opening up new sources of revenue.⁵⁷ For providers, they are the answer to chronically “underfunded” public systems. Since all expenditures on health care are, as a matter of basic accounting, equal to the total incomes of providers, their interests are completely transparent. It is hard to see why they would so consistently advocate such fees if the standard economic story were valid. But as the evidence above indicates, it is not.

The persistence of the topic of user fees on the health care policy agenda—its zombie-like ability to re-emerge from the realm of intellectually dead ideas—is readily

⁵⁴ Barer, et al. “Old and Sick,” 1992.

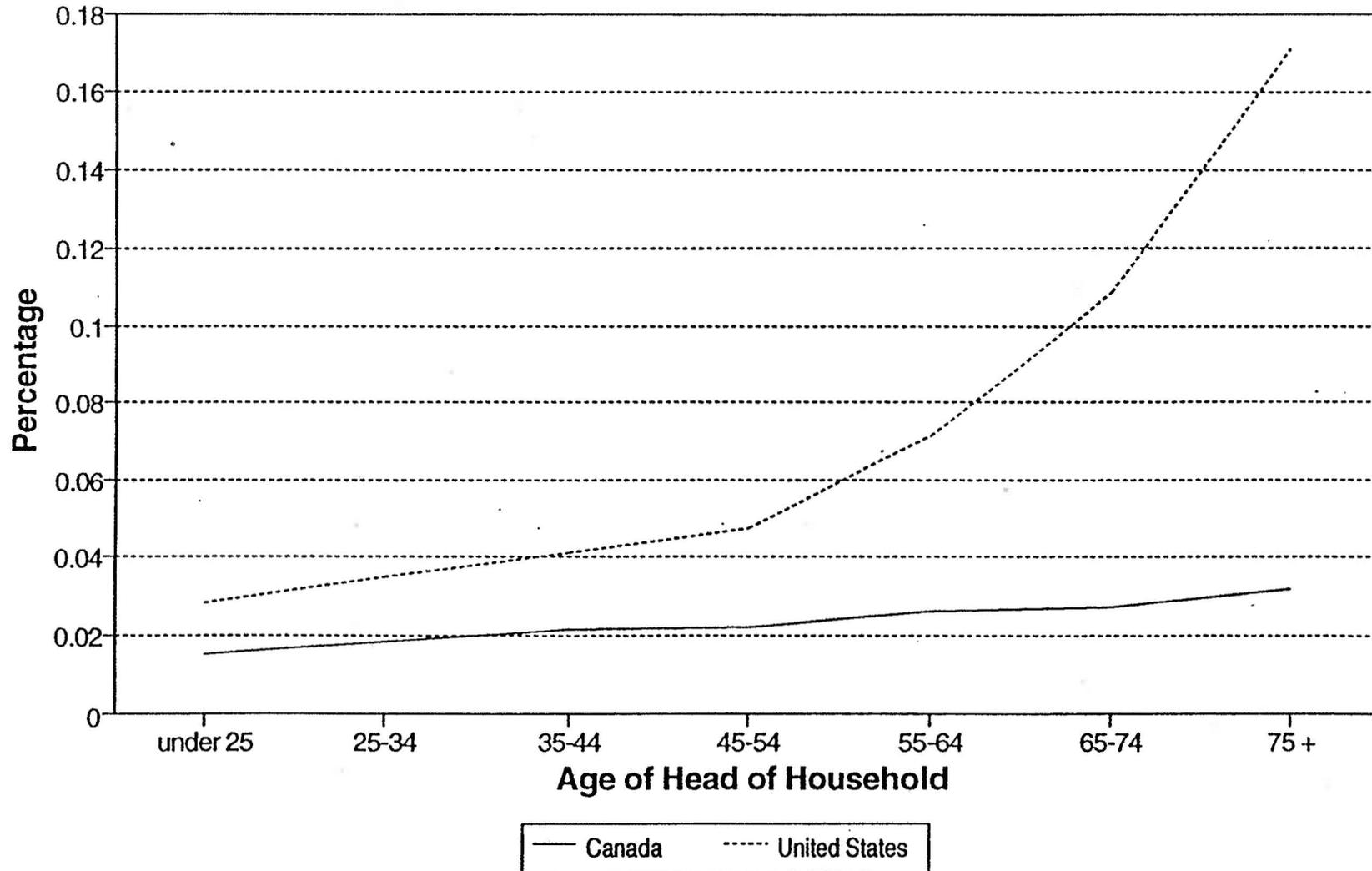
⁵⁵ B.B. Torrey and E. Jacobs (1993), “More Than Loose Change: Household Health Spending in the United States and Canada” *Health Affairs*, Vol. 12, No.1; pp. 126.31.

⁵⁶ Stoddart et al., Why Not User Charges...?

⁵⁷ Barer et al., The Remarkable Tenacity.

FIGURE 8

Household Spending on Health Care As Share of Personal Consumption, 1986



Source: Torrey and Jacobs (1993)

understandable when one appreciates the number of people who do stand to gain from the introduction or expansion of such fees. There are vested financial interests here, and the potential winners are more powerful and concentrated than are the potential losers.

Moreover, user fees allow governments to shift some of the burden of financing health care out of the public domain and onto the private sector. In this era of fiscal restraint, any opportunity for cost-shifting is almost irresistible. However, there is a price to be paid: to the extent that government turns health care over to the private sector, it loses control over the issues of who receives care, how much, and at what price. It also relinquishes control of the only levers that have been successfully demonstrated anywhere to control the overall share of aggregate economic resources given over to health care.

Zombie # 3: “You can't get something for nothing—covering America's uninsured would add substantially to costs that are already very high.”

The principal task of the Clinton health care reform initiative in the United States was to provide secure and portable insurance coverage—portable across geography, employment and the life cycle—to ease the anxiety of the middle class. There is no politically practical way to accomplish this without also providing coverage for the fifteen percent of the population that is currently uninsured, and upgrading benefits for those with inadequate coverage.

US discussions of health care reform have focused on the cost implications of extending coverage; it has generally been taken for granted that this would be “very costly.” If the average cost of care in the US is \$3,500, and there are 40 million uninsured, then their coverage will cost an additional \$140 billion. Full coverage for the under-insured would add substantially to this estimate. An important corollary of this notion, which also feeds the fears of the insured, is that one cannot reduce costs without reducing coverage.

Once the notion that extending coverage will inevitably be very costly has been solidly implanted, two routes appear to be open. One may increase coverage despite the added cost. This means that additional revenues will need to be raised, and, most

crucially, the American public will need to be convinced of the need for assuming the added burden. This was the issue that dominated the Clinton health reform debate. The critical question was how to get to (only) 17% of gross domestic product (GDP) by the turn of the century, rather than whether such a level of spending was reasonable, necessary, or inevitable. Alternatively, one may decide that such a cost increase is not supportable, and that therefore the United States cannot afford universal coverage. One interpretation of the failure of the Clinton health reform efforts is simply that the American public was not prepared to assume this spending increase.

But the idea that extending coverage to all *must* lead to cost expansion is another zombie. It serves to promote a series of interrelated messages: about the inevitability of cost-increases were coverage to be increased; about threats to security of coverage were expenditures to be reduced; and, somewhat paradoxically, about the notion that the richest nation on the face of the earth cannot afford to ensure that all of its population has health insurance. These arguments are rooted in a particular set of assumptions about how the health care system will look after reform.⁵⁸ They persist because they are promoted by powerful, concentrated interests with a great deal to lose from any substantial system change, and because they have a superficial plausibility that resonates widely with the American public.

As illustrated by the simple calculation above, estimates of the likely costs of universal coverage are generally derived by focusing on revenue requirements. But if extending coverage really does require \$140 billion or more dollars in additional health care revenue, what will happen to expenditures and incomes? Both total expenditures, and provider incomes, must also rise by \$140 billion. Existing incomes would rise, new incomes would be created, or both. So why were so many organizations representing these potential income recipients so determined to reject this shower of gold?

Perhaps they recognized that the story is not quite this simple. In fact, there are a number of compelling reasons for believing that extending coverage need not create a

⁵⁸ A more detailed critique of these assumptions can be found in Morris L. Barer, Robert G. Evans, Matthew Holt, and J. Ian Morrison (1994), "It Ain't Necessarily So: The Cost Implications of U.S. Health Care Reform," *Health Affairs*, Vol. 13, No.4; pp.88-99.

windfall for providers. But if it does not, then it follows that the revenue requirements are also overstated.

1. International experience

Perhaps the strongest *prima facie* evidence that extending coverage need not increase costs comes from international experience. The US health care system has both disproportionately higher costs and lower coverage than any other health care system in the GECD. Ironically, over the last two decades, the United States has managed to reduce the share of its population with coverage, while dramatically increasing its costs.

How do these other countries provide coverage for all while holding costs to levels well below those found in the United States? The common thread in the international approaches to cost containment is some form of monopsony control over health care budgets, or at least their most significant segments. Collective decisions are made about total expenditures, through controlling financing mechanisms or available sources of revenue. This in turn must imply control over total incomes, although it does not necessarily fix the numbers of income recipients, or the levels of individual incomes. In fact, each country has adopted its own unique type of control over the number and mix of income recipients, and over the level of expenditures in particular sectors. But the common theme remains the use of budgets or other controls to establish spending limits.

2. A closer look at the assumptions in those cost estimates

One need not rely on the experience of others to question the assumption that extending coverage must increase costs. The reasoning underlying this claim is that extending coverage will result in a significant increase in the use of services, which will in turn require a proportionate increase in overall expenditures. The idea is encapsulated by Mark Blumberg:

Changes in use of health services by the uninsured, when covered after health reform, are a key to the costs of reform...[i]f uninsured persons obtained

private coverage....their visits to physicians' offices would increase 52 percent.⁵⁹

Such projections, however, embody two separate sets of assumptions about the impact of increased coverage, that it will increase aggregate use of care, **and** that it will have no effect on the price of care. The first is questionable; the second is plain wrong.

Taking the second point first, it is notorious that the administrative costs of the fragmented private insurance system in the United States add a very large amount to the total cost of health care – far larger than in any other country. Current estimates of the excess administrative burden, relative to a universal, single-payer system such as that in Canada, are of the order of over \$100 billion per year, or more than 10% of total health care costs.⁶⁰ This is sufficient to cover the whole cost of providing care to the currently uninsured population.

All of these excess administrative costs are reflected in the higher prices of health care services in the United States. Accordingly, rationalizing the payment system on the lines of that found in every other developed society would free up sufficient funds to extend coverage to the entire population, with little or not impact on total costs. Service volumes would rise, but their average costs would fall.

Such a change would, however, reduce the incomes of the insurance and financial services industry—by about \$100 billion. Accordingly, while no one has seriously questioned these estimates, strenuous and successful efforts have been made by the insurance industry—and more generally by specialists in management and finance— to focus attention elsewhere.

But if extension of coverage, through a universal public program, has the potential to transfer \$100 billion, more or less, from the paper pushers to the care givers, why do not the latter, from all professions, rally behind the cause? The explanation is again found in the extraordinarily high prices of health care in the United States. In part these reflect the higher costs of administrative overhead, but they also reflect the bargaining

⁵⁹ M.S. Blumberg (1994), "Impact of Extending Health Care Coverage to the Uninsured" *Health Affairs*, Vol. 13, No.4; pp. 181-92.

⁶⁰ These include both the costs of the insurance and prepayment system itself, and the additional costs generated in hospitals, clinics, and physicians' offices in coping with that system.

power of providers.

The radical reform that would eliminate private insurers would almost certainly leave providers face to face with powerful monopsony funders. In the public systems in other countries, governments or other public agencies have become very hard bargainers, on behalf of their publics, over fee levels and increasingly over total expenditures.⁶¹ Since US expenditures are already so far out of line in international terms, providers may quite reasonably fear that a more rational system, instead of turning over to them another \$100 billion in incomes, would instead push overall costs down.

But providers now find themselves on the horns of a dilemma. Private, for-profit payers are learning the techniques of the public sector monopsonists, and are able to apply them much more rigorously in the absence of political restraints or public scrutiny. Clinicians are beginning to see their incomes squeezed, while the administrative sector grows and shareholder profits fatten. This is precisely the reverse of the potential scenario above - the transfer of income from clinical to administrative functions rather than the other way. They have chosen to resist public sector reform; their reward appears to be increasingly aggressive private sector control.

These fundamental realities have, however, been largely crowded out of the public debate by the micro-estimations of the likely effects of coverage extension on use by the currently uninsured. Most such estimates assume that those without insurance will, when they have it, adopt the care seeking patterns of those with adequate insurance and that this will increase overall use and costs of health care services by a corresponding amount.

Those now without coverage, however, generally receive care when they are in serious need-albeit less care than those with insurance. In fact, because they are usually sicker than those with coverage by the time they do access the system, their health care needs, per episode of illness, are greater than they would have been if they had had coverage. This utilization is already reflected in existing health care costs, although it may not be accounted for explicitly. Public institutions already receive a variety of forms of public and private subsidy, part of which is used to pay staff and purchase the supplies

⁶¹ On the Canadian experience, see M.L. Barer (1991), "Controlling Medical Care Costs in Canada", *Journal of the American Medical Association*, Vol. 265, No. 18; pp. 2393-4; and M.L. Barer, J. Lomas and C. Sanmartin (1996), "Re-minding our Ps and Qs: Medical Cost Controls in Canada", *Health Affairs* Vol. 15, No.2; pp. 216-234.

and services necessary for the provision of care to those unable to pay (“uncompensated care”). Other patients cross-subsidize these expenses through higher prices; the taxpayer pays higher taxes.

Accordingly it would be highly misleading to assume that use by the currently uninsured is zero, and therefore that it would increase by the full amount of care for the currently insured. Indeed the argument that the uninsured are not in fact going without care is often used by defenders of the current American system.

The impact of increased coverage on use by the uninsured themselves could nonetheless be substantial. The Congressional Budget Office (CBO), for example, estimated that their use might rise by as much as 60%.⁶² However, since the uninsured make up only 15% of the population, this apparently large figure translates into an aggregate increase of around 5%.⁶³ This is not particularly imposing relative to historical rates of expansion in US health care costs.

And this estimate may be biased upwards as well, as it oversimplifies the real world in several crucial respects. These estimates are based in one way or another on information drawn from analyses of partial populations. They are, in other words, generated using observations of utilization based on situations in which only a small proportion of patients and providers are affected. As we have previously pointed out, it is a simple (and common) fallacy of composition to generalize from such situations to population estimates -- another instance of the “RAND fallacy.” Changes at the population level may be substantially different from what is suggested by partial analyses.

Consider the following hypothetical situation. Providers are fully employed (no excess capacity), and use by those without insurance is substantially less than that by those with insurance. If we now extend coverage to the uninsured, the implications for overall use are very different from what partial analyses might suggest. In this case there

⁶² United States, Congressional Budget Office (1993), *Behavioral Assumptions for Estimating the Effects of Health Care Proposals*, Staff Memorandum, Washington, DC: CBO.

⁶³ For example, if one assumes that per capita costs for the uninsured, are, say \$2,000 per annum, and would rise to \$3,200; that the per capita costs for those with full coverage are \$3,500 and would remain so; that the uninsured represent 15% of the population; and that the US population is 250 million, then the overall impact of the increase from \$2,000 to \$3,200 is about 5.5%. If one assumes per capita costs for the uninsured are lower, the overall impact of the 60% increase is lower. But even if extending coverage brings the uninsured to the \$3,500 level, the overall increase is only 6%.

might be a redistribution of services used, but aggregate service use would not increase because the supply side would already be at “full employment.”

More generally, the impact of extended coverage on **total** use will inevitably depend critically upon conditions and behaviour on the supply side of the health care system. Like the RAND experiment, virtually all of these micro estimates of the impact of extending coverage simply slide right by this fundamental reality.

Whatever else ensues, however, extending coverage to the uninsured *would* require substantial changes, in a direction opposite to the way the American health care system is currently evolving. Such a substantial restructuring would threaten powerful corporate and shareholder interests, particularly those in the insurance and financial management sectors. The stakes are high, and so the lengths to which those interests will go to ensure the *status quo*, or at least evolution under favourable terms and conditions, should surprise no one.

Extending coverage to the un- and under-insured in the United States need not increase costs. Indeed if it were part of a more general reform, costs would most probably (on the international evidence) decline. But a betting person might be excused for insisting on odds. Those who would lose from such a reform—recalling that all expenditures make up someone's income—have been remarkably successful in focusing attention on the “unacceptable costs” of extending coverage—the zombie.

Zombie # 4: “America's Health Care Does Not Come Cheap – But the Best in the World is Worth Paying For.”

The United States spends far more (per capita) on health care than any other country in the world. What does the US get for what it spends? A widespread—and heavily promoted—assumption is that, in health care as in other walks of life, you get what you pay for. Greater health care expenditures lead to better health; any reductions would put this at risk. Or so say the promoters of this zombie.

It should then follow that Americans receive the highest quality health care in the world, enjoy unparalleled good health, and are the happiest with their system. In fact

Americans appear to be less healthy than people in other comparable nations, and to be less satisfied with their system. So what *do* they get, for all the extra spending?

If more spending is to lead to better health, it is necessary and critical (though not sufficient) that more services are being provided to patients. But expenditures on health care, as on everything else, are the product of *prices* as well as quantities. What the US primarily purchases for its higher levels of spending is, it turns out, higher priced services rather than more care. Moreover, much of the care that Americans do receive is of questionable benefit to their health.

A number of recent cross-national studies have explored the differences in health care spending between the United States and other countries. The most comprehensive compared per capita expenditures and use of services across all OBCD countries, using two different approaches to estimating relative levels of use in each country.⁶⁴ Per capita levels of expenditure in each country were converted to US dollars using exchange rates based on purchasing power parities (or “PPPs”).⁶⁵ Purchasing power parities are estimates of the costs, in each country, of a specified basket of commodities. This study used two different baskets of commodities, one representative of the whole of the national product and the second of a market basket specific to the health care sector.

Figure 9 shows health care spending per capita in a number of OBCD countries, converted into US dollars using each of these alternative “PPPs” and then divided by the corresponding US value. The first bar of each country's pair represents that country's level of (per capita) health care expenditure, relative to that in the US, when adjusted for the general level of prices in that country relative to the US. The second bar, however, shows the result of adjusting for relative price levels specific to the health care sector of each country.

When the conversion is made using an economy-wide market basket (so-called “GDP PPPs”), US health care spending per capita was indeed much higher than that in all other countries. When spending is adjusted using prices of the health care related basket

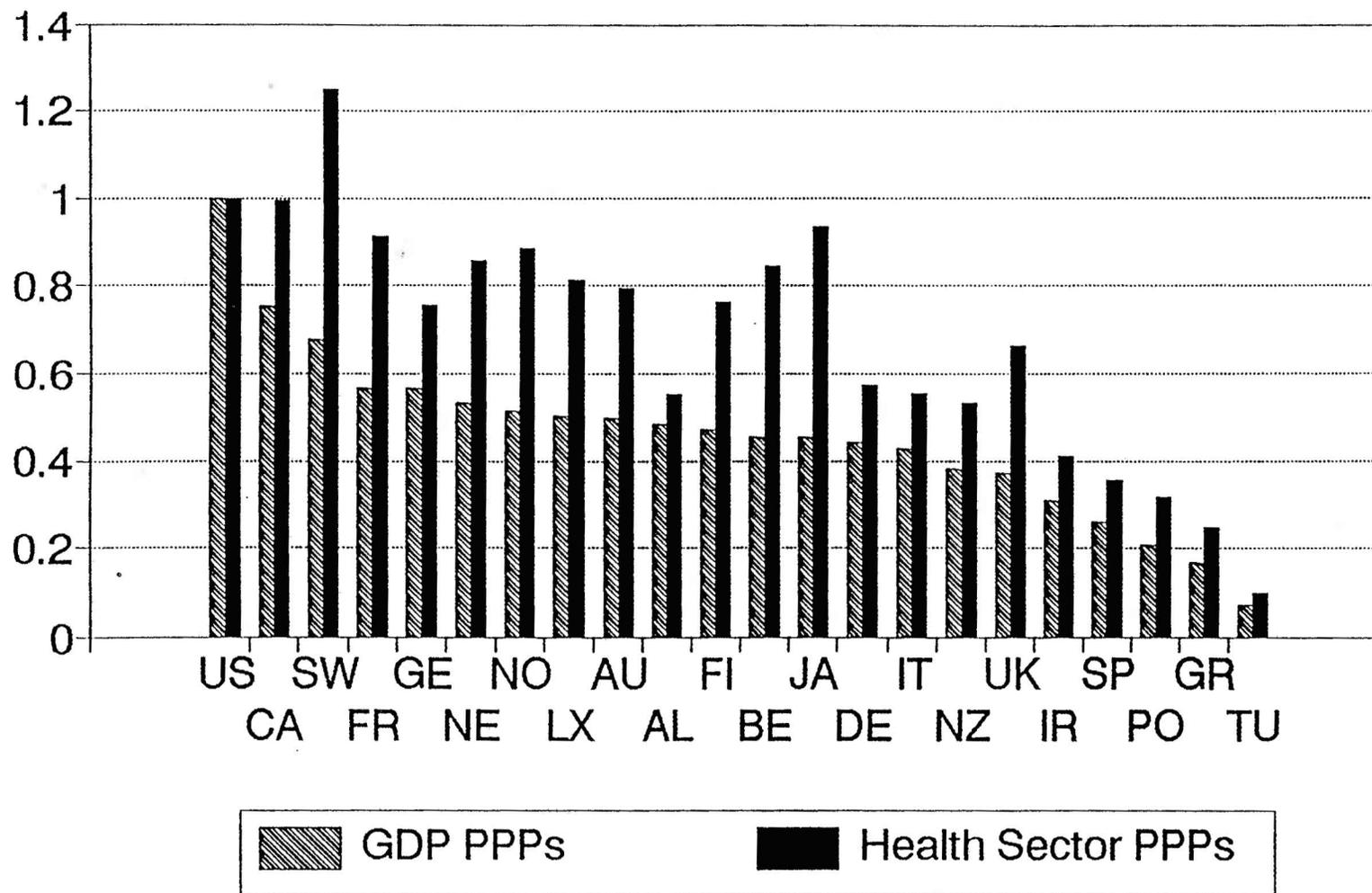
⁶⁴ Ulf-G. Gerdtham and Bengt Jonsson (1991), “Price and quantity in international comparisons of health care expenditure” *Applied Economics*, Vol. 23; pp. 1519-28.

⁶⁵ Since currency values change, and exchange rates often have more to do with international trade flows than with the ability to purchase goods within a country, purchasing power parities are generally used in preference to exchange rates for the purposes of international comparisons.

FIGURE 9

Health Care Spending per Capita, 1985

As % of US, in Purchasing Power Parity



Source: Gerdham and Jönsson (1991)

of goods—that is to say, using health care sector specific adjustment—it turns out that other countries used comparable, and sometimes greater, amounts of health care. The level of prices in the health care sector, relative to prices in other sectors, is higher in the United States than in any other country examined.⁶⁶

Japan, for example, spent less than half of what the US spent per capita in the health care sector when measured in terms of GDP PPPs, but more than 90% of what the US spent per capita when measured using health care sector PPPs. Similarly, Sweden spent less than two thirds of what the US spent per capita measured in terms of GDP PPPs, but spent almost one quarter *more* when expenditures were compared using health care sector PPPs. And Canada spent just over three fourths of what the US spent in terms of GDP PPPs, but spent exactly the same amount on health care per capita when converted using health care sector PPPs. Most of the extra spending in the US was thus not purchasing more health care services—Americans just pay higher prices.

Several studies comparing spending patterns in the United States and Canada corroborate the finding that health care prices are higher in the United States. Fuchs and Hahn⁶⁷ estimated that expenditures per capita for physicians’ services in the US were 1.72 times the Canadian level in 1985 (1.75 in 1987). Patients in the United States were found to receive significantly *fewer* physicians’ services—on average, only 72% of those received by their Canadian counterparts. But American fees are much higher. Their findings are summarized in Table 4.

Confirmation is provided by Welch, Katz and Zuckerman (1993), who found that Canadian physicians’ fees in 1992 were, on average, 59% of US Medicare fees.⁶⁸ Moreover, based on 1990 data, Urban Institute researchers estimated US Medicare fees to average 76% of private “submitted charges.”⁶⁹ Taken together, these data suggest

⁶⁶ Although the information is not included in this graph, relative prices in the health care sector are higher in the US than in every country but the Mediterranean countries such as Turkey, Greece and Portugal, where relative prices of goods outside the health care sector are quite low.

⁶⁷ Victor R. Fuchs and James S. Hahn (1990), “How Does Canada Do It? A Comparison of Expenditures for Physicians’ Services in the United States and Canada” *New England Journal of Medicine*, Vol. 323, No. 13; pp. 884-90.

⁶⁸ W.P. Welch, S.J. Katz and S. Zuckerman (1993), “Physician Fee Levels: Medicare Versus Canada” *Health Care Financing Review*, Vol. 14, No.3; pp. 41-54.

⁶⁹ M.E. Miller, S. Zuckerman and M. Gates (1993), “How do Medicare Physician Fees Compare with Private Payers?” *Health Care Financing Review*, Vol. 14, No.3; pp. 25-39.

Canadian fees were, at the time, approximately 45% of private submitted charges in the US. Canadians thus spend less on physicians' services, and receive more.⁷⁰

These aggregate differences tell only part of the story. Fuchs and Hahn (Table 5) found that submitted fees for surgical services in the United States were 3.21 times the comparable Canadian rates. Anesthesiology and radiology services cost 3.73 and 3.59 times their Canadian equivalents. The weighted average of physicians' fees charged for all procedures was 3.34 times the Canadian rate. Consultative services also cost about 60% more in the US than in Canada. Overall fees for physician services averaged 2.39 times higher in the United States, a result remarkably consistent with data emerging from the work of the Urban Institute.

But higher physicians' fees are not the only, or even the largest, factor underlying higher US expenditures. Fuchs and Redelmeier⁷¹ examined the differences in hospital costs between Canada and the United States, and found that expenditures per admission were about forty percent higher in the latter (Table 6). The patient load in American hospitals was, on average, somewhat more complex, but much of the difference remained even after adjusting for differences in case-mix. Furthermore, the difference in input prices between the two nations is remarkably small, and cannot explain the discrepancy in costs per admission. There are, it would seem, simply more things being done per (comparable) patient in the United States. The crucial question is what sorts of things are these, and what contribution do they make to health?

Woolhandler and Himmelstein (1991) shed some light on this question.⁷² They examined 1987 spending per capita in the United States and Canada on various categories of health administration (Table 7). For each of the categories examined, spending per capita was greater in the US. Spending on insurance administration, for example,

⁷⁰ One might argue that submitted charges by American physicians did not represent fees actually paid to physicians, because there was a significant amount of discounting involved in US physician services reimbursement. As a result it was difficult to determine the relationship between reported prices and transaction prices. Although we would not dispute that this was going on, it is important to recall that US Medicare fees represent an approximate lower bound on the discounting process. This would still leave American fees well above Canadian levels.

⁷¹ Donald A Redelmeier and Victor R. Fuchs (1993), "Hospital Expenditures in the United States and Canada" *New England Journal of Medicine*, Vol. 328, No. 11; pp. 772-78

⁷² Steffie Woolhandler and David U. Himmelstein (1991), "The Deteriorating Administrative Efficiency of the US Health Care System," *New England Journal of Medicine*, Vol. 324, No. 18; pp. 1253-8

Table 4

Estimated Relative Physician Fees and Services, United States vs Canada, 1985 and 1987.

VARIABLE	1985 RATIO OF UNITED STATES TO CANADA	1985 RATIO OF IOWA TO MANITOBA	1987 RATIO OF UNITED STATES TO CANADA
Expenditures per capita	1.72	1.51	1.75
Fees	2.39	2.18	2.61
Quantity of services per capita	0.72	0.69	0.67
Prices of resources	1.30	1.43	1.32
Ratio of quantity of resources to quantity of services	1.84	1.53	1.98

SOURCE: Fuchs and Hahn, *New England Journal of Medicine*, Sept. 27, 1990

Physicians' Fees, 1985

Table 5

SERVICE	RATIO OF UNITED STATES TO CANADA	RATIO OF IOWA TO MANITOBA
Surgery	3.21	2.76
Anesthesiology	3.73	2.86
Radiology	3.59	4.19
Procedures (weighted average)	3.34	2.99
Moderate office visit	1.56	1.44
Extensive office visit	1.55	1.50
Moderate hospital visit	4.77	3.56
Extensive hospital visit	2.57	2.70
Consultation	1.60	1.64
Evaluation and management (weighted average)	1.82	1.72
All services	2.39	2.18
Net income per physician	1.35	1.61

SOURCE: Fuchs and Hahn, *New England Journal of Medicine*, Sept. 27, 1990

Accounting for Ratios of Expenditures
per Admission in the United States and Canada, 1987

Table 6

	Ratio, U.S. to Canada	Ratio, California to Ontario
Expenditures per admission	1.39	1.63
Diagnostic case mix	1.14	1.11
Expenditures per adjusted admission	1.22	1.46
Prices of resources	1.04	1.05
Resources per adjusted admission	1.17	1.39
Inpatient share of expenditures	1.06	1.05
Inpatient resources per adjusted admission	1.24	1.46

SOURCE: Redelmeier and Fuchs, *New England Journal of Medicine*, 1993

Cost of Health Care Administration in the
United States and Canada, 1987

Table 7

COST CATEGORY	SPENDING PER CAPITA	
	U.S.	CANADA
Insurance administration	106	17
Hospital administration	162	50
Nursing home administration	26	9
Physicians' overhead and billing expenses		
Expense-based estimate	203	80
Personnel-based estimate	106	41
Total costs of health care administration		
High estimate	497	156
Low estimate	400	117

SOURCE: Woolhandler and Himmelstein, *New England Journal of Medicine*, 1991.

averaged \$106 per person in the United States versus \$17 in Canada. For hospital administration, the figures were \$162 versus \$50; for nursing home administration \$26 and \$9; and there were also significant differences for physicians' overhead and billing expenses. Overall, the total cost of health care administration is significantly higher in the US than in Canada (Figure 10). Woolhandler and Himmelstein estimated total per capita costs of health care administration in Canada to fall somewhere in the range of \$117-156, while those in the US fell between \$400-497.

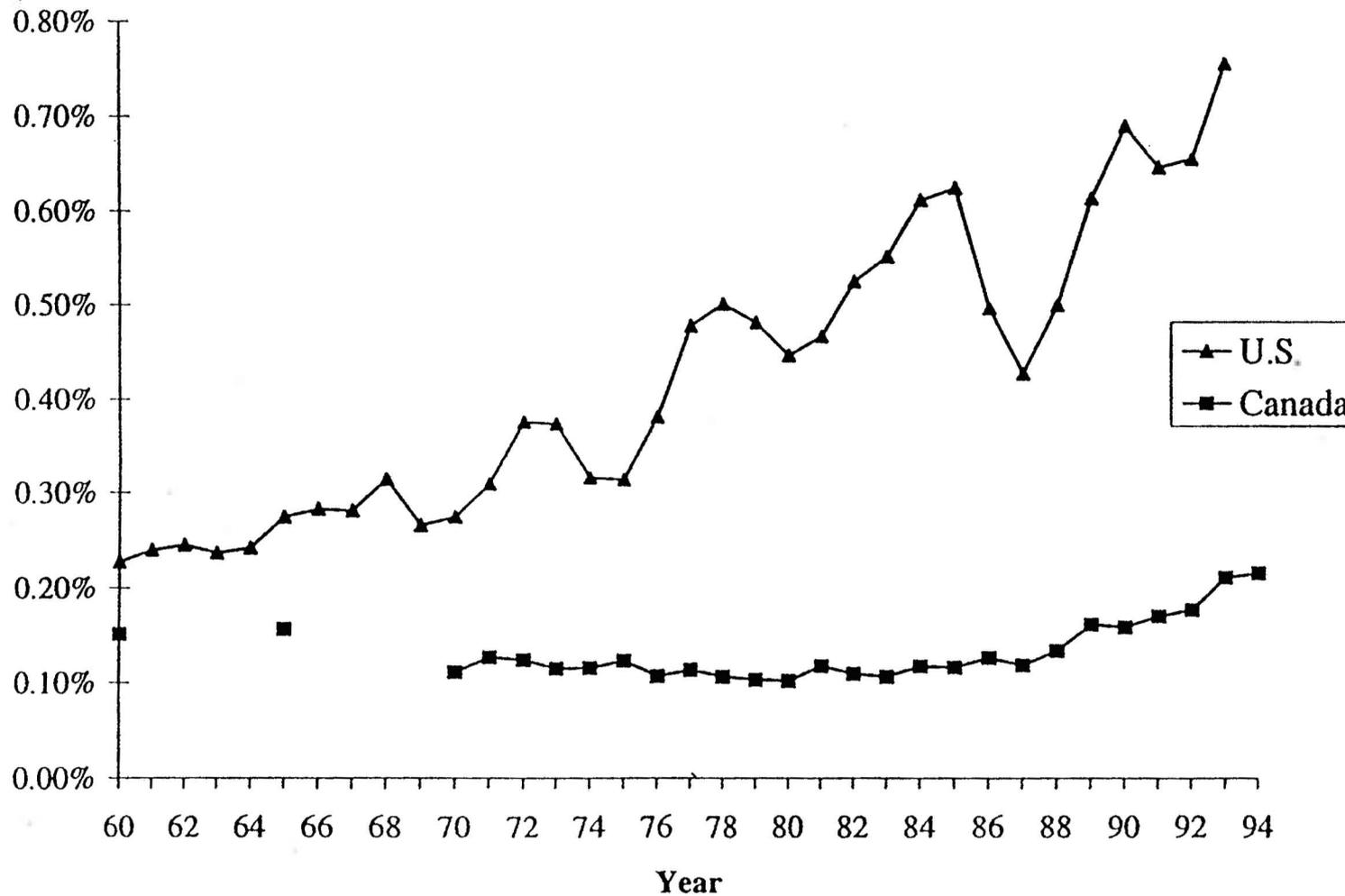
The inescapable conclusion is that there is a lot more paper pushing going on in the US. And no one has yet offered evidence in support of the hypothesis that pushing more paper produces greater health. There are occasionally claims that the extra administrative effort yields benefits in the form of more appropriate or higher quality care. But these have not been substantiated, and the complaints of clinicians trying to cope with a plethora of different and ever-changing benefits managers provide sobering counterpoint.

In sum, you *do* get what you pay for. But what you get may not be what you think you're getting. Americans pay more for health care primarily because they pay higher prices. And they pay higher prices partly because providers themselves enjoy higher fees and incomes, and partly because in the US "providers" include a much larger number of people being paid for non-clinical services. The administrative overhead costs of the payment system—which are hidden in service prices—are higher than anywhere else in the world.

The zombie is the simple-minded notion of a direct connection between health care expenditures and health outcomes—"gee, if we don't spend enough we're all going to die." Its persistence is attributable both to the role of vested interests in perpetuating the myth, and to the fact that the idea is on the surface a plausible one, and is even seen in many quarters as obviously true. The role of interests should be relatively clear. There are those who win, and win big, from current American expenditure patterns. Each dollar of expenditure is simultaneously a dollar of income for someone. The present mode of organizing and financing the delivery of health care in the United States creates jobs, incomes, and profits. Physicians and other clinical staff, hospital administrators and

FIGURE 10

**Costs of Prepayment and Administration
as a Share of GDP, Canada and U.S.**



Source: Health Canada, National Health Expenditures in Canada, various years; HCFA National Health Expenditure Estimates.

financial support staff, insurance agents, marketing analysts, managers and consultants, and shareholders of various corporations—all have a stake in maintaining and increasing expenditures in this sector. Their representatives are well organized, and acutely aware of where their interests lie. Quite understandably, they work hard to promote ideas that support those interests.

But the phenomenon of popular resonance is more complex, and is relevant not just to this zombie, but to all. Why is it that some ideas seem instantly credible, apparently independent of their evidentiary basis, while others never take hold in the popular imagination? We return to this puzzle after describing one final zombie—perhaps the biggest of them all.

Zombie #5: The health of a population depends primarily upon its health care; more care is the surest route to better health

This zombie overlaps to some degree with #4 above, that more (less) expenditure necessarily leads to better (worse) health. The flaw identified in that claim was that higher costs, at least in the United States, are largely associated with higher prices rather than more care. But that observation leaves open the question of the relationship between the level of health care—whatever its cost—and the health of the population served. This final zombie embodies the notion that more care necessarily leads to better health, and moreover that it is the only, or at least the most effective, route.

The emphasis placed by modern societies on the production and distribution of health care, and the resources, attention and effort devoted to this task, gives testimony to the widespread acceptance of this belief. In all modern states, the health care industry is one of the largest clusters of economic activity.⁷³ The truly remarkable achievements of modern medicine also serve to reinforce the impression created by this story; namely, that health care is the most important determinant of health.

⁷³ G.J. Schieber and J.P. Poullier (1989), “Overview of International Comparisons of Health Care Expenditures,” *Health Care Financing Review* (Annual Supplement): pp. 1-7; and OECD Secretariat (1989), “Health Care Expenditure and Other Data: An International Compendium from the Organization for Economic Cooperation and Development” *Health Care Financing Review* (Annual Supplement): pp. 111-94.

This impression persists in the face of a vast and growing body of contradictory evidence, accumulating over the last thirty years, that demonstrates the lack of a relationship between the amount of care provided to a population, and either the needs or the health outcomes of that population. The Dartmouth Health Atlas of 1997 is but the latest and most comprehensive example of such observations.⁷⁴

Our intention here, however, is not to re-review this well-known literature. Rather we focus on a less extensive and younger, but also rapidly expanding body of evidence demonstrating the importance to the health of populations of factors other than health care. The sources of better health turn out to be remarkably complex, and are as yet far from fully understood. The simple story—that more health requires more health care—is more easily understood, and wrong. It is another zombie, and as such there are good and understandable reasons for its persistence. But they have nothing to do with the evidence.

Human populations display a marked diversity in their patterns of health and disease. This diversity is not due simply to differences among individuals, which one could expect to be (more or less) randomly distributed.⁷⁵ Rather, there are significant differences in terms of health status between entire populations, or among sub-groups of the same population, which can consistently be correlated with other distinguishing characteristics of these groups, such as income level, educational attainment, and social class.⁷⁶

During the nineteenth and early twentieth centuries, human life expectancy in industrialized countries increased from approximately 40 to over 60 years. It is commonly thought that medical interventions were the principal factors driving this trend.

⁷⁴ J.E. Wennberg, M. McAndrew, et al. (1997), *The Dartmouth Atlas of Health Care* 1998. Hanover, New Hampshire: Dartmouth Medical School. Centre for the Evaluative Clinical Sciences.

⁷⁵ More or less, because (1) there may be some groups which differ from each other genetically, such as males and females, in ways which lead to differing health status. Our account makes no presumption that under ideal conditions such differences would disappear. However, comparisons of populations which are relatively genetically homogeneous also show non-random distributions of health status; and (2) for some conditions, the hypothesis that illness generates group affiliation (e.g. chronic mental illness and low-income status) may be plausible.

⁷⁶ A more complete account is given in C. Hertzman, J. Frank, and R.G. Evans (1994), "Heterogeneities in Health Status," in *Why Are Some People Healthy and Others Not?* Robert G. Evans, Morris L. Barer, Theodore R. Marmor (Eds.), Hawthorne, NY: Aldine de Gruyter; pp. 67-92. See also the introductory chapter to the same volume.

But the major decline in mortality from most infectious diseases took place prior to the development of effective medical therapy.⁷⁷ As shown in Figures 11 and 12 the greatest portion of the decline in annual death rates from tuberculosis in the UK occurred well before successful medical interventions (chemotherapy and BCG vaccinations) were introduced. Similar patterns have been found for most other infectious diseases.

This is not intended to suggest that medical therapy played no role. Although the decline in mortality from most infectious diseases predated effective therapy, once effective measures were developed they did indeed play an important part in accelerating the rate of decline. The decrease in mortality from tuberculosis, for example, was about 50% during the 1940s. But medical treatment neither initiated the decline in mortality, nor represented the most significant factor in that decline.

There is an obvious correlation between declining mortality, and the increasing wealth and living standards of industrialized populations. Nutritional deprivation, contaminated water supplies, overcrowding, poor sanitation and other aspects of daily living environments are by-products of poor economic conditions with direct health consequences. Improve those conditions, and health improves. Improve the economic circumstances of a population, and those conditions improve.

But this account of the relationship between health and wealth is also too simple. Appreciation of its true complexity has been increasing over the past several decades. Early on in the century, the association appeared to be a straightforward monotonic linkage between wealth and health. In international comparisons, for example, it was clear that life expectancies were low in low income countries and high in high income countries, and that they rose over time in each country along with average incomes. In more recent years, however, the cross-sectional relationship has broken down. Above a certain level of average income, currently surpassed by the world's richest nations, there is no apparent relation between a nation's average income level and the most commonly used measures of health status. Within each country, however, rising average incomes continue to be associated with improving health. And across the population within each country, health status is still closely correlated with income.

⁷⁷ T. McKeown (1979), *The Role of Medicine: Dream, Mirage or Nemesis?* 2nd Edition, Oxford: Basil Blackwell.

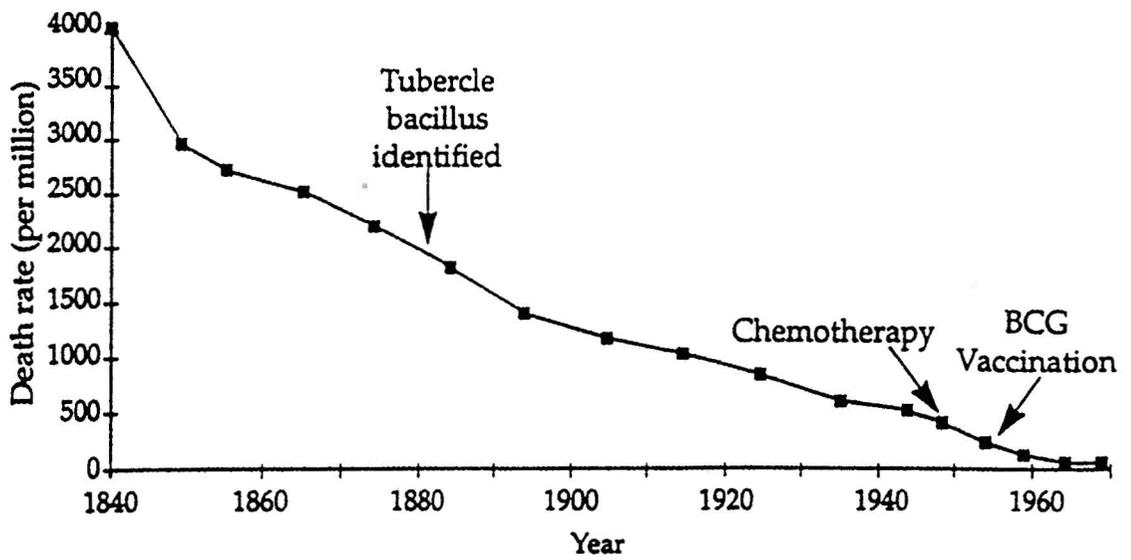


Figure 11 Respiratory tuberculosis: mean annual death rates (standardized to 1901 population), England and Wales, 1840–1970.
Source: McKeown (1979)

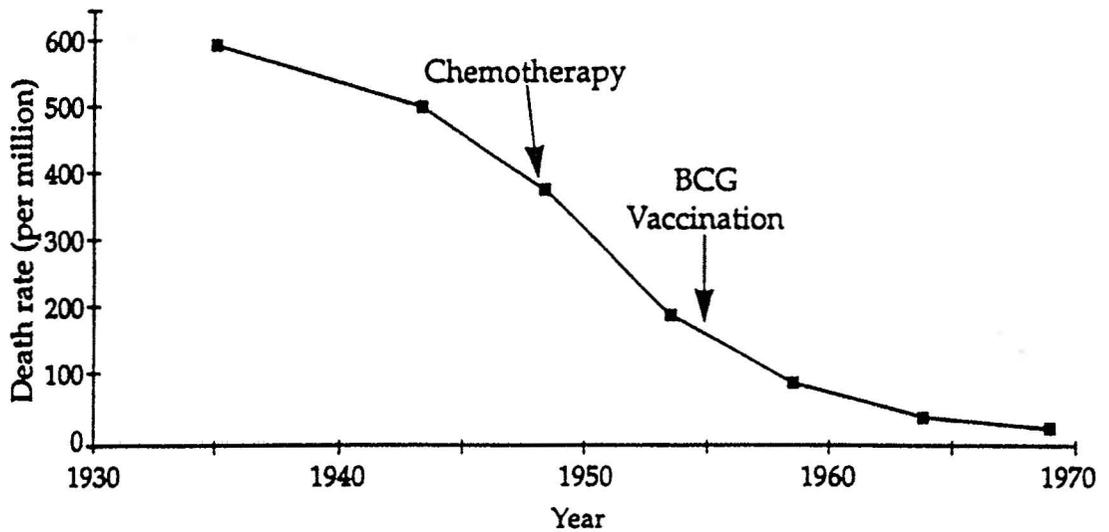


Figure 12 Respiratory tuberculosis: mean annual death rates (standardized to 1901 population), England and Wales, 1935–1970.
Source: McKeown (1979)

A line of research pioneered by Wilkinson⁷⁸ suggests a potential reconciliation. At lower levels of income, health status is primarily determined by the material circumstances of life--adequacy of food and shelter, and protection from infections and environmental hazards. But past some threshold level, material conditions become less important and the distribution of further increments of wealth becomes a dominant factor. Wilkinson, and more recently other researchers, find evidence that in wealthier societies, greater wealth is neither necessary nor sufficient for improved health. Whether it makes a positive contribution appears to have a great deal to do with how that added wealth is distributed. Inequality *per se* leads to illness.

Among countries at the top of the income scale there is a prominent relationship between life expectancy and the income distribution characteristics of a country. Cross-sectional evidence shows lower mortality rates in countries with a more egalitarian distribution of income (Figure 13).⁷⁹ A large part (about two-thirds) of the variation in life expectancy is accounted for by variation in the proportion of income going to those below the seventh income decile. Variation in average income levels (gross national product per head) contributes less than 10% to the proportion of the variance explained. Similar findings have emerged from studies of different regions within the United States.⁸⁰

The relationship appears to hold up when examined dynamically as well. Wilkinson compared changes in income distribution over time in different countries with changes in life expectancy. Increases in the share of income going to the least well off were associated with faster increases in average life expectancy (Figure 14). It also appears that all income classes may be healthier in societies with more even income distributions. Having systems in place that buffer the emergence of inequality may be good not only for the vulnerable, but also for the privileged.

⁷⁸ A recent synthesis of this work, which dates back over a decade, can be found in R.G. Wilkinson (1994), *Unfair Shares*, Ilford: Bamarado's.

⁷⁹ R.G. Wilkinson (1992), "Income Distribution and Life Expectancy," *British Medical Journal*, Vol. 304; pp. 1651-168.

⁸⁰ G.A. Kaplan, E.R. Pamuk, J.W. Lynch, et al. (1996), "Inequality in income and mortality in the United States: Analysis of mortality and potential pathways", *British Medical Journal*, Vol. 312; pp. 999-1003; B.P. Kennedy, I. Kawachi, and D. Prothrow-Stith (1996), "Income distribution and mortality: Cross-sectional ecologic study of the Robin Hood index in the United States", *British Medical Journal*, Vol. 312; pp. 1004-1007.

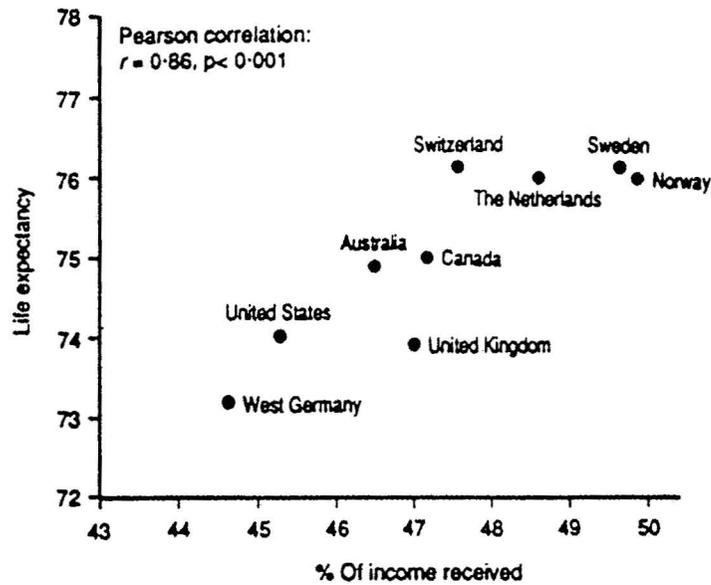


FIG 13-Relation between life expectancy at birth (male and female combined) and percentage of post tax and benefit income received by least well off 70% of families, 1981

Source: Wilkinson (1992) P.165-168

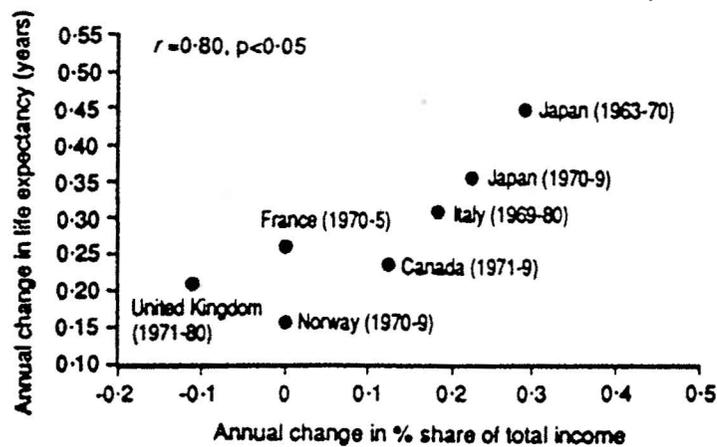


FIG 14-Annual change in life expectancy and percentage of income received by least well off 60% of population. (Two figures for Japan were combined when calculating correlation coefficient)

Source: Wilkinson (1992) P.165-168

Wilkinson offers the contrasting recent experiences of Britain and Japan as a particularly dramatic illustration of these relationships. In 1970 these countries were fairly similar in terms of both income distribution and life expectancy. However, Japan now has the highest life expectancy in the world. Marmot and Davey Smith found no obvious explanation, such as changing diet or availability of health services, for this rapid improvement in Japanese life expectancy.⁸¹ They did, however, observe that Japan now has the most egalitarian income distribution of any OECD country. Moreover, the tendency for mortality to fall most rapidly among the upper classes in Britain, but among the lower classes in Japan, is consistent with the notion that these divergent trends in mortality are related to changing socioeconomic differentials.

Powerful evidence of a link between socioeconomic status and health comes from the Whitehall studies, that have followed more than ten thousand British civil servants for nearly two decades. An extensive array of information, including health status measures, has been accumulated on each of the participants. The hierarchy of income and rank in the British civil service is particularly well-defined. These studies, being longitudinal, offer important insights into causality and patterns of illness over the life cycle.

The first phase of the Whitehall study gave rise to two important findings. First, there was a strong link between status and health (Figure 15).⁸² The (age-standardized) mortality rate, calculated among males aged forty to sixty-four over a ten-year period, was about three and a half times as high for those in the lowest grades, as for those in the senior administrative grades.⁸³ Second, there was an unmistakable *gradient* in mortality from top to bottom of the hierarchy. Mortality was significantly higher in the second classification of professional and executive personnel, than in the top administrative grade, and increased further as one went down the scale.

This result is inconsistent with a simple model of health as determined by access to health care. Britain offers universal health care to the entire population through the

⁸¹ M. Marmot and G. Davey Smith (1989), "Why Are the Japanese Living Longer?", *British Medical Journal*, Vol. 299; pp. 1547-51

⁸² M.G. Marmot (1986), "Social Inequalities in Mortality: The Social Environment" in R.G. Wilkinson (Ed.), *Class and Health: Research and Longitudinal Data*, London: Tavistock; pp. 21-33.

⁸³ M.G. Marmot and T. Theorell (1988), "Social Class and Cardiovascular Disease: The Contribution of Work," *International Journal of Health Services*, Vol. 18; pp. 659-74.

National Health Service. Nor is it adequately explained by poverty and material deprivation, since none of these groups of British civil servants could be characterized as impoverished or deprived in any common sense of these terms. All are employed, most in office jobs with low occupational health risks (which are, moreover, generally comparable to those senior to them), and the professional and executive grades are relatively well-paid as compared to the population as a whole.

Furthermore, the gradient is not satisfactorily explained by differences in individual risk factors. While the risk factors for coronary heart disease (smoking, high blood pressure, and elevated blood cholesterol) were more prevalent among people in lower ranks, these explained very little of the variation in CHD mortality between grades (Figure 16).⁸⁴ There is something correlated with hierarchy or status that powerfully influences CHD mortality, yet that lies beyond the standard explanations of disease causality. Similar gradients were observed for most (but not all) causes of death.

A number of other studies of differential mortality by socioeconomic status have also shown that patterns of mortality in relation to social class seem to express themselves through different diseases over time. Whatever the mix of diseases accounting for mortality, the gradients persist. The Black report, for example, provides data on mortality rates by social class in the UK over most of the twentieth century, on a decade by decade basis. These data too show a gradient.⁸⁵ But what is perhaps most interesting is that they show the gradient persisting, with little change, over most of the period since data were first collected in 1911, and apparently increasing in recent years.

Yet the principal causes of death have changed radically during that period. At the beginning of the century, infectious diseases were the big killers. At the end of the century, the major killers are chronic conditions such as heart disease and cancer. But at both times, (age-standardized) mortality rates are higher in the lower classes. Over this

⁸⁴ M.G. Marmot, G. Rose, M.J. Shipley and P.J.S. Hamilton (1978), "Employment Grade and Coronary Heart Disease in British Civil Servants" *Journal of Epidemiology and Community Health*, Vol. 32; pp. 244-249.

⁸⁵ D. Black, J.N. Morris, C. Smith, P. Townsend, and M. Whitehead (1988), *Inequalities in Health: The Black Report*, P. Townsend and N. Davidson (Eds.), Middlesex: Penguin. See also R.G. Wilkinson (1986), "Socio-economic Differences in Mortality: Interpreting the Data and Their Size and Trends," in *Class and Health*. R.G. Wilkinson (Ed.), London: Tavistock; pp. 1-20.

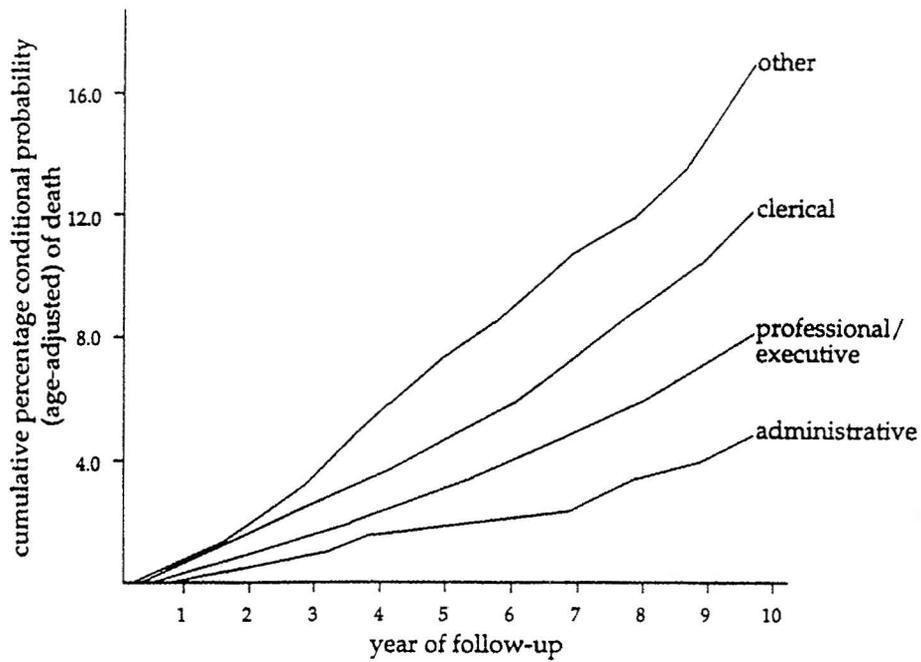


Figure 15 Whitehall study: all-cause mortality among total population by year of follow-up.

Source: Marmot (1986:23)

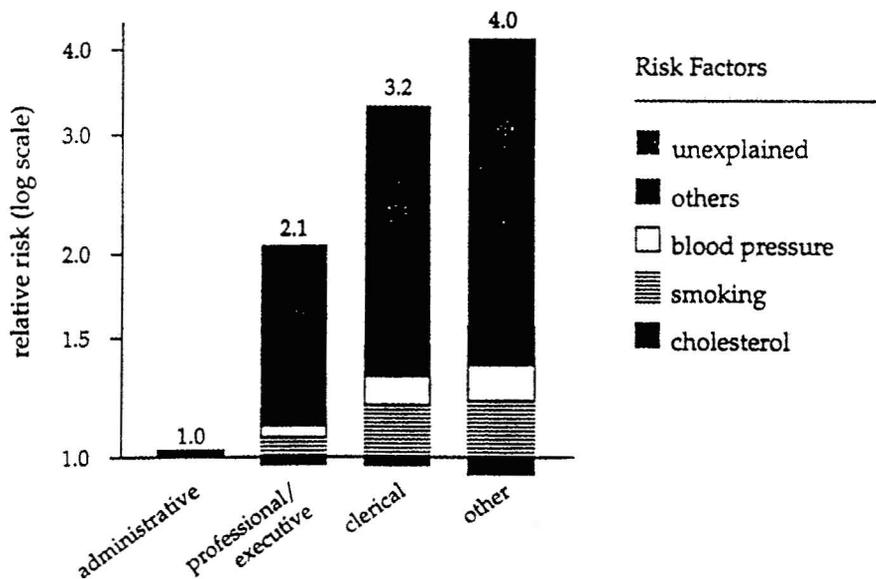


Figure 16 Relative risk of CHD death in different grades "explained" by risk factors (age-standardized).

Source: Marmot et al. (1978:248)

period the scale and sophistication of the health care services have vastly increased, and the NHS, established at mid-century eliminated economic barriers to access. These observations suggest that approaches to the study of disease which look at individuals in isolation from their social environments are likely to miss this underlying process.

These insights derived from longitudinal studies of the evolution of health status are consistent with those found in a number of cross-sectional analyses in the countries of the OECD and the former Communist Bloc. During the last 25 years of the Cold War, life expectancy⁸⁶ in Central and Eastern Europe (CEE) fell behind that in the countries of the Organization for Economic Cooperation and Development (OECD). Life expectancy in the OECD countries has continued to rise, while that in CEE countries has remained static or even declined.

Examination of the trend line for the CEE shows three distinct phases. During the first, lasting until the mid-1960s, life expectancy for these two groups of nations converged to near equality, due principally to rapidly declining infant mortality rates in the CEE. The second phase, dating from the mid-1960s to 1989, witnessed the reemergence of differentials in survival rates, as a result of growing differences in mortality from chronic diseases in mid-life. Male life expectancy in the CEE underwent no sustained increases after age 30, and there were even modest declines, while life expectancy for men in the wealthy societies of the OECD rose. Female life expectancy after age 30 rose in the CEE over the same period, but by much less than in the nations of the OECD.

A third phase, dating from 1989 to the present, is particularly intriguing. It marks a significant decline in life expectancy in most countries in Central and Eastern Europe. The pattern has been most apparent in Russia and some of the other Newly Independent States, but has also shown up to a lesser degree in Central Europe. Adults under 60 years old, and particularly unmarried males, appear to be most vulnerable.

The reasons for these divergent experiences are not yet clear. Differences in health

⁸⁶ Mortality and its corollary, life expectancy, are used as the principal indicators of health status in these analyses, because mortality has historically been the only health status indicator routinely collected in comparable fashion in both OECD and CEE nations. It has the added attraction of not being influenced by cultural context, as self-reported measures of health status tend to be.

care, environmental pollution, and greater prevalence of individual risk factors such as smoking and poor diet, have all been proposed. Although these variables appear to be a part of the story, much of the difference appears to be rooted in the rapidly diverging social and economic circumstances. A review of the evidence bearing on each of the competing hypotheses will make this clear.

1. “Historical inevitability”

The historical inevitability hypothesis asserts that patterns of life expectancy in the CEE lag behind those of OECD nations by about a generation, for reasons rooted in human biology. Proponents observe that differences in life expectancy between CEE and OECD countries have been the norm for over a century. On this basis, they regard the phase of convergence in life expectancy after World War II, rather than the (re)emergence in recent decades of the differences, as the anomaly.

This anomaly is then explained as a historical artifact, the product of two simultaneous trends. The explanation runs as follows: in the post-war period of apparent trend convergence, the CEE experienced a decline in infant mortality similar to the one that had occurred earlier in the countries of the OECD. However, it did so at a time when chronic disease mortality was increasing in the West. The differential was re-established when the countries of the OECD entered a period of declining mortality from chronic disease, while the CEE, inevitably, started to experience increased mortality from chronic diseases.

This explanation is implausible, as there is no sound reason to believe that rapid declines in infant mortality in a given society are biologically incompatible with declines in chronic disease mortality in mid-life. The assumption is, in fact refuted by the example of Japan, where these two trends occurred simultaneously. Comparisons of the evolution of life expectancy between the Baltic countries and Finland also provide an important counter-example to the historical inevitability hypothesis. The example of Finland is of particular significance, as its historical experience closely paralleled that of the Baltic states until the end of World War II. Finland was part of the Czarist empire until 1917,

and enjoyed a brief period of independence between World Wars I and II. At this point, unlike the Baltic countries, Finland was not incorporated into the Soviet Union.

This makes the Finnish case a sort of natural experiment in life expectancy, with all important variables except the nature of the political and economic system held constant. Life expectancies in these countries were similar until the late 1970s and early 1980s, when Finland, which developed the pattern of the OECD in life expectancy, pulled ahead of the Baltic countries, which followed the CEE pattern. This seems to show, in contrast to the thesis of historical inevitability, that the causes of the life expectancy gap are not inevitable.

2. “Physical environment”

This approach takes as its point of departure the observation that the gap in CEE-OECD life expectancy re-emerged around the time that environmentalism and worker health and safety began to have an impact on the public policy agenda in the OECD nations. In contrast, inhabitants of CEE countries were exposed to increasingly polluted air, water, soil, and food. The physical environment theory hypothesizes that these exposures are not only the proximate causes of specific forms of mortality, such as asthma and chronic lead poisoning, but are also the direct causes of mortality from respiratory diseases, certain cancers, congenital anomalies, and perhaps, cardiovascular diseases. Its proponents also note that living in a highly polluted environment is damaging to a population's sense of well-being, and to its citizens' feelings of being in control of their lives.

As it fits well into our existing models of health and disease causation, this explanation finds a ready reception. Moreover, it has a clear measure of evidentiary support. Available air pollution monitoring data seem compatible with this theory, as relative changes in air quality over time between CEE and OECD nations parallel the re-emergence of the life expectancy gap. Data from OECD cities between 1970 and 1985 show that sulfur dioxide concentrations in urban air dropped precipitously. During the same period, however, sulfur dioxide concentrations tripled in certain parts of the CEE.

Considerable evidence has accumulated to show that in heavily polluted regions, the physical environment **was** a significant contributor to ill health.⁸⁷ For example, life expectancies in rural areas of Poland exceeded those in urban areas over the 25 years prior to political liberalization in 1989. This highly unusual demographic pattern was more pronounced in the province of Katowice, the most polluted region of the country. However, the trend toward declining life expectancy during what we have called the second and third phases took place in both heavily polluted and relatively unpolluted parts of the CEE, a fact which puts a ceiling on the potential importance of environmental pollution in explaining the emerging gap with the west.

3. “Health care and the economy”

This approach postulates that, despite its large supply of health care professionals, the CEE was unable to match advancing western standards in diagnosis and treatment of the chronic diseases that are the principal causes of mortality. More specifically, the political rigidities and weak economies of the CEE did not permit the necessary local autonomy, professional independence and strategic capital formation fundamental to a successful health care system. In this, their health care systems simply paralleled their more general inability to make the transition from an industrial to a service-information society. These problems were exacerbated by a shortage of hard currency in the CEE, which became increasingly necessary to purchase medical equipment and drugs not available from CEE producers.

This approach sounds plausible to western ears, since it links health status to advances in health care, and general economic failure to “communism.” But in fact it accounts for little of the life-expectancy gap: CEE-OECD differences in case-fatality rates for medically avoidable causes of death are not large. Calculations based upon the historical evolution of rates of medically avoidable causes of death for these two groups

⁸⁷ Estimates of the contribution of pollutants to mortality through airborne dust and sulfur dioxide are in the range of 4 to 22% for heavily polluted regions. This forms an upper limit to estimates of the contribution of pollution to the life expectancy gap; in many regions of the CEE, air pollution is no greater than in OECD nations. See C. Hertzman (1995), “Health and Environmental Pollution”, Chapter 3 in *Environment and Health in Central and Eastern Europe*, Report for the Environmental Action Programme for Central and Eastern Europe, Washington, D.C.: World Bank.

of countries show that no more than 9 to 15% of the life expectancy gap can be attributed to less effective health care services in the CEE.⁸⁸

4. The differential “risk factor” approach:

This approach begins from the observation that a large share of the life-expectancy gap is attributable to a single family of causes of death: heart disease, stroke, and circulatory disorders. It is then hypothesized that the life expectancy differentials can be explained by differences in the established risk factors (smoking habits, blood pressure control, and diet) for these conditions.

While this hypothesis may seem intuitively plausible, especially since so much emphasis is placed on “lifestyles” in the western popular media, it has not been supported by the evidence. To the contrary, a series of international comparisons have failed to show a strong relationship between the established risk factors and international differences in heart disease mortality. In fact, these differences are better predicted by differences in gross domestic product *per capita*.⁸⁹

Of particular importance is a 15-year prospective cohort study that compared several thousand men in Kaunas, Lithuania, and Rotterdam during the Cold War period. Mortality was higher in Kaunas than in Rotterdam, as one might expect. But the established risk factors for heart disease, which were measured at the beginning of the study, did not predict this difference.⁹⁰

Finally there is no evidence that one of the major risk factors, smoking behaviour, was in fact higher among citizens of the CEE during the 1960s and 1970s, when the life expectancy gap first emerged. Among women in particular, smoking rates have historically been low, and they remain lower than those in the OECD. On the basis of this particular risk factor differential one would expect lower, not higher, rates of mortality in the CEE.

⁸⁸ C. Hertzman, S. Kelly and M. Bobak, Eds. (1996), *East-West Life Expectancy Gap in Europe: Environmental and Non-environmental Determinants*, London: Kluwer Academic Publishers.

⁸⁹ M. Marmot and M. Bobak, personal communication, 1997.

⁹⁰ H. Bosma (1994), *A Cross-Cultural Comparison of the Role of Some Psycho-Social Factors in the Etiology of Coronary Heart Disease*, Maastricht: Universitaire Pers.

5. “Socioeconomic Conditions”

This approach maintains that the political, social and economic conditions in the CEE created a climate of alienation and powerlessness that simulated, in varying degrees, the conditions of deprivation experienced by people of low socioeconomic status elsewhere in the world. This line of explanation is promising for two reasons: first, it deals directly with issues which show large differences between the CEE and OECD nations, such as material deprivation, people's perceptions of control of their lives, and their levels of social support. Second, these factors have been demonstrated in numerous studies to be important determinants of the health of both individuals and populations.

This is the only explanation that is consistent with the evidence of mortality increases after 1989. Within three years of the political and economic changes of 1989, death rates had risen dramatically in all of the former Warsaw Pact countries except Hungary and the Czech Republic, where increases were detectable but modest in size. The greatest increases were found in those aged 25 to 59. Although the trend affected both males and females, males were affected to a greater degree, and single and divorced people were affected to a much greater degree than married people.

A number of significant socioeconomic changes were observed over the same period. Real wages in every country of the former Warsaw Pact fell between 15 and 35%. This precipitous drop changed living conditions in important ways—increases in the proportion of household income spent on food in some countries in the region, especially Russia, Ukraine, Bulgaria and Romania, and declines in the average per capita consumption of meat, fish and dairy products. There was also large-scale disruption in the social environment. Marriage rates declined by 19 to 35% in the CEE countries, and there were more modest reductions in pre-primary school enrollment. Large-sample opinion surveys showed that despite the popularity of the freedoms associated with the transition to democracy, trust in the new political institutions was low, and free market reforms were widely viewed with suspicion.

Pollution of the physical environment, the weakness of the formal health care systems, and individual risk factors all contribute towards explaining lower life expectancy in the CEE countries. But the *widening* of the gap with the OECD countries

since 1989 cannot be accounted for without also considering the health effects of the widespread socioeconomic changes accompanying the economic and political restructuring in these countries. Even if rapid changes occurred in smoking and diet, for example, it is biologically impossible for these to have had such an immediate impact on mortality. And injuries and heart disease, the principal contributors to increased mortality since 1989, are not particularly sensitive to environmental pollution, or to medical care provided after the life threatening event has occurred. In fact, pollution has declined across the region since 1989, as many polluting industries have shut down.

Final Observations - Waking Up to Bacon

The zombies described above come from only a couple of cages in a much larger zoo. All are “commonsense” ideas that turn out, on examination, to be much more common than sense. But this is not news--in each case the critique we have provided above merely synthesizes widely understood counter-evidence. What makes these ideas “zombies” is their remarkable persistence in the face of apparently decisive refutation. Like the bunny in the battery commercial, they go on and on and on.

One obvious explanation is that the perpetuation of these ideas serves the economic interests of readily identifiable groups. These groups are sufficiently well placed, and well-resourced, to keep re-vitalizing their zombie ideas no matter how many times they are “killed off” and buried. “Cherchez l'interêt,” as the French might say--and did, long ago.

But this cannot be the whole explanation. Why do these ideas take hold in such a way that they survive? When their sponsors resuscitate them, why does anyone else listen? Why do some ideas resonate more broadly than others, what makes them **seem** like “common sense”?

This is not a new question. Perhaps the most elegant exploration of

“[t]he idols and false notions which are now in possession of the human understanding, and have taken deep root therein...[and] so beset men's minds that truth can hardly find entrance”

was provided by Sir Francis Bacon, nearly four hundred years ago.⁹¹ The explicit typology that he offered still serves, to a remarkable degree, to identify the reasons why certain types of erroneous ideas continue to resonate and to survive refutation. Their life expectancy seems unaffected by evidence. Bacon identified four “idols” explaining the human tendency to misinterpret evidence.

The “**Idols of the Tribe**” are those sources of misunderstanding that have their foundation in human nature itself. Bacon argued that the human understanding is “like a false mirror.” It imposes structure on the information it takes in, and this structure can, at times, serve as a source of distortion. Examples of the tendency towards systematic bias in human thinking come to us through studies of judgement conducted by Kahneman and Tversky, which show that individuals respond to complexity and uncertainty by employing simplifying strategies.⁹²

One such strategy is the tendency to generalize broadly on the basis of what is, or is perceived to be, true at the individual level. Another is to overestimate the likelihood of situations that are readily called to mind and easily understood. Causal stories that are easy to visualize, or particularly vivid, are good examples of this. The idea that fat clogs the arteries and thus causes heart disease, for instance, is relatively easy to comprehend. The role of social hierarchy in the genesis of heart disease is much harder to understand. How does one visualize the causal pathway?

The “**Idols of the Cave**” are the direct experiences of individuals. Personal experience and education “refracts and discolours the light of nature.” It causes us to overestimate the relevance and importance of instances most easily recalled, because they are most frequent, most vivid, or most familiar to us. Public ‘understandings’ of the factors influencing health care use are a good example of this.

Most of us, fortunately, do not have extensive contact with the health care system. We find it plausible, therefore, to think that high system costs must be driven by unreasonable or frivolous patient demands. But as noted above, health care services are used most intensively and extensively, at any given time, by a very small proportion of

⁹¹ Sir Francis Bacon, *Novum Organon* Book I: XXXVIII-XLIV, as found in *Sir Francis Bacon, Selected Writings*, Hugh G. Dick (Ed.), New York: Random House (1955); pp. 469-471.

⁹² D. Kahneman and A. Tversky (1982), *Judgement under Uncertainty: Heuristics and Biases*, New York: Cambridge.

the population who are very sick. The majority of the population, making occasional, and occasionally discretionary, visits to physicians, generalize by analogy to imagine that such behaviour makes up a significant part of health care use and costs. It does not.

At the same time, all of us know of family or friends who really are sick, and who are using a great deal of care, even if we do not always realize how much. But these account for a small minority of our personal observations, and so represent a relatively unimportant part of our broader understanding. Of course this understanding is fed and 'fortified' by interested parties to support, e.g. user fees.

The “**Idols of the Marketplace**” are those barriers to understanding created through the process of communication. This requires the translation of ideas into words, and the re-translation of words back into ideas. When this goes on among groups in the wider marketplace, “words are imposed according to the apprehension of the vulgar,” and “the ill and unfit choice of words wonderfully obstructs the understanding.” Just as the human mind displays inherent biases in the processing of information, so the public discourse of the "marketplace" simplifies and distorts the information that passes through it.

Bacon might well have had in mind the contemporary distortions of complex policy ideas filtered through the lenses of modern communications media. Policy discussions tend to be compressed and over-simplified, to meet the needs, not of common understanding, but of the media themselves. Sound bite explanations sell. The zombies above are all simple—and they pass easily through these filters. The refutations are more complex, detailed, and nuanced. They take longer for the media to understand, and are inherently more difficult—sometimes impossible—to translate into sound bites. They are at a permanent disadvantage—and so the messages rarely get through and do not stick. The zombies always come back.

Finally, the “**Idols of the Theatre**” “have immigrated into men’s minds from the various dogmas of philosophies, and also from wrong laws of demonstration.” These “received systems” are so named because, like theatrical performances, they represent worlds created in the human imagination. Religious systems of belief are obvious examples - to those who do not share them.

Perhaps the most obvious and influential such system in the world of public policy is neoclassical economics, which often functions more like a belief system than a body of scientific theory. It offers a consistent and self-contained explanation, not only of how human beings interact with each other, but of how they **should** interact, and how their institutions should be designed and resources allocated to yield “optimal” results.

Although these accounts appear to be empirically based, and are heavily promoted as such by their architects, they are in fact deduced from a series of *a priori* propositions. So long as one accepts the underlying assumptions, the belief system cannot be refuted by any form of evidence. The critical step, as with any belief system, is the “leap of faith” that must be made in accepting that its fundamental tenets represent “truth” –that its assumptions actually correspond to the world of experience. In the case of health care, as demonstrated by the evidence assembled above (and much more besides) they do not. The shoe does not fit.

Why then is this particular theatre so influential? The simple accounts of human behaviour that it offers, resonate with “common sense” understandings based on analogies with other walks of life, and particularly from “the marketplace.”⁹³ Yet when it comes to health care, most people do not make the full leap of faith. The proposition that people should receive only that health care that they are willing and able to pay for, is overwhelmingly rejected even in the United States. And most of us, economists included, are well aware of the powerful role played by providers in determining the level and pattern of care use. Health care is not bought off the supermarket shelf by “sovereign consumers” –and everyone knows it.

The neo-classical story, however, generates policy recommendations that are consistent with the agendas of powerful interest groups. By labeling those needing care as “consumers” rather than “patients,” it side-steps embarrassing questions about the effectiveness of care and the powerful role of providers. After all, the consumer knows best. By treating health care as just another commodity, it side-steps embarrassing questions about the distribution of access and cost. After all, we do not raise these questions about Range Rovers. By focusing on the behaviour of the individual, it side-

⁹³ Mark Schlesinger, personal communication, Summer 1997.

steps embarrassing questions about social context, and the effects of collective ‘purchase’ decisions. After all, ‘society’ is just a label for a collection of individuals.

Faced with this collection of idols, “now in possession of the human understanding,” it is not surprising that we find ourselves surrounded by such a zombie zoo. Yet Bacon’s purpose was not merely to describe, but to overthrow these idols—and to a remarkable extent he succeeded. The history of modern science bears witness. Faulty ideas, however powerful, **can** be killed off and buried—permanently. What then are the prospects for burying the current collection of zombies?

Quite frankly, we do not know. What we do know is that to date, these zombies have developed a remarkable degree of resistance to countervailing evidence. The wrong pesticide is being applied—continuing to accumulate evidence is unlikely to be effective. Our hope is that by analyzing what they really are, the purposes they serve, and the harm they bring, we may expose the zombies themselves to the light of day and so finally lay them to rest.