Achieving Health Care Reform in the United States: Toward a Whole-System Understanding

Gary Hirsch
Consultant, Wayland, MA, USA
GBHirsch@comcast.net

Jack Homer
Consultant, Voorhees, NJ, USA
jhomer@comcast.net

Geoff McDonnell
University of New South Wales, Coogee, Australia
gmcdonne@bigpond.net.au

Bobby Milstein
Centers for Disease Control and Prevention, Atlanta, GA, USA
bmilstein@cdc.gov

Abstract
The U.S. health care system faces serious problems of high cost, limited access, unequal treatment, and inadequate health protection. Though these problems have persisted for decades and various reforms have been attempted, the overall impact of reform efforts has been only modest. This paper examines potential types of reform and the history of reform efforts. Causal-loop diagrams are presented which together comprise a theory to explain what created the set of problems that exist and why efforts at reform have largely failed. Different philosophical bases for reform and the need for an eclectic approach are discussed, and a sequential “bootstrapping” approach to comprehensive reform is outlined. The diagrams and discussion of this paper are intended as a starting point for further collaborative work on health care reform among system dynamics practitioners and health policy experts, leading to simulation modeling and further insights.

Key words: health care, health reform

Disclaimer
The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention, the Department of Health and Human Services, or the U.S. Government.
Introduction

U.S. health care is in bad shape. When compared with health care in other developed nations, the American system is notable for its high cost, limited access, unequal treatment, and, by some measures, poorer health outcomes (OECD 2002; Scott 2001; LeBow 2003; Institute of Medicine 2003a). New proposals for reform appear regularly from all ends of the political spectrum, generating heated debate and reaction. Various reforms have in fact been attempted since the 1940s, but with only modest overall success. Heirich (1999) argues that most such efforts have fallen short of their goals because they are either made in piecemeal fashion and do not address the full scope of the problem, or are crafted in such a way that they fail to satisfy the interests of powerful stakeholders. He contends that successful reform policies must come from seeing the entire health care system as a large dynamic enterprise, complex unto itself and linked to changes occurring in the political economy of communities, the nation, and the world. In line with this dynamic perspective, Heirich also stresses that reformers must first understand the forces that have brought the health care system to its current state.

Our purpose here is to follow Heirich’s suggestion to try to understand the health care system as a whole. Although this is a daunting challenge, it is just the sort of “big problem” that system dynamics should be useful for addressing, and plays well to the ability of system dynamics to integrate many perspectives on an issue. Toward that end, we have developed an initial set of causal-loop diagrams to (1) describe the development of U.S. health care over the past several decades, (2) explain the fate of past reform initiatives, and (3) consider future reform possibilities. The intent of this paper is to present the diagrams as a starting point for further discussion.

Types of Reform Initiatives

Before presenting our diagrams, it is important to set the stage by describing and giving some historical context for the types of health system reform that have been attempted or proposed in the U.S. Elements of reform initiatives may be classified in four major categories:

1. Access: Expanding and creating more equitable access by providing (a) broader and/or deeper health insurance coverage, or (b) incentives and allocating resources to overcome geographic and other types of maldistribution of care.

2. Cost: Containing cost by controlling (a) the volume and nature of services provided, (b) prices of services, or (c) capacity to provide services.

3. Quality: Improving the quality of care by (a) imposing standards for care and the training of personnel who provide it, (b) implementing review processes that assure standards are being met, (c) providing better information for managing patient care and supporting the practice of evidence-based medicine, or (d) providing better information that let payers and consumers make better informed judgements about provider quality and the value of care.
4. Health protection: Preventing disease by (a) promoting healthier lifestyles and better services for managing behavioral and biological risk factors, or (b) improving living conditions that affect health.

The U.S. has made some progress in each of these areas, but is still behind other industrial nations in measures such as life expectancy and infant mortality and has costs that are considerably above those of other countries (OECD 2002, Scott 2001). As a general comment on past reform efforts, and as a preview to our causal diagrams, one may say that health care reform has been less than successful because the various objectives of reform are in conflict with each other. Rather than looking at an overall set of objectives and agreeing on some deliberate tradeoffs, reform has been done in a piecemeal manner with the implicit tradeoffs seen as unacceptable drawbacks rather than as constructive compromises.

Consequently, reforms are resisted. Attempts to expand insurance coverage are resisted because of the potential cost increases they represent. Cost-containment efforts are resisted because they threaten the autonomy of patients and providers to receive the most advanced possible care. Quality improvement efforts are resisted because they add to health care costs and administrative burden and require investments in information systems and training. And health protection efforts are resisted because they also often represent costs to employers and individual taxpayers, and are not seen as paying off as quickly or as being as urgent as treating illness and providing investment returns for the existing health care infrastructure.

Expanding Access

Commercial health insurance was introduced in the 1930s to help people pay for care when they needed it. The greatest progress in expanding coverage was in 1965 when the federal Medicare and Medicaid programs were passed to provide insurance coverage for elderly and lower income people, respectively. Prior to those programs, poorer people only received care as charity or in tax supported institutions.

Expansions in coverage since then have been limited, incremental, and uneven, varying greatly from one state to another. A few states including Hawaii and Massachusetts passed “employer mandates” that required the provision of health insurance or payment into a fund for this purpose, but these were typically not implemented or enforced. A Health Insurance Portability and Accountability Act (HIPAA) passed in 1996 provided uniform standards that improved coverage under employer-funded plans.

A provision for Medicaid “waivers” implemented during the 1990s, and adopted by 19 states to date, allowed states more flexibility to offer coverage to people below a certain income level. One limitation of the waiver programs is that they are required to be “revenue neutral” over a five-year period and not cost the Federal government any more than would have been spent without the waiver.

A number of states including Minnesota and Washington implemented programs independent of Medicaid that expanded coverage for families and children. After 20 states had passed such
programs, the Federal government passed a State Children’s Health Insurance Program (SCHIP) that assists states in offering coverage for uninsured kids.

These initiatives have been important, but approximately 15% of the population is still without public or private coverage, the highest of any industrial country. Hospital services received by the uninsured are covered by a patchwork arrangement of mechanisms, such as uncompensated care pools, which are inadequate and cumbersome and effectively create a disincentive to care for the uninsured (Institute of Medicine 2003b).

Another factor affecting access is the geographical distribution of health care resources. The U.S. has made modest progress over the last 40 years in this area. The National Health Service Corps initiated in 1972 was one attempt to bring health professionals to underserved areas by forgiving loans to recent graduates. Community health centers that began in the 1960s brought resources to inner city and rural areas. The large number of health professionals that graduated during those years led to some natural movement toward underserved areas. Yet, as work by Wennberg and others has pointed out, distribution of health resources in the U.S. has improved somewhat but still remains very uneven (Rosenthal et al. 2003).

**Containing Cost**

Efforts to contain costs over the past 40 years have dealt piecemeal with every possible determinant of cost. The need for cost-containment began to be apparent in the 1970s, particularly as the impact of advanced medical technology grew and Medicare and Medicaid-related costs started to soar. Early efforts to control costs included mechanisms such as Health Systems Agencies and “Certificates of Need” that attempted to control the expansion of health care capacity and adoption of new technology. Price controls implemented as part of a broader anti-inflation program in the Nixon Administration slowed cost increases for a while. Rate-setting commissions in some states including Massachusetts, New York, and Maryland also attempted to control health care prices. Benefit package design constrained health care spending as well, although the benefits ironically excluded many preventive services and would not pay for care until people got sick. Utilization review and pre-admission screening sought to limit the quantity of services, but also triggered the first backlash against cost-control mechanisms that were seen to interfere with patient care.

The 1980s saw the implementation of a Prospective Payment System under Medicare that used Diagnosis Related Groups (DRGs) to set fixed amounts for each admission rather than simply reimbursing providers for their costs as had been done until then. Providers found ways of “gaming” the system, but Medicare was able to use the DRG payments as a way of ratcheting down payments for patients with government coverage.

The reduction in government payments caused providers to shift costs to patients with private coverage. Employers, in reaction, began to rely more on health maintenance organizations (HMOs) and other forms of managed care to control costs. Growth of managed care accelerated in the 1990s and was able to temporarily slow increases in health care costs through capitation and other mechanisms designed to make providers more cost-conscious.
**Improving Quality of Care**

Initiatives to improve the quality of care have tried to limit avoidable errors, reduce the overuse of unnecessary services and the underuse of cost-effective services, and control risky variations in the way health care is delivered. As with other attempts at reform, this has not been accomplished by a single, coherent program, but by a fragmented constellation of mechanisms carried out by state, Federal, and private organizations that has evolved over time.

States have departments of public health that regulate facilities and boards of professional licensure that monitor health professionals (with a great deal of variability from state to state). States also have insurance commissions that monitor the insurance aspects of health care and do some quality control through their Medicaid programs.

At the Federal level, the Centers for Medicare and Medicaid Services enforce quality standards on services delivered under those programs. Non-governmental agencies such as the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) set standards and do reviews. The National Committee on Quality Assurance develops data on health plans and compares and evaluates HMOs and other managed care organizations. Much of the regulation of health care quality such as that done by the JCAHO and state licensure boards deals more with the inputs to care (e.g., training, credentials, having the right equipment) than with the outcomes that arise from it. The lack of coordination among these mechanisms also makes them less effective in assuring high quality.

**Protecting Health**

Dealing with the causes of illness, rather than just trying to treat symptoms and syndromes more effectively and economically, could be the most fundamental kind of health reform and one that has the greatest impact on health care outcome measures and costs. Better risk management can help to reduce illness incidence and severity at the individual level. In addition, there are a whole host of changes that citizens can make through their work in other sectors of society to enact health-enhancing policies, promote healthful lifestyles, and reduce our collective vulnerability to illnesses of many kinds. These changes range from improved education to safer workplaces, better housing, and reduced urban sprawl.

As with other types of health reform, there have been major achievements and significant setbacks in health protection over the past decades. For example, reductions in smoking and changes in diet to control cholesterol have contributed to declining heart disease mortality; a greater emphasis on family planning and parenting have led to healthier mothers and babies; motor vehicles were made safer, as were many workplaces; and greater awareness of the dangers of obesity has resulted in action by some school boards to take soft drink machines out of schools. At the same time, we have seen the infrastructure for health protection erode in other areas, leading to the rise of drug-resistant microbes, violence, environmental decay, and a host of modern threats. Adding to these challenges is the fact that convincing evaluation studies often take a long time to reveal the benefits of health protection efforts. Also, the elements of successful health protection are often context-specific and thus do not always transfer well from one region or locality to another.
A Causal Framework

In this section we present a series of six causal-loop diagrams (Figures 2-7) intended to capture the essential feedback relationships that (1) have driven the U.S. health care system into its current unsatisfactory state, (2) have resisted past attempts at reform, and (3) may offer insights into more effective routes to reform. Before delving into those detailed diagrams, we present in Figure 1 an overview that identifies the main factors included in the detailed diagrams and shows how those factors are associated with various stakeholders in the health system. At the center of the diagram are health care providers who interact with patients, employers, insurers/payers, drug/device makers, and regulators and monitors of health care practice. The providers care for acute medical conditions as well as more routine matters. Routine care includes monitoring and management of patients who have clinical risk factors for disease, such as obesity and high blood pressure, as well as the diagnosis and management of existing disease conditions. The intensity and efficacy of care affect the health status of patients, and also drive health care costs. Individuals are also affected by their physical and social surroundings, which, if adverse, can lead to a greater fraction at risk for disease. The demands on the health care system and its costs could be much reduced if people were more protected from health risks through the improvement of adverse living conditions.

Population stocks and flows

Our causal loops are built around the backbone of a highly aggregated stock-flow structure, seen in Figure 2, dividing the entire population into three categories: people with disease, people at elevated risk for disease, and “safer, healthier” people not at elevated risk for disease. (This structure is similar to those presented in Homer and Hirsch 2005 and Milstein and Homer 2003.) Some subset of people with disease are acute cases who may benefit from immediate medical attention. Some fraction per year of these acute cases, in turn, die from their disease. (The diagram does not indicate deaths for reasons other than disease. Neither does it indicate births, nor migration flows. Although one would want to include these other flows in a simulation model that accounts fully for population changes, they are not part of the endogenous feedback structure we are considering and so, like all other exogenous factors, have been left out of our diagrams.)

Death from disease creates a balancing loop (B1) that acts to limit the number of people with disease and their demands on health care resources. To the extent that acute patients receive effective care, this death loop is weakened, thereby prolonging life but also increasing health care demands. Sometimes, as also indicated in the diagram, acute care may actually cure a disease and not simply prolong it, but these curable cases are in the minority—particularly in the realm of chronic disease, which accounts for an increasing majority of disease burden in all developed countries.

Figure 2 also indicates the impacts of risk and disease management on population flows. Effective risk management (primary prevention) reduces the flow of people developing disease, while effective disease management (secondary prevention) reduces the acute fraction of disease. These are both good things, of course, and while risk and disease management do have costs of
their own, the reduction of acute care demands may in many cases make these preventive care practices net cost savers (see, for example, Homer, Hirsch et al. 2004).

Moving further “upstream” in the population flows, Figure 2 indicates that the amelioration of adverse living conditions can protect people from eventual disease by reducing the flow of people becoming at risk and increasing the flow of people becoming safer and healthier (Milstein and Homer 2003). Like all other interventions in the health system, efforts to improve living conditions may have a cost. But, like risk and disease management, efforts to improve living conditions are likely in many cases to be ultimate cost savers when taking into account the downstream acute care costs they can prevent.

The growth of high-tech medicine

Figure 3 shows how the growth of advanced high-technology medicine affects and is affected by the growing prevalence of disease and risk. The prevalence of acute cases (toward the right side of the diagram) provides the incentive for development of new technology for acute care, its purchase by physicians and hospitals, and for the training of specialists in the use of the new technology. This building of capability leads to more use of advanced technology for acute cases, which, in turn, prolongs the life of acutely sick patients (loop R1), or, less commonly, cures their disease (loop B2). The prolongation of life causes the prevalence of disease and disability to grow further than it would do otherwise, and thereby provides even more demand and incentive to develop high-tech capacity for lifesaving and amelioration. (The further growth of prevalence that occurs when an outflow of patients is reduced has been called “backing up”; see Jones, Homer et al. 2005.)

Innovation and capacity building have also occurred in the area of non-acute patient management, as with drugs to lower cholesterol and blood pressure, and as with the many devices in common use today for routine testing and monitoring of at-risk and chronically diseased patients. The prevalence of risk and disease (toward the left side of the diagram) provides the incentive for the development of such capability and its use in patients. Such use tends to make patient management more efficacious (although we will see that quality problems can undermine such efficacy) and thereby helps to prevent the onset of new disease as well as the progression to acute status of existing disease. The result is that the periods of risk and disease are prolonged, and the prevalence of these conditions grows further than it would do otherwise (loops R2 and R3). This growth in prevalence (another instance of backing up) provides even more demand and incentive to develop high-tech capacity for patient management.

Figure 3 also indicates several loops involving the growth of an industry and infrastructure around high-tech medicine, and the impacts of that industry on health care. As it grows, the industry exerts increasing influence on providers, insurers, and regulators. The influence on providers and insurers leads to increasing investment in high-tech equipment and training, and increasing utilization of high-tech products (loop R4). The industry and the providers who use its products also press for minimal regulation and guidelines, so as not to constrain or limit use of the new technology. Weak regulation and monitoring allow for inappropriate overuse of the technology, often associated with inadequate training and experience, and may result in harm to
patients as well as unnecessary cost (Fisher and Welch 1999). Such overuse allows the high-tech industry to grow even further (loop R5).

Another side effect of inadequate regulation is defensive medicine (also seen in loop R5), which refers to the ordering of additional tests and visits as a precaution against medical malpractice lawsuits. Malpractice may be claimed when the patient believes that a bad outcome has occurred as a result of care that does not meet the prevailing standard. Advanced technology has raised prevailing standards and expectations to the point that there are many more opportunities to fall short than there used to be, and thus, more opportunity for claims of malpractice. The combination of rising standards and weak regulation means that bad outcomes lead to more use of high technology, not less.

Figure 3 indicates an important aspect of high-tech medicine that may help to limit its proliferation. To the extent that advanced technology is able to improve risk and disease management, that will reduce the number of acute cases, and thereby reduce the demand for additional high-tech developments for acute care. Reduced demand may then translate into slower growth of the high-tech industry (loop B3).

*The fragmentation of care*

Although regulatory and oversight bodies can act to limit the inappropriate use of medical technology in a narrow clinical sense, there are other negative consequences of advanced technology over which they have no control. As shown in Figure 4, these consequences include the fragmentation of medical care into numerous specialties and subspecialties, and the natural tendency of specialists to congregate in more populous and attractive areas. The proliferation of specialists has been shown to lead to numerous quality problems in patient management, often because of poor information transmission among providers and logistical difficulties for patients (Institute of Medicine 2001). These quality problems undermine the potential benefits of more advanced tools for risk and disease management (loop B4).

The geographical concentration of specialists has tended to leave some poorer inner cities and remote rural areas short of physician capacity (Wennberg et al 2002). The fraction of patients under proper management for their health risks and chronic diseases in these underserved areas is less than it is in fully served areas, contributing to the problem of inequitable health care in America (Institute of Medicine 2003). Because the underserved often do not receive the preventive care they should, they suffer more disease and more fatal complications, and progress to the acute stage of disease faster than the well-served portion of the population who remain backed up in early stages of the disease (loop B5). Some of these underserved acute cases will receive only basic treatment at local hospitals, and in that way will at least not exacerbate the trend toward geographic concentration (loop B6). But other underserved may travel or be transferred to larger and better endowed hospitals, where advanced technology may save their lives, thereby seeming to justify the continued growth of high-tech medicine. Unfortunately, it is precisely this high-tech trend, leading to the greater geographical concentration of physicians, that helps to create the problem of underserved patients in the first place (loop R6).
Rising costs and attempts to contain them

Figure 5 traces the ways in which advanced technology contributes to higher costs of acute care and patient management, and also indicates how insurers and employers may attempt to contain those costs. The purpose of most new technology in medicine is to provide more effective care, which usually ends up costing more rather than saving money. Thus, an increase in the utilization of high-tech care tends to increase costs. However, acute care and patient management are not equal in this respect. Acute care typically costs much more, on a per-capita basis, than do risk and disease management. It follows that technology-driven increases in patient management costs may be justifiable on both health and cost grounds if they can reduce the need for expensive acute care.

Other costs related to high-tech medicine, although not explicit in Figure 5, should also be noted. First, we previously mentioned in connection with Figure 3 the fact that an influential high-tech industry has the indirect effect (due to increased standards of care combined with weak regulations) of increasing malpractice claims and the practice of defensive medicine, both of which increase costs. Second, we previously mentioned in connection with Figure 4 how high-tech medicine tends to fragment care among multiple specialists, which increases the possibility that tests and other procedures will be repeated by one provider that had already been performed by another provider, thereby increasing the costs of patient management.

Figure 5 indicates how insurers and employers may attempt to contain rising health care costs. The two basic mechanisms we have pictured, and combined for the sake of simplicity, are reductions in insurance reimbursement rates to providers, and reductions in the availability or breadth of insurance coverage to consumers. Reduction in the availability or breadth of insurance coverage has in recent years become a cost-containment tactic of employers relative to their employees. The fraction of the population covered by employment-based insurance (through an employer or union) declined each year from 2000 to 2003, starting at 64.1% and ending 60.4% (U.S. Census Bureau 2005). Reduction in the breadth of coverage—consisting of actual exclusions or low payment percentages for specific conditions and treatments—may also be considered a tactic of government to the extent that Medicare and Medicaid have forced many individuals into managed care plans with fewer options than traditional fee-for-service plans.

Figure 5 shows two possible consequences of reduced reimbursements or reduced insurance coverage: less high-tech acute care (loop B7) or less high-tech patient management (loop B8). Let us consider the relative strength of these two loops, for each of the two mechanisms of containment. First, in regard to reimbursement, insurers have historically reimbursed acute care relatively more generously than patient management—and have tended to cut patient management reimbursements more readily than they have cut acute care reimbursements. This bias may reflect in part the fact that the benefits of many acute care interventions are easier to demonstrate than those of patient management activities and, as a result, may be in greater demand by the public. Another reason may be the fact that there is a stronger medical-industrial lobby pushing for acute care than for patient management. This bias may finally be changing, now that evidence for the benefits of patient management has become more plentiful and definitive.
Second, in regard to insurance reductions, the kind of care likely to be most affected is—as in the case of reimbursement reductions—patient management. Patients lacking any health insurance or the means to pay out of pocket are unlikely to see a provider routinely for preventive care, but can still receive emergency care from most hospitals, including many if not most types of high-tech acute care. Similarly, patients with limited health insurance or whose provider reimbursements are reduced are likely to be limited not so much with respect to acute or catastrophic care, but rather with respect to regular visits to specialists for appropriately aggressive management of risk factors and chronic illnesses.

Loop R7 indicates an unintended effect of cost containment that undermines its intended effect, and is also harmful from a health standpoint. To the extent that best-practice patient management becomes less regularly used as a result of cost containment, one will see increases in the prevalence of disease with acute complications, as well as an increase in acute care costs. This cost increase undermines the original intent of cost containment and, given the high cost of acute care, could plausibly end up increasing overall health care costs rather than decreasing them.

Another important unintended effect of cost containment measures is shown in loop R8. In the multi-payer system that characterizes the U.S., the demand for health care cost containment has translated into aggressive competition among many insurers trying to offer employers acceptable benefits for their employees at the lowest price. One aspect of this competition is the creation of a broad and ever-changing menu of plans with different exclusions and different payment percentages for different health services. Another aspect of the competition is the effort to lower reimbursements to providers to the extent possible, which often involves intense negotiations with the larger specialty practices and hospitals whose business the insurer cannot afford to lose.

Facing this cacophony of payer fee schedules and arrangements, the administrative overhead of providers has increased significantly since the 1980s. The growth of overhead costs has, in turn, reduced the productivity and income of many providers. These losses have undoubtedly caused some providers to find creative ways to increase their billings, and have also caused providers to consolidate to form larger entities so that they may have greater negotiating clout with the insurers. The net effect of both responses is to undermine the ability of insurers to reduce their costs and to offer employers a better price for health care.

*Living conditions and citizen involvement*

In Figure 6 we move further upstream in the health stock-flow system to consider another detrimental effect of health care costs: a reduction in the funds available for ameliorating health-threatening living conditions. As public and private funds have been increasingly absorbed for supporting and improving the care of existing disease, funding agencies have tended to reduce their support for creating safer and healthier communities and environments. As a consequence, more individuals have become at risk for disease and have developed disease—see, for example, the case of rapidly increasing obesity and diabetes since the 1980s. The response has been to put an even greater investment of health-related funds in patient management and care at the expense of living conditions (loop R9).
In Figure 6 we also see that the protection of health is determined not only by funding, and the management of disease is made possible not only by clinical effort, but that the involvement of ordinary citizens may provide important support and impetus to both activities. Responding to an increasing prevalence of health risks and disease in their communities, some citizens may choose to lend a hand to improve living conditions, thereby helping to prevent further increases in risk prevalence (loop B9). Citizens may also contribute by assisting needy or socially isolated patients in the daily self-management of their risk and disease conditions (loop R10; a reinforcing loop, because it helps patients remain backed up rather than progressing to acute disease.)

**Why Hasn’t Reform Had More Impact in the U.S.**?

The historical discussion of health reform early in the paper indicated some early victories in the 1960s in terms of more people receiving insurance coverage through Medicare and Medicaid, but little since then other than small advances resulting from specific programs. It is useful to return to the question raised earlier in the paper about why reform has not had more impact in the U.S. and use the causal loops that have just been presented to help explain the lack of impact. Figure 7 summarizes these loops and shows how different types of reforms can be expected to impact variables in the loops.

One overarching reason that reform efforts have not succeeded, as Heirich (1999) suggests, is that they have been attempted piecemeal rather than in a coordinated fashion. As we examine each type of reform, it is especially helpful to examine how this sort of reform, by itself, has generated “pushback” from the system and undermined its potential impact. Another possible reason that reforms have not succeeded is that they have been attempted without resolving some rather basic questions, presented below, about each type of reform itself.

**Reforms to improve the distribution of care**

Reforms to improve access via better distribution of care raise some fundamental questions. For example, is it possible to achieve more equitable distribution in the U.S., geographically and by income level, or is some maldistribution a natural consequence of the country’s sprawling geography and emphasis on provider autonomy? What mechanisms work best for achieving some modicum of redistribution? Is equitable distribution even a desirable goal as long as everyone has access to an acceptable level of care? How can one ethically decide what is an acceptable level of care, and who is in a position to make that determination?

Other questions concern how improving distribution relates to other goals of health reform. Ideally, as indicated in Figure 7, improved distribution of care would occur in a way that expands the management of risks and diseases. Expanded risk and disease management may ultimately reduce costs if additional cases of illness and complications can be prevented. However, this effect cannot be achieved selectively. Providers attracted to remote areas will also provide more acute care including elective care that may be constrained when supply is low, but will expand when there are more providers available. Expanded acute care will generate costs immediately while improved risk and disease management services may take longer to have their impact. Availability of expanded acute care services in previously underserved areas and populations
may even unleash a backlog of demand that constrains providers’ ability to deliver risk and disease management services, and thereby further dilutes their impact. Improved distribution may mean that people who previously lacked access to care now have the same ability to get overly expensive acute care as the rest of the population already does. There is also the danger that attempts to improve distribution, if they involve shifting resources among areas, may end up reducing access for some people.

The result of these efforts is higher overall cost since improved distribution is rarely achieved by redistributing resources, but rather by adding resources to underserved areas. Any redistribution typically occurs very slowly, if at all, because of how location decisions are made by providers and how large fixed investments are tied up in existing facilities. The potential for higher overall cost acts as a disincentive to improving the distribution of care.

Reforms to expand insurance coverage

Improved access through expanded insurance coverage poses questions similar to those of improved distribution. For example, does everyone need access to the same set of services or benefit package, or is some variation acceptable as long as everyone has access to care that meets certain standards? What level of services represents a point up to which additional care improves health and beyond which services add only to cost and not to value? Should providers be obligated to serve everyone or can they select the patients they serve?

Figure 7 indicates that the expansion of insurance coverage would impact health care in essentially the same ways as improved distribution of care would. Increased availability and breadth of insurance coverage can expand both the managed percentage of risks and diseases and the percentage of acute patients receiving high-tech care. A higher percentage of risks and diseases being effectively managed will have a desirable effect by reducing the rate at which people develop diseases and the rate of acute episodes among those with existing disease. The ultimate effect is improved health and lower cost, but these effects take a while to develop. Meanwhile, expanded patient management and acute care services made possible by broader and deeper insurance coverage will generate immediate costs. These short-term costs may create a disincentive among employers to adopting expanded coverage, especially when health care costs are already considered to be quite high. Employers may respond to mandatory coverage requirements by selecting plans for their employees with higher co-pays and deductibles that effectively make some forms of care unavailable to lower-paid employees.

Attempts to improve coverage are somewhat like bailing out a leaky boat. As long as there is some ultimate upper limit on how much employers and taxpayers are willing to spend on health care or some maximum rate at which they will allow spending to grow, giving some people better access to coverage will lead to a reduction in coverage for others, with little net improvement overall. Reform that produces expanded coverage or better distribution of care requires some offsetting reduction in cost if it is to be acceptable.
Reforms to contain costs

The regulation of costs and service volumes is the subject of fierce debate, with suggested solutions ranging from a fully-regulated single-payer system to one that depends entirely on the free market (see below). Central questions in this debate include: How much cost is too much, and are we getting fair value for our money? What are fair returns to physicians, drug companies, and other providers? Is it necessary to put the brakes on high-tech acute care, for example, by establishing more stringent criteria for Medicare and Medicaid reimbursement of specific technologies?

Figure 7 indicates where attempts at cost containment come up against other goals for health reform. Efforts to contain costs by limiting the volumes, prices, and capacity to provide care have typically had limited, temporary effects in slowing the rate of increase in health care costs. Reductions in acute care may eliminate some unnecessary procedures that result in both cost savings and protection of patients from potential harm. Unfortunately, cost-containment measures also tend to reinforce the bias toward acute care services since these are typically regarded as more urgent and less discretionary. Risk and disease management can suffer or be deferred as a result. Cost-containment efforts may reduce risk and disease management directly by not paying for certain services or they may simply keep providers very busy with heavier patient loads and excess paperwork that get in the way of better patient management. Although cost containment that reduces risk and disease management will result in higher cost in the long run, this is typically beyond the decision making time horizon used by managers, politicians, and government officials. Cost containment can also lead to potentially detrimental effects on quality if staffing levels are cut below what is really needed. Cost-consciousness may also prevent investments from being made in such things as better information systems that could improve quality and ultimately reduce cost.

Reforms to improve quality

The most basic question about quality is how to measure it and what constitutes better quality rather than simply a more elaborate product. What measures are most meaningful at the individual and population levels? The next most basic question is about what mechanisms help to improve quality. Does the current system of malpractice law that gets so much attention in the public debate actually improve quality or potentially reduce it while adding to costs through the practice of defensive medicine? What combination of autonomy and regulation will assure that providers give the best evidence-based care while having the flexibility to adapt care to the needs of individual patients? Who should regulate care? What mix of regulation by providers themselves, government, insurers, and other agencies will yield the best result?

Figure 7 indicates where quality improvement mechanisms have their effects and how they affect other health reform goals. One dimension of improved quality would be maintaining better information for patient management, allowing for reduced onset and progression of disease. Another dimension would be to improve the quality and efficacy of acute care, allowing more patients to recover from acute episodes. However, quality improvement presents the same sort of short-versus-long-term cost dilemma as improving access to care does. Monitoring to assure provision of necessary but currently underused services will increase the volume of services.
These might pay off in the long run if they allow for better risk and disease management, but will cost more initially. As indicated in Figure 7, regulations and quality assurance procedures themselves add to cost and require up-front investments in information systems and training. Consumers, employers, and government must be willing to pay more in the short term, but have been reluctant to so far. Better quality assurance can ultimately save money if it reduces procedures done in the context of defensive medicine by establishing more objective standards for care rather than having providers feel that they need to do everything to avoid being sued. However, it will take time to change patterns of behavior to take advantage of these more objective standards.

Reforms to protect health

Improving the adverse living conditions that contribute to illness could ultimately be the most effective sort of reform in terms of improving health outcomes and reducing costs. However, as Figure 7 indicates, most funds that might be devoted to further improvement in this area are already spoken for by the costs of treating illness and paying for the existing health care infrastructure. Savings that might be produced by healthier living conditions will take a while to develop while the costs of these measures are immediate. The needed investment is likely to be resisted by the employers and government officials who would be asked to fund these measures. The measures may also be opposed by some taxpayers and consumers, especially those who do not live in the communities most affected by adverse living conditions. Another difficulty with funding health protection is that there exists no mechanism for capturing health cost savings and re-investing them in healthy lifestyles and environment. Without some explicit mechanism for reinvestment, cost savings from initial modest investments in health protection are likely to lead to lower taxes and slower growth in health insurance premiums, leaving insufficient funds for appropriate expansion of the investment. Because of the difficulties of obtaining consistent government funding for such long-term investments, it may fall to citizens and community groups and faith-based organizations to help achieve some of the required changes themselves without much government support.

Can Unifocal Health Reform Succeed?

Insights gleaned from the causal diagrams suggest that an effort concentrated in any one of these areas alone will create enough “pushback” to jeopardize its chances of succeeding. To illustrate further, we present a thought experiment in which one might try to reduce costs simply by improving quality alone.

In the discussion of cost-drivers above, we touched upon a couple of key factors that should be amenable to quality improvement initiatives. These include inappropriate use of technology in acute care, and the poor state of information sharing among providers which leads to errors and duplication in patient management. It is also possible that quality initiatives, such as those that support evidence-based medicine, could raise the fraction of patients whose health risks and diseases are under proper management. One might reasonably ask whether a package of initiatives that effectively corrects these quality problems could also stem the rise in health care costs.
The answer to this question hinges on the fact that acute care accounts for the majority of health care costs. Even if one assumes that the quality improvement measures could be implemented at reasonable cost, health care costs still will not be reined in unless the growth of high-tech acute care itself is controlled. Once a patient develops disease with acute complications, the modern medical imperative is to save and prolong that person’s life using the most advanced tools available. Assuming that this imperative is unaffected by the quality initiatives, the only way to stem the growth of acute care is to reduce the growth in acute disease itself. Thus, the more precise question to ask is whether quality initiatives could improve the effectiveness of patient management sufficiently to reduce the growth of acute disease.

We suspect that for a few basic reasons the answer to that question is, unfortunately, no. First, recall the basic dynamic by which the successes of acute care continue to prolong the lives of the ill, helping to build their numbers and justify further expansions in acute care. Second, quality initiatives by themselves would not address the problem of geographic concentration leading to lack of health care access for underserved populations. Third, quality initiatives by themselves would not address the lack of access to patient management for the portion of the population who are uninsured or whose insurance does not adequately cover preventive services. Fourth, the quality initiatives would do nothing about improving living conditions and protecting people from health risks, but would only attempt to limit the damage consequent to those risks. If the growth of risk is not addressed, improvements in quality might end up looking like an effort to keep the shore clean and dry during a rising tide.

Is One Philosophical Approach to Reform the Best?

The health reform debate in the U.S. revolves around the philosophical approach that should guide more detailed aspects of health reform. Heirich (1999) identifies three such approaches that dominate the debate: single payer, market competition, and managed competition. He characterizes the three as follows:

- **Single Payer** attempts to assure access to care for everyone, drawing funding from taxes or other broad streams rather than being employment based. Benefit package design, global budgeting, and annual budgets for hospitals would help to constrain the utilization of costly technology while the single-payer approach also allows for tighter control of reimbursement rates.

- **Market Competition** does not address the issue of coverage for all nor does it purport to share the cost burden equitably among different segments of society. Costs are expected to be contained through the inclusion of some managed care mechanisms, as well as by the fact that consumers would pay for care directly out of their own health savings accounts. Capitated payment and voluntary purchasing pools may also help to limit reimbursements.

- **Managed Competition** may include mandates requiring that region-wide providers serve all who want coverage. Funding comes from a mix of mandated payments or contributions by employers, employee co-payments, and tax revenues for certain patients who cannot pay on their own. Benefit package design and the various mechanisms of managed care would help
to control technology utilization while monopoly purchasing by alliances and capitated payment would limit reimbursements.

The unfortunate thing about the debate is that proponents of each approach seek to have the models implemented in their pure form. However, any one of the approaches implemented by itself has shortcomings that would impede its adoption and its effectiveness as a vehicle of reform.

**Evaluating the single payer approach**

In a perfect world, a single payer plan has a great deal of appeal. It is the simplest to administer and avoids the administrative cost spiral (first seen as loop R8 in Figure 5) made possible by the existence of multiple competing health plans. Single payer also provides the greatest possible availability of coverage, thereby increasing the managed fraction of risks and diseases, which, in turn, reduces the number of acute episodes. A single payer plan also confers great bargaining power on the insurer to control reimbursements, and thereby avoid the high-tech service volume spirals seen in Figure 3 (loops R4 and R5).

While the single payer approach may represent a desirable steady state, getting there is the greater challenge, given the prevailing political and economic environment in the U.S. The growth in influence of health care providers, insurers, and drug and medical device companies has created formidable vested interests that are likely to oppose any proposal built on a single payer platform. In addition, even with the greater control on volumes and reimbursement rates of a single payer system, greater availability of insurance coverage could trigger costly increases in demand for high-tech care, based on the need to satisfy latent demand that is not currently being met. Pent-up demand may be enough to overwhelm the capacity of health care providers and keep them from increasing the fraction of risks and diseases managed, losing an important benefit of broader coverage. Resources needed to meet this surge in demand could also detract from the ability to invest in improved living conditions, and thereby allow the further growth of risk and disease (see loop R9 in Figure 6).

**Evaluating the market competition approach**

Market competition seeks to maintain control on service volumes and reimbursements by making purchasers and consumers of health care more aware of their costs. This would ideally weaken the drivers of increased volume and cost depicted in our diagrams. Some people, perhaps even a majority, might reduce their health care costs as a result of the control granted them through mechanisms such as health savings accounts.

However, pure market competition would abandon a significant minority who could not afford to pay for adequate insurance coverage. As indicated in Figure 5 (loop R7), reduced availability and breadth of coverage would mean that fewer people have access to risk and disease management, and the acute care costs of the underinsured could subsequently skyrocket. Unless we are willing to abandon these people altogether, their emergency care costs will either bankrupt providers or be spread over the consumers who can still afford insurance, wiping out any savings produced by anticipated efficiencies of the free market approach. There is the additional problem with market
competition—to the extent that such an approach lacks managed care mechanisms—that most consumers do not have the knowledge needed to always make the best choices, and may end up opting for care they do not need or opting out of care that could be cost saving. Without clinical skills, consumers are also hard-pressed to assess the quality of the care they receive, which means that such quality could suffer under pure market competition.

Evaluating the managed competition approach

Managed competition attempts to strike a balance between the two extremes. This seems sensible, but we have to remember that this is the system that has been in place in the U.S. for a number of years and appears to satisfy no one. A major problem with managed competition is that it is relatively weak on cost control, and those cost controls that do exist bias care in the direction of high-tech acute care. Third-party payment insulates individual consumers from the costs of their care and removes any free market incentives to seek less expensive or more cost-effective care. Multiple insurers and managed care companies each lack the leverage to effectively resist cost increases by providers. Insurers who push back too hard in trying to control rate increases or limit access to procedures of questionable merit typically lose access to providers for their subscribers or lose subscribers who insist upon access to the latest technology. Medicare and Medicaid are large enough that they can exert some leverage. However, because there are still enough individuals not covered by such effective cost-controlling managed care plans, providers have been able to shift many of their costs to these individuals. Through such cost-shifting, providers have largely been able to sustain their ability to profit from the practice of high-tech, acute-care oriented medicine.

The forces that have driven cost escalation in the past will continue to operate under a managed competition regime. As cost escalation is concentrated on people not covered by effective managed care, there is a greater danger that employers will not be able to offer plans with proper risk and disease management, or that the number of uninsured will continue to increase.

These drawbacks suggest that an approach heavily dominated by any one of these philosophical approaches will not make much headway in achieving health reform. In addition, none of them addresses living conditions that affect health and may ultimately be the best way to control costs while producing better health status. An eclectic approach seems to be called for. The Clinton plan in the 1990s attempted something like this, but, according to Heirich, encountered resistance from insurers who felt they would be losers under the plan, as well as from certain business interests, the public, and Congress. As discussed in the next section, this sort of resistance to rapid, comprehensive change may suggest a more gradual, sequential approach as well as one that is eclectic in drawing from different segments of the ideological spectrum as well as diverse public viewpoints.

A Bootstrapping Approach to Comprehensive Reform

Aside from the question of philosophical approach, another key question in health reform is “How much change is enough?” Responses to this question lie on a continuum from doing nothing at one end, to a very comprehensive package of reform at the other, with various types of incremental reform in the middle.
A do-nothing approach may be advocated by those who feel that the current situation is not that bad or, alternatively, despair of achieving significant reform. However, a do-nothing approach is a recipe for increased cost and illness. As our causal loops indicate, uncontrolled growth in health care costs will lead to fewer people having insurance, less attention to managing risks and disease, and less money spent on the living conditions that affect health. Employers and taxpayers already spending large sums of money on health care will not be willing to spend more on risk and disease management and promotion of healthy lifestyles, even if those measures are shown to be ultimately cost saving.

Incremental change may take any of a number of forms, but as explained above, each of these partial approaches—whether the emphasis is on access, cost, quality, or health protection—has serious shortcomings. A comprehensive approach is thus ultimately what is needed, but the experience with the Clinton health plan mentioned above suggests caution. Efforts to pass and implement a comprehensive health reform plan all at once are fraught with peril because of the propensity of the system to resist major changes, especially those that have higher initial costs associated with them. Thus, although a comprehensive plan should inform and inspire a long-range vision, it might best be achieved in a sequence of steps that achieves and builds on early successes to create momentum and drive more significant change.

This kind of “bootstrapping” could, for example, start with relatively low-cost interventions, perhaps mostly in the form of informational campaigns aimed at providers, patients, and insurers, that would improve the quality of disease management for chronically sick patients through greater adherence to best-practice guidelines for chronic care. As shown in Figure 7, more effective disease management could pay off rather quickly in the form of fewer acute cases and their associated costs. These initial savings from disease management could be used to help pay for health information systems, which could, in turn, further improve the effectiveness of disease management and savings from it (Homer, Hirsch et al. 2004).

The resulting savings from disease management that accrue to employers and insurers may reduce their tendency to cut back on insurance coverage and on reimbursement rates for risk and disease management. Moreover, these further savings might next be applied as a matter of public policy to expanding access to those still underinsured and underserved, and making certain that they too receive effective disease management so as to minimize any increase in acute care costs associated with such coverage. Ultimately, this expansion of coverage should reduce taxpayer outlays required to provide emergency medical care for the underserved.

The savings generated by the previous steps of improved disease management and access could next be applied to providing more diligent risk management, again perhaps through a combination of informational campaigns and information systems investments. Improved risk management would gradually lead to more cost savings, because it prevents disease and the greater cost of managing that disease. In a final step, the further savings could be applied to improving living conditions, which ultimately may be the best way to reduce health care costs over the long term.
Under the bootstrapping approach, more and more savings would be generated by moving increasingly upstream in the health population system we have described and diagrammed. The result would be better use of resources and better health for the population. Initial successes may produce a sense that change is possible and lead to a groundswell of political support for meaningful change in the way health care is financed and delivered.

The main challenge in the bootstrapping approach will be to create effective mechanisms for reinvesting savings in health that do not allow these savings to be spent on other things or returned to consumers and employers in the form of lower taxes or higher profits or wages. The American people will have to be willing to make a commitment to an initial investment in better care for the underinsured and the underserved, as well as a commitment to long-term funding for health promotion and protection. A part of this funding should seek to leverage the energies of individuals and community organizations to create healthier communities. Improvements in health, especially those that arise from improvements in living conditions, do not have to rely on the formal health care system, but can come from citizens working together in their own communities.

**Unresolved Questions and Next Steps**

Causal diagramming has led us to some initial conclusions about what will be required for successful health care reform in the U.S. However, our analytic effort is at present still exploratory. Because the diagrammed hypotheses are not yet in the testable form of a simulation model, we can only say that our conclusions thus far are tentative. On the other hand, these conclusions are at least logically derived from an integrated framework, which represents progress over the incomplete and slanted arguments typically presented in the debate over health care reform.

Simulation modeling supported by empirical data could help to confirm or reject the conclusions presented above, and could also address some key questions our diagrams did not touch upon. These questions include, for example:

- What is the critical mass of change needed to begin achieving sustainable reform?
- What is the right sequence for “bootstrapping” change?
- Which health problems offer the best payoff in terms of making early progress?
- What mechanisms can be used to capture savings and where should they be reinvested?
- What stakeholders need to be involved at each step?
- Which changes provide a basis for agreement rather than conflict?

We hope to attract other members of the system dynamics community to this effort to investigate the dynamics of health reform, and to make it an ongoing initiative of the SD Society’s Health Policy Special Interest Group, with future extension into a more international perspective and consideration of global health problems. We also hope to attract non-SD health policy experts to join us in this exploration and look forward to collaborating with them in the writing of papers and the development of concrete policy proposals.
References


Milstein B, Homer J. The dynamics of upstream and downstream: Why is it so hard for the health system to work upstream, and what can be done about it? Atlanta, GA: Centers for Disease Control and Prevention; December 3, 2003. Powerpoint presentation.


Figure 1. Overview: Stakeholder roles and interactions
Figure 2. Population health dynamics
Figure 3. Growth of high-tech medicine
Figure 4. Fragmentation as a side effect
Figure 5. Attempts to contain costs
Figure 6. Living conditions and citizen involvement
Figure 7. Intervention impacts