S. HRG. 103-603

MEDICAL PRACTICE PATTERNS AND APPROPRIATENESS OF CARE

HEARING

BEFORE THE

COMMITTEE ON FINANCE
UNITED STATES SENATE

ONE HUNDRED THIRD CONGRESS

FIRST SESSION

OCTOBER 26, 1993

Printed for the use of the Committee on Finance

U.S. GOVERNMENT PRINTING OFFICE
78-699—CC WASHINGTON : 1994

For sale by the U.S. Government Printing Office
Superintendent of Documents, Congressional Sales Office, Washington, DC 20402
ISBN 0-16-044586-8

COMMITTEE ON FINANCE

DANIEL PATRICK MOYNIHAN, New York, *Chairman*

MAX BAUCUS, Montana	BOB PACKWOOD, Oregon
DAVID L. BOREN, Oklahoma	BOB DOLE, Kansas
BILL BRADLEY, New Jersey	WILLIAM V. ROTH, JR., Delaware
GEORGE J. MITCHELL, Maine	JOHN C. DANFORTH, Missouri
DAVID PRYOR, Arkansas	JOHN H. CHAFEE, Rhode Island
DONALD W. RIEGLE, JR., Michigan	DAVE DURENBERGER, Minnesota
JOHN D. ROCKEFELLER IV, West Virginia	CHARLES E. GRASSLEY, Iowa
TOM DASCHLE, South Dakota	ORRIN G. HATCH, Utah
JOHN B. BREAUX, Louisiana	MALCOLM WALLOP, Wyoming
KENT CONRAD, North Dakota	

LAWRENCE O'DONNELL, JR., *Staff Director*
EDMUND J. MIHALSKI, *Minority Chief of Staff*

CONTENTS

OPENING STATEMENTS

COMMITTEE PRESS RELEASE

PUBLIC WITNESSES

ALPHABETICAL LISTING AND APPENDIX MATERIAL SUBMITTED

COMMUNICATIONS

MEDICAL PRACTICE PATTERNS AND APPROPRIATENESS OF CARE

TUESDAY, OCTOBER 26, 1993

U.S. SENATE,
COMMITTEE ON FINANCE,
Washington, DC.

The hearing was convened, pursuant to notice, at 10:06 a.m., in room SD–215, Dirksen Senate Office Building, Hon. Daniel Patrick Moynihan (chairman of the committee) presiding.

Present: Senators Rockefeller, Daschle, Packwood, Chafee, Durenberger, and Grassley.

[The press release announcing the hearing follows:]

[Press Release No. H–40, October 21, 1993]

FINANCE COMMITTEE RESCHEDULES MEDICAL PRACTICE PATTERNS HEARING

WASHINGTON, DC—Senator Daniel Patrick Moynihan (D–NY), Chairman of the Senate Committee on Finance, announced today that the Committee has rescheduled a hearing regarding medical practice patterns and appropriateness of care originally scheduled for October 21.

The hearing has been rescheduled to begin at *10:00 a.m. on Tuesday, October, 26 1993,* in room SD–215 of the Dirksen Senate Office Building.

"Geographic variation in medical practice patterns appears to suggest that the number of expensive medical procedures may be inappropriate," Senator Moynihan said in announcing the hearing. "The Committee wants to examine the circumstances under which unnecessary procedures are performed, and inquire into how we can change the incentives in the system in order to reduce inappropriate care."

OPENING STATEMENT OF HON. DANIEL PATRICK MOYNIHAN, A U.S. SENATOR FROM NEW YORK, CHAIRMAN, COMMITTEE ON FINANCE

The CHAIRMAN. A very good morning to our distinguished panel and our guests. This is, I suppose, the last in a succession of hearings that we have been holding preparatory to receiving the administration's proposal on health care.

This particular meeting is a hearing on medical practice patterns and appropriateness of care, a subject about which both physicians and economists are concerned. Indeed, so is the rest of the Congress.

It happens that one of our panelists, Dr. Wennberg, who is the Director of the Center for the Evaluative Clinical Sciences at Dartmouth is featured in a long story in the New York Times this morning on this very subject. It sounds very good.

We are also going to hear from Dr. Mark Chassin. Dr. Chassin is our Commissioner of Health in New York State. And Dr. Charles Phelps, who is a professor of political science and chairman of the

Department of Community and Preventive Medicine at the School of Medicine and Dentistry at the University of Rochester. We have a very wide range of experience and competence and interests here.

Senator Packwood and I are going to have to leave in a short period because the extended unemployment insurance bill is on the floor and we are, by definition, the bill managers. So the sooner we get started the better.

Senator Packwood?

OPENING STATEMENT OF HON. BOB PACKWOOD, A U.S. SENATOR FROM OREGON

Senator PACKWOOD. I have had a chance to read the testimony of Dr. Wennberg and Dr. Chassin already. I find it most interesting. You are the first group of witnesses we have had who said there is a lot, and they do not mean because of malpractice or anything else, there is a lot that can be squeezed out legitimately. Not from people who are cheating, but from people who perhaps are practicing doctoring to a higher degree than is necessary under normal circumstances.

As you are all aware, Oregon has tried a variant of some of the things you are talking about and we tried to rank medical procedures on the basis of effectiveness, not just on the basis of cost. We have not said an 8y-year-old cannot get a hip replacement because it costs a lot. But we have attempted to rank it on the basis of effectiveness.

And many of the treatments at the bottom of the Oregon Medicaid waiver list are treatments for which there is no rational likelihood of effectiveness. So if they are not going to be effective, why pay for them?

It also turns out that the Oregon list—and this is more coincidental than intentional—the things at the top of the list are relatively inexpensive medicine, a lot of it preventive, a lot of it prenatal. There is no question in terms of bang for the buck it pays off.

But one of the reasons they are at the top of the list, the principal reason is, that the effectiveness is very high. And again, cost was not the factor. The effectiveness was high. And it also turned out to be relatively and comparatively inexpensive to some things that do not work at all and are very expensive.

I may have some questions. I may have to submit them in writing because the Chairman and I will be on the floor. But I was very intrigued with the two statements I read. Dr. Phelps, I did not see yours as of last night. But I was very intrigued with the statements I read and, frankly, quite impressed with them.

[The questions appear in the appendix.]

The CHAIRMAN. Dr. Wennberg, unbeknownst to us but for the record, on the patterns of care, the economists identify something called "Say's Law" after a French economist of the 18th Century, which simply states that supply creates demand.

I believe you find that in villages where there is a doctor who can perform tonsillectomies there will be quite a few more tonsillectomies performed than in villages where they do not have them.

Dr. WENNBERG. That is right.

The CHAIRMAN. It stands to reason, but it is something to be noted.

Senator Durenberger, a friend of all of our distinguished witnesses.

OPENING STATEMENT OF HON. DAVE DURENBERGER, A U.S. SENATOR FROM MINNESOTA

Senator DURENBERGER. Mr. Chairman, thank you. I want to just make the observation that last Tuesday's hearing and today's hearing are not in the classic mold, but they are probably greater contributors to where we go on health care reform than any others.

I was even at a lobbyist breakfast this morning and somebody raised his hand and said, "What good is it if we guarantee universal coverage for medical access and have not dealt with crime in America?" I said, you should have been at the Finance Committee hearing a week ago. The Chairman has asked us to examine the social condition of the country.

Today he has asked us to examine the issue of productivity, if I can just use a one-word description of what we are doing here. If, in fact, the goal in American health care reform is to have more and better but have it cost less, productivity is the answer.

The question is going to be how do we get to it. Do we get to it through something called managed competition or managed competition modified by some form of rules, either national rules or state-by-state rules or whatever? What is the role of budgets? In which I disagree with a couple of my colleagues' presentations, although we have discussed that before.

But it is interesting also that the Washington Post this morning, for which we make plans for what the people care about inside the beltway, has this interesting story by Michael Wisekoff, "From Eyes, to Ears, to Nose, to Back: Specialists Battle to be Part of Clinton Plan."

One of the association representatives says, "Rather than have the government set the fees, they have chosen to have providers to pull out the knife and do one another in." I could not think of a better way to describe the concerns that some people in the fee-for-service piecemeal practice of medicine have for the issue of productivity.

Mr. Chairman, you have chosen well in these three witnesses, and Dave Eddy, who also could not be with us today. I know all of us are going to learn a great deal from their testimony and their continued cooperation.

[The prepared statement of Dr. Eddy appears in the appendix.]

The CHAIRMAN. Well, your being the one who knows the most already, that is very generous of you to say, sir.

Senator Grassley?

OPENING STATEMENT OF HON. CHARLES E. GRASSLEY, A U.S. SENATOR FROM IOWA

Senator GRASSLEY. Mr. Chairman, I believe the subject we are talking about today on the appropriateness of care is a very critical issue to deal with when we talk about health care reform. That is because Americans, of course, have a right to expect that the health care they receive will be of the highest quality.

But it is also critical because in the managed competition system, the consumer has to be able to evaluate the quality of medical care provided through their health care plan.

There appears, however, to be some skepticism as to whether we actually have such measures available. We are going to hear today whether we do or not and be able to evaluate those measures. I hope that our witnesses today will be able to give us some assurances that appropriate quality of care measures will be available in a timely fashion because I think this is important.

Our citizens also should be able to expect that they will actually receive the health care services that they need. The incentives in a managed competition plan are liable to encourage under-utilization. In fact, if one of the core problems of the fee-for-service—third party payer systems is, in fact, over utilization, then one of the core problems of the capitated managed competition system is liable to be one of under-utilization. This point was made by two of our earlier witnesses—Karen Davis and Stuart Altman, in fact. It is a point also made testimony written today by Mr. Brook, who was unable to be with us. I understand that his testimony is going to be included in the record.

The CHAIRMAN. It will be included in the record.

[The prepared statement of Mr. Brook appears in the appendix.]

Senator GRASSLEY. Thank you.

I am looking forward to learning whether our witnesses believe that appropriateness of care measures can be signed that will help us track not just the quality of care and the prevention of unnecessary care, but also those situations in which care may not be given when it should be.

Thank you.

The CHAIRMAN. Thank you, Senator Grassley.

Senator Rockefeller, who will soon be chairing our hearing.

OPENING STATEMENT OF HON. JOHN D. ROCKEFELLER IV, A U.S. SENATOR FROM WEST VIRGINIA

Senator ROCKEFELLER. Mr. Chairman, highly abbreviated, it is incredible that people would agree that either 20 to 25 percent of all that we do is unnecessary and inappropriate. It is stunning. It is absolutely stunning.

So the question is to understand why this is so. Then if we can to change the incentives so that this will not be so because those dollars are so scarce. Then the question, of course, is what really does constitute inappropriate or unnecessary care. That has to be answered honestly and it is a very, very difficult one.

I think many of us miss the importance of this. I think this issue is a cornerstone, if not the cornerstone, of the problem of our health care system. But it is almost never looked at systemically.

The question of whether one should undergo bypass surgery or modify their personal behavior and is there an incentive which leans one way or another, and then the sum of all of these millions and millions of decisions all across the country on a daily basis add up to these enormous amounts of money.

The question is, are we doing the right thing and what should we do? The fact of the matter, in concluding, Mr. Chairman, is that we really have not been able to get the data up until Senator

Durenberger, and Senator Mitchell, and the Chairman, and myself, and some others got something going called AHCPR back in 1989, which is now turning out practice guidelines, which hopefully will be adapted across the country.

But it is terribly, terribly important and it is a marvelous area and I am very glad the Chairman called it. Thank you, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Rockefeller.

So we will at once do our work. Dr. Chassin, and Commissioner, good morning, sir. You are first on our list.

STATEMENT OF MARK R. CHASSIN, M.D., M.P.P., M.P.H., COMMISSIONER OF HEALTH, STATE OF NEW YORK, ALBANY, NY

Dr. CHASSIN. Thank you very much, Mr. Chairman and members of the committee. In the 20 years that I have been a physician, I have been privileged to look at the American health care system from a number of different vantage points.

I have practiced emergency medicine for 12 years. I was the Deputy Director of the PSRO Program at HCFA. I did health services research at Rand for 9 years. And now I'm the Commissioner of Health in New York State.

I have seen the problem of inappropriate care in all those settings and have witnessed the harm that it does. There is no question that people suffer needless death and injury every day because they receive unnecessary health care services.

Inappropriate health care, therefore, is a quality problem. And what motivates me to eliminate it is primarily a desire to improve health. It clearly has substantial financial implications as well. Indeed, as Senator Rockefeller indicated, I believe that it is the key to our ability to control health care costs safely, without jeopardizing health.

Virtually all of the other more administrative cost containment approaches run the significant risk of adversely affecting health. But I also agree it is important to be clear about what constitutes inappropriate care.

I use the term to refer to circumstances where health care services are provided when their risks exceed their benefits. It is important to note that it does not include, in my view, the very difficult clinical decisions that physicians and families agonize over because reasonable physicians can reasonably disagree and there is really no clear answer. That is a separate category of clinical care.

The CHAIRMAN. Doctor, could I just say, you used a term that is new to me certainly. You said, where risks exceed benefits as against costs exceeding benefits which is the way an economist might view this instance.

Dr. CHASSIN. That is right.

The CHAIRMAN. You are talking about something different.

Dr. CHASSIN. That is exactly right. I am talking about the clinical decision to use a particular health service. Every health service that we provide can produce harm, whether it is the immediate risk of a surgical complication or in a diagnostic test the risk of falsely labeling someone as having a diagnosis when it is not present.

And inappropriate care refers to a circumstance where the sum of those harms is actually greater than the benefit that might be conveyed by that service.

Research, using that definition, has demonstrated that inappropriate care exists in virtually all sectors of the health care delivery system—from the use of medications, the use of diagnostic tests, hospital stays, surgery, et cetera. It is difficult to be sure about what the total magnitude of the problem is.

My own reading of that literature and my experience suggests that we could eliminate about 20 percent of what we do in health care and quality would actually improve because patients would be spared the risk of those inappropriate services.

Now, I would distinguish two different kinds of inappropriate care. The first is the inappropriate use of commonly done health care services. A number of studies have looked at this part of the problem. Some of them have shown results like 14 percent inappropriate coronary bypasses; and 20 percent inappropriate pacemakers. A recent Rand study found 16 percent inappropriate hysterectomies in a selected group of HMO's.

This part of the problem tells us something about how much might be saved of what we spend today on health care. But as you well know, there is at least an equally important, if not a more important, problem in cost and that is the rate at which health care costs rise.

Health care costs have been rising faster than costs in every other sector of the economy for almost 50 years. Most studies suggest the main explanation lies in all the new tests and treatments that we add into the system every year.

This phenomenon is not limited to high technology, to big machines like MRI scanners or PET scanners. It also encompasses the next new antibiotic that costs $90 instead of $3.50 for a 2-week course, the next new sleeping pill, the next new blood test, and so on.

We are terrifically effective at inventing this stuff. But we are perfectly awful when it comes to evaluating it, figuring out when it produces good outcomes for patients; and we are even worse at limiting its use to those indications where effectiveness has been proven.

Let me just rehash one example. It is an old one. About 20 years ago CAT scans were introduced to diagnose head injury. And it really allowed us for the first time in effect to look inside the skull without having to open it up to find out what was going on.

It immediately revolutionized the treatment of patients with head injury and immensely improved the quality of their care. Well, we produced these machines in abundance and what happened? Patients with head injury benefited. No question.

But as we produced more and more of them, idle capacity began to be a problem and we started to use them for confusion, for dizziness, for chronic dementia and all sorts of vague and nonspecific complaints in circumstances where benefit has never—had not been then and is not now—demonstrated.

That is the common pattern with medical innovation. It is shown effective for a very small group of patients under very specific clinical circumstances, but when it is widely disseminated it gets used

for far more indications where effectiveness has not been demonstrated.

Virtually every recent significant innovation has followed this pattern—coronary angioplasty, kidney stone lithotripsy, magnetic resource imaging, to name a few.

This kind of inappropriate care compounds the first. Today's commonly done, inappropriately provided health service was often yesterday's poorly evaluated innovation. The less well we evaluate today's innovations, the more inappropriate care we build into the health care system in years to come.

A couple of words about the causes of inappropriate care. I think fee-for-service reimbursement is certainly part of the problem. But its role, I think, is exaggerated.

The enthusiasm of physicians and others in advocating a particular treatment or procedure plays an important part. It is enthusiasm that drives the expansion of indications for specific health services from those where effectiveness has been proven into areas of no proven benefit. It is quite infectious.

Patterns of care that make specialists dependent on referrals from primary care physicians for patients are another cause. Factors outside medicine are important. As a society, we are infatuated with technology and innovation. The media write and broadcast stories that wildly exaggerate the potential benefits of new tests and treatments. Patients bring those expectations to their physicians, often demanding the latest new drug or procedure, even when there is no proven benefit in the circumstances they present; and physicians often find such demands difficult to resist.

How then might inappropriate care be eliminated? Because its causing are multiple and reinforcing, we cannot rely on any single approach. Altering physician reimbursement by changing fee-for-service to salary or capitation will not work by itself.

HMO's are prone to inappropriate care in part because their physicians are prone to the same enthusiasm as fee-for-service physicians. We have to use other tools.

For some services limiting the capacity of the health care system may work. I believe that this mechanism has allowed New York State, for example, to contain the inappropriate use of invasive cardiac procedures. Through a strict certificate of need process we limit the number of hospitals that are permitted to offer coronary angiography, bypass surgery and angioplasty.

Only 31 hospitals in all of New York State offer these services, compared with about four times that number in California as a comparison. New York also experiences—and I do not think it is coincidental—very little inappropriate care in these services.

Rand studies recently showed that only 2 percent of New York State coronary bypasses were inappropriate, 4 percent of inappropriate angiographies and angioplasties.

The CHAIRMAN. Could I ask you at this point, Doctor, we are trying to be free in these occasions. We are trying to learn. Maybe the whole panel could. Is there a confident procedure for determining what is inappropriate?

You do a double blind test and say, let us see about the inappropriate procedure. Will you come up right most of the time? Uniform, not necessarily right.

Dr. CHASSIN. This measurement is clearly in the beginning stages of its development. There are lots of refinements that need to be made. But I do think that the most clearly inappropriate care can be identified with a high degree of reliability.

The CHAIRMAN. This fellow gets a little nervous—2 percent inappropriate.

Dr. CHASSIN. Yes.

The CHAIRMAN. Two percent, plus or minus 5.

Dr. CHASSIN. Well, I do not think we know enough yet to be quite so quantitative. We don't have a gold standard. There are only a few methods for addressing this issue. Most of them depend on the expert opinions of physicians. Nevertheless, we can pinpoint inappropriate care quite accurately.

The CHAIRMAN. You have a protocol. You have a form to fill out.

Dr. CHASSIN. Yes.

The CHAIRMAN. Okay.

Dr. CHASSIN. That is right.

Senator GRASSLEY. The Chairman's question is very appropriate though in light of the fact that the Clinton plan is supposed to be up and running on January 1, 1997 and the appropriateness of care leg is very important. So the question is: If it is a long time from being developed, is it going to be ready when it has to be ready for the Clinton plan?

I think you are indicating that it is very difficult to say and consequently then it probably is not.

Dr. CHASSIN. Well, I think the principal problem with advancing the measurement methodology and producing appropriateness guidelines has really been a lack of investment. We have not made the appropriate investment in the technology, in the development of the measures, to get them to the point where they could cover a wide variety of procedures.

It is technically feasible to do that and to do that with relative expedition.

Senator GRASSLEY. So it could be done by January 1, 1997?

Dr. CHASSIN. We could get very close. We could develop between 25 and 50 appropriateness guidelines if we started work immediately with a relatively modest investment—$25 to $30 million a year.

Senator ROCKEFELLER. Mr. Chairman, if I could just add on to what Senator Grassley said.

The CHAIRMAN. Sir.

Senator ROCKEFELLER. This is slow because it needs to be slow. I mean this started really with the resource-based relative value scale legislation back in 1989, which is already 4 years. It needs to be slow.

When he says 40 to 50 practice guidelines, you are talking out of thousands and thousands of medical codes. I mean, it is really, really tough stuff.

The CHAIRMAN. Yes.

Dr. CHASSIN. I do not want to leave the impression that—I certainly agree that it is difficult. But if one focuses on the most commonly done services where the most common problems in overuse, are, I think we can make tremendous progress relatively quickly. But we have not made the investment.

Just a couple more points and then I will wrap up. I think limiting capacity can be important. I think focused quality improvement processes in hospitals, in physician practices, and in HMO's can also be effective. But I think that they will have to be supplemented by other incentives.

We should explore, for example, ways to use the reimbursement system to reward excellence. Today it is largely neutral or, in fact, perverse with respect to quality. I think publishing data on quality of care rendered by physicians and institutions is another potentially effective mechanism. It has to be coupled with specific quality improvement efforts and focused on specific clinical areas.

In New York for example, we have used this approach to improve operative mortality rates for coronary bypass surgery. Public education is also important to dispel the widespread belief that more health care always leads to better health. People suffer when they receive inappropriate care. The media need to be less effusive and more thoughtful in reporting about new tests and treatments.

I think a combination of these approaches can be very effective. Any one alone is likely to fail. If we fail, we will have little choice but to resort to the other more administrative ways to control costs, each of which has a substantial risk of doing harm. Rationing of effective health care services, drastic reductions in hospital and physician payments, substantial increases in consumer out-of-pocket cost sharing are all likely to have that effect. The effort to eliminate inappropriate care has to come first.

Thank you.

The CHAIRMAN. Thank you so much, Doctor.

[The prepared statement of Dr. Chassin appears in the appendix.]

The CHAIRMAN. As we told you ahead of time, Senator Packwood and I have got to be on the floor to deal with a measure reported out of our committee on extended unemployment benefits. That is where we are supposed to be. We are deeply grateful to you all.

I am very happy to turn over the chair to Senator Rockefeller who, along with Senator Durenberger, has been involved with this particular subject for so very long. And if you will my dear colleague. Thank you very much.

Thank you all, gentlemen.

Senator ROCKEFELLER. Dr. Wennberg.

STATEMENT OF JOHN E. WENNBERG, M.D., M.P.H., DIRECTOR, THE CENTER FOR THE EVALUATIVE CLINICAL SCIENCES, AND PROFESSOR OF EPIDEMIOLOGY, DARTMOUTH MEDICAL SCHOOL, HANOVER, NH

Dr. WENNBERG. Thank you. It is a pleasure to be back here again and have a chance to testify once again in this ongoing saga of practice variations. The basic facts are quite simple—how much care patients use depends more on where they live than on the disease they have. And in health care, geography really is destiny. That is what I have learned through 20 years of studying practice variations.

The interpretation that I think is reasonably taken from these facts is that the capacity of the health care industry, particularly the hospital industry, and the capacity of the physician specialty

work force are well in excess of the amount that is required to provide services that are known to work and patients actually want. I am going to emphasize "actually want" and I will come back to that.

More care is clearly not necessarily better. The policy implications seem to be quite straightforward. The nation can and should limit inflationary growth. We can do this without fear that such action will induce significant rationing of services if we do it correctly.

It is not only safe, but it is actually in the public interest to place restrictions on the growth of hospitals and specialists in the country.

If this sounds overly optimistic, let me give you a sense of the magnitude of the opportunity. If in 1989 the utilization patterns of Boston had been like those of New Haven—incidentally, these cities are quite similar demographically and all other relevant characteristics and they are both served by large, well-known teaching hospitals—if, however, the patterns in Boston had been like New Haven, 1,000 beds in Boston would have gone unused; 7,800 health workers could have been allocated to other tasks in the health care system; and a half billion dollars and some change would have been the difference in cost.

There are also big differences in the amount spent for physician services. The Medicare program, for example. If the rates in Miami, FL had been as low as they were in San Francisco in 1989, the program would have saved some $192 million. Again, it is no small change. Such opportunities exist throughout the country.

Practice variations occur in virtually every medical condition. They do that because almost every medical condition could be treated more than one way. In my written testimony I have listed Table 1, which contains some nine conditions, giving you the options that are available in the medical traditions for treating those conditions.

Some of them involve surgery, some will involve medical management, sometimes simply watchful waiting is a valid course of treatment. Those nine conditions account for well over half of the major surgery done in this country.

Now we've seen the tremendous variation that occurs in the rates of procedures for those that are listed in that table. For example, bypass surgery is twice as prevalent on a population base for residents of New Haven as it for residents of Boston.

The option being prescribed in Boston more commonly to treat this condition, namely angina, is medical management. On the other hand, the hip replacement rates are 75 percent higher for residents of Boston. It is a flip—higher on one condition for surgery and lower on the other.

Hysterectomy rates are twice as high in New Haven as they are in Boston. I am talking about the population-based rates. And on the other hand, carotid artery surgery is twice as common for Bostonians. Just to give you an idea of the kind of flip-flopping of rates.

Now each of those examples represents not necessarily underservice or overservice but different service. And the question really is which rate is right. And I want to come back to that.

Outcomes research is beginning to clarify the underlying structural problems, what is at stake in these decision processes. It is not correct to assume that it is already known. There is a good deal of dispute among physicians about what is appropriate medical theory and what the actual outcomes are that matter to a patient.

But probably the most important thing I want to say today is that better science is not enough. We really do not know how much surgery or other form of high tech medicine patients actually want. The true demand for care becomes apparent only when patients are informed about the options, about what is known and what is not known about the outcomes of the option.

When this happens most patients, in my experience, willingly accept responsibility to choose the treatment they want according to their own preferences. So reducing unwanted practice variations is not simply a matter of choice of health plan or even choice of doctor.

The fundamental problem is that if we want to discover the true demand for health care, we need a new relationship between doctors and patients, one in which the patient becomes a full partner with the physician so decisionmaking is shared. Shared decisionmaking goes way beyond the current notion of informed consent.

Now these statements come really from some of our own research in which we have been conducting outcomes research for patients who have prostate disease. When we began this research, we noted that in some parts of Maine, well over half of men were having proctectomies by the time they reached 85, whereas in other parts the rates were 15 percent to give you a sense of the variation.

Through a series of studies we could see that the theoretical basis for doing prostate surgery was poorly understood. We clarified much of that thinking and came to the clear understanding that the major benefit of this surgery is improvement in the quality of life, not the length of life. Tradeoffs, therefore, are between impotence, continence, operative mortality and a really good chance of urinating better if you have the surgery.

But which rate is right? This question really has to be answered by the patient. We set up some experiments to do that.

We learned that for this condition, the surgical resources needed to provide the care that actually works and that patients actually wanted was substantially less than now being provided in the United States.

When patients became full partners with their physicians in choosing their treatments, nearly 80 percent of men with severe symptoms, all of whom were eligible for surgery under appropriateness criteria, for example, actually passed up the chance for surgery. They chose watchful waiting.

When patients share decisions with their doctors, the link between supply and utilization is broken and the per capita rates of surgery actually declined in this case. The trend toward conservative treatment choice was evident, even in HMO's, where the rates of surgery were already relatively low, where patients faced no cost barriers at the point of delivery.

The per capita rates at Kaiser Permanente in Denver and at Group Health Cooperative in Seattle fell a full 50 percent when shared decisionmaking was instituted.

Let me direct the committee's attention to a second form of practice variations. This is the kind that arises in the decision to hospitalize patients or to treat them outside of the hospital.

In 1989 the number of staffed hospital beds in Boston was 3.8; for New Haven, it was 2.6. That difference in capacity is the principle source of the differences in expenditures which I mentioned a moment ago. The savings are there. Bostonians, by virtue of their excess capacity, experienced nearly 60 percent higher rates of hospitalization for a host of chronic and acute medical conditions, such as congestive heart failure, pneumonia and low back pain.

The residents of New Haven were much more likely to be treated for these conditions outside of the hospital, for example, at home, in a hospice, a nursing home or a clinic.

Variations such as these are typical of the entire USA. Per capita number of beds typically range from fewer than 2.5 per 1,000 to over 5 beds per 1,000. Numbers of employees follow the same pattern of variation.

What about the outcomes? In the case of hospitalization, we find no differences that point to the interpretation that more is better. In fact, mortality rates tend to be higher in the high rate areas than when differences in demographic characteristics and age structure are taken into account.

In my written testimony I ask, why, indeed, should greater spending be expected to bring about better results. Hospital capacity is not based on explicit theories about what works in medicine. The optimal number is unknown. In my opinion probably it will never be known in any strong sense.

Senator ROCKEFELLER. Dr. Wennberg, what page are you on?

Dr. WENNBERG. I have just been reading from my nonwritten material.

Senator ROCKEFELLER. Okay. I am sorry.

Dr. WENNBERG. I actually do not know how to refer that to you. In the text, it is towards the end. [Laughter.]

Senator ROCKEFELLER. That was not my point.

Dr. WENNBERG. The point is that the supply of resources that we have in the system really is based on custom and tradition and has nothing to do with explicit ideas about what works in medicine.

For example, the number of beds we need to treat pneumonia has never even entered discussion. They are just here and they are used. The same is true for the number of physicians. Currently the number of physicians we have in the United States is traceable to the policies of the 1960s, much of which began on Capitol Hill in response to a perceived scarcity of physicians.

But the numbers of specialists are really arbitrary. Indeed, the numbers working in an HMO is arbitrary from the point of view of true patient demand. Our study of prostate disease shows that even the relatively low number of urologists per capita hired by an HMO is in excess of the number they actually required when patients shared decisionmaking.

Learning what works and what patients want will have enormous consequences, in my opinion, for the health care economy. It breaks the link between supply and utilization. I believe that shared decisionmaking will lead to a generalized decline in demand

for high tech procedures. I do not know that, but that is my belief, my hypothesis.

We find that patients are more risk averse than physicians. But when they know they have an option, and that option includes more conservative treatment, many will prefer the conservative treatment to avoid the risks of invasive treatment, even though the benefits along some dimension may be better.

Prostate surgery, for example, clearly improves symptoms more than watchful waiting. And yet when patients were informed, 80 percent chose watchful waiting, wishing to avoid the risks of impotence, continence, operative mortality associated with the chance of feeling better.

I believe that if we generalize shared decisionmaking, many who are now accepting the recommendation for evasive treatments will choose more conservatively. If I am right, then a lot of money we are currently spending in health care can be reallotted without ration. If I am wrong, at least we will know the true demand, patient demand, rather than supply induced demand is the source of the issue.

Let me just quickly summarize. As I said health care, geography is destiny. The kind of care you get now depends on where you live. For those with chronic illnesses, those who are terminally ill, it depends to a large extent on the capacity of the hospital, whether you will be treated as an inpatient or an outpatient. This is a source of enormous differences in cost per capita.

The rates for surgery and for other technologies depend on how many physician specialists are in your area, and on their practice pattern; but we know how to solve these problems.

First of all, we can deal head-on with excess capacity, the numbers of specialists, the investments in hospital care. We also can expand the scope of outcomes research to find out what works in medicine. I agree with Dr. Chassin, continually patrol new technology. It is not just a matter of figuring out what works today, but it is setting in place the apparatuses for continually doing this.

Finally, if we give patients a real choice about their treatments, by implementing shared decisionmaking, we will, in fact, approximate a classic liberal market in this area where the choice is between treatments and options rather than leaving it up to the physician.

These steps really do break the link between supply and utilization; and I think offer the opportunity for a reformed health care system along a dimension which has not been seen any place else in the world.

Thank you very much.

Senator ROCKEFELLER. Thank you, Dr. Wennberg.

[The prepared statement of Dr. Wennberg appears in the appendix.]

Senator DURENBERGER. Mr. Chairman, as you know, I have to leave to make a commitment that we both have at 11:00 and you will be delayed. I wonder if I might just make two observations by the way of perhaps questions to leave behind that the panelists can comment on.

There is so much good in what they say that I do not want to suggest by these questions that it is a critique of any of the panelists because they are making a tremendous contribution.

I know Dr. Wennberg and I did have a chance to talk about some other things ahead of time, so I will not ask all my questions. Dr. Wennberg, looking at your testimony, when you get near the end you have a series of principles, one of which is, and I will just read from it, "an understanding of the epidemiology of medical care leads to the prediction that their entitlement would permit them to be absorbed into the health care system without loss of benefits for those now in and without any special increase in aggregate expenditures."

In other words, you are saying that we can absorb the demands of universal coverage, bringing in the uninsured without any special increase in aggregate expenditures. The capacity to treat the uninsured is already there. I understand that part.

And your conclusion or your premise I think is, in a steady state situation the increases in costs for treating the uninsured will be offset by savings realized by reduced utilization among those now insured. That is something we have heard commonly.

The question I am going to leave behind for you to respond to—and I regret that to fulfill a mutual commitment here I am going to get up and leave—is, I do not think we can use all of the consumer savings for expanded benefits.

In other words, if you want consumers and accountable health plans and doctors and hospitals in the system to begin to save you money, you have to leave a little of the savings with them. You cannot take all of the savings and send them off to somebody else's care.

The second point is that you cannot use all the savings in Minnesota, which is already saving you lots of money—I mean, compared to Long Island, New York you can get the same thing done in Duluth for about 40 percent of the cost.

You cannot use all of the Minnesota savings to extend coverage on Long Island or Miami or some place like that. So perhaps you could elaborate on that. And all three of you might, because you all make contributions to our understanding of the lack of evaluative science that we have in this country.

Using Dr. Phelps last comment as a premise for the question, Dr. Phelps says, "You must also understand that no single private entity—no health alliance, no health plan, no HMO, no hospital, no doctor's group, and certainly no single doctor—has sufficient incentive to provide the full amount of research in this area that our country should choose."

The Federal Government has a clear and crucial role here in the production and dissemination of information. I would suggest to you—and you know more about the country than I do, all three of you—that in Minneapolis-St. Paul, the combination of employers' accountable health plans and the physician community and the hospital community is already producing a lot of the information that Dr. Chassin had talked to us about.

There are incentives in the existing system to get that information out there as quickly as you can. And, yes, it does begin with

things we already think we understand and we just need to firm up.

But if, in fact, you do not penalize providers in the health plans for developing this information, they will develop it very quickly because it is in their interest, as well as the consumers to do it. So I would not by way of argument, but just suggest a question, that the right incentives in the market will produce—that part of the question—will produce a lot of the evaluative information that consumers need.

I would certainly agree that the dissemination of that kind of information and perhaps certain select kinds of research should be the function of a Federal agency of some kind. But I just want to lay on each of you the question, which is, with the right set of incentives, not paying them anything, but just rewarding them for doing the right thing, is it not true that some health alliances, meaning employer coalitions, for example, or health plans or HMO's like Kaiser and some others, the Group Health in many communities, would they not have and do they not already have the self-interest and the capacity to develop the kinds of evaluative information that we as consumers need?

Again, this is unusual. I hate to ask a question. Leave, not listen to the answer. But the Chairman is going to meet up with me in a little while and will give me the answer.

Thank you all very much.

Senator ROCKEFELLER. Dr. Phelps.

STATEMENT OF CHARLES E. PHELPS, PH.D., PROFESSOR AND CHAIR, DEPARTMENT OF COMMUNITY AND PREVENTIVE MEDICINE, SCHOOL OF MEDICINE AND DENTISTRY, AND PROFESSOR, POLITICAL SCIENCE AND ECONOMICS, UNIVERSITY OF ROCHESTER, ROCHESTER, NY

Dr. PHELPS. Thank you. I just received my invitation to attend here yesterday afternoon.

Senator ROCKEFELLER. Yes, and I want to comment on that. You are very, very good to accept so quickly and to already have your testimony up here and also to be from Rochester.

Dr. PHELPS. I was editing the testimony at 6:30 this morning, which is why some of you have not had a chance to see it previously.

The administration's proposed changes that could markedly alter our health care system, and with the goals of universal insurance coverage, improving quality of care and high on the agenda, reducing health care costs.

These goals potentially create a confining box from which it may be difficult to escape because expanded insurance coverage and higher quality inevitably add to costs, despite wishful thinking to the contrary, unless we do something to break out of this box.

A major new emphasis on cost effectiveness of health care interventions does provide a way out of this box. Now much has been written about inappropriate care. And Dr. Chassin's testimony this morning discusses elimination of medically inappropriate care defined as where medical risks exceed the medical benefits.

Removing such care from the health care system is laudable. I think most people would agree with that. Today I wish to impress

upon you an equally important goal: We must not only eliminate care where the medical risks exceed the benefits, but also care that creates little improvement in patients well-being per dollar of resources expended.

It simply makes no sense to introduce or mandate a budget process that ultimately limits per person spending on medical care in this country and then ignore in our considerations of what is appropriate or inappropriate.

Unfortunately, we do not possess the knowledge of how to do this in many areas right now and we will need to make major new investments to produce and disseminate information about how to carry out cost effective medical care. I will return to this, as Senator Durenberger's question later as well.

The administration's proposed health care plan and many others eventually propose a cap on total spending. Once such caps are in place, it is fundamentally necessary to introduce considerations of cost as well as medical risk into the analysis of what is appropriate and inappropriate in order to achieve the best possible health outcomes for the people of the United States.

How can we achieve this goal? Newly developed methods provide specific ways of measuring benefits to people's health using the same yardstick, no matter what the medical interventions—quality adjusted life years and some variants of this.

Dr. Wennberg has talked about how considerations of quality instead of just life expectancy are very important in some of these issues. It is quite easy to show that for a given budget for caring for a defined population, the best health outcomes emerge if, and only if, the use of such interventions is organized in a way so that the added benefits relative to the added costs for each care or each service you provide are brought to the same level.

Let me provide the intuition behind this result. People commonly think about cost effectiveness of an intervention as something cast in bronze like a statue in a park. Thus, people say that coronary artery bypass or "CABG" surgery costs $30,000 per life year saved, while mammography for a woman costs only $10,000 per life year saved, or quality adjusted life year.

Alas, neither of these statements is really correct in a meaningful sense. CABG surgery is incredibly cost effective, less than $8,000 per quality adjusted life year, when performed on a subset of very high-risk patients with a specific definition of disease.

However, extending that surgery to people at less risk, with milder symptoms, brings quality adjusted life years cost up to $65,000 and more per quality adjusted life year.

Annual breast examinations for women in the 55 to 64 year old age group costs only $11,000 per quality adjusted life year. Adding mammography to the same group brings the cost per daily to about $30,000 per life year. Screening younger women would increase the cost per quality adjusted life year. Increasing the frequency of screening would increase the cost per quality adjusted life year.

There is no such thing as a good or bad intervention. In a sense, there is always somebody who will benefit from a medical intervention as we know it now, as Dr. Chassin has pointed out. The question is finding out for whom we would get the greatest benefit for dollars of resources expended.

If we have a fixed budget for treating the defined population, what happens, for example, if we eliminate one coronary artery bypass surgery on these patients with moderate angina, saving say $20,000 and losing on average about three-tenths of a quality adjusted life year.

You could take that same $20,000 investment, shift it to some mammograms for women, say, costing $200 a piece, buy 100 mammograms, and they will create about seven-tenths of a quality adjusted life year on average for the population of women. So we get an increase in health outcomes by shifting resources from relatively expensive procedures in cost per quality adjusted life year to those that are lower relative cost.

We can adjust every intervention in our health care system to a common rate. It is not just an issue of acute therapy versus prevention. The same idea works equally well if we start out with the premise that all women have been receiving mammography as well as breast examination but the CABG is only performed on people with left main artery disease, a high-risk group.

In that case, proper allocation of resources would shift resources away from mammography towards doing a little more coronary bypass surgery, as puzzling as that might seem to some people.

But these examples only hinted at a pervasive phenomenon that Drs. Wennberg and Chassin have also described. The cost per quality adjusted life year differs hugely across medical interventions in the United States, sometimes at $100 per quality adjusted life year and ranging up to hundreds of thousands or millions of dollars per quality adjusted life year.

The ways we use these interventions can and should come under increasing scrutiny in the new cost-conscious environment. What does this mean for managed competition and other forms of cost-conscious health care? As the intervention spreads to wider groups, as I said the bang for the buck invariably declines.

We do not need to mandate the use of cost effectiveness criteria in order to get their use in managed competition environment, because a group that is trying to provide good care for their people and compete on quality and cost will automatically begin to use this type of analysis.

Unfortunately, we do not have a complete understanding of what the cost per quality adjusted life year are for many interventions in the health care system. Thus, eventually a limiting feature of managed competition appears to be a new form of knowledge that our society does not now in general possess—an understanding of cost effectiveness of medical interventions as they are used for different populations, different intensity and different frequency.

We continue to invest in the production of basic biomedical knowledge at a high rate. The NIH budget this year will be around $9 billion. By contrast, AHCPR, their new budget this year, is just under $150 million, about $1 per each $60 spent on the production of basic biomedical knowledge.

Studying the relationship between medical interventions and outcomes provides the best way to determine cost effectiveness of various types of interventions. Studies to do this, just as basic biomedical research, require important funding.

Immediately funding is needed to carry out the most important studies. We cannot do them all immediately. However, also of urgent need will be increased research funding for people to carry out this research. We will not get it all done immediately because we do not have the capacity. So we need not only to invest in new studies but in training of people to do this sort of research.

I have conducted some studies through the Institute of Medicine suggesting that the returns——

Senator ROCKEFELLER. What is the training for that research?

Dr. PHELPS. Pardon?

Senator ROCKEFELLER. What is their training for that research?

Dr. PHELPS. They have a variety of trainings. Masters of Public Health would be the most common path now, commonly with an MPH degree. Dr. Wennberg's group has organized master's degree in clinical evaluative sciences. Ph.D. training in many of the social sciences, blended with the medical training can provide this I think M.D. Ph.D. training is terribly important. Epidemiology is important. Psychology is important in order to understand patient's preferences. Almost any of the social and behavioral sciences and epidemiology are the key training areas.

Senator ROCKEFELLER. But those specialties are attracting students, not so much because of something called AHCPR or National Health Policy, but because of academic health centers and their own particular pursuits? Or have they heard of this thing called AHCPR?

Dr. PHELPS. I think it is going to take at least in the short run a fueling of training grants to produce opportunities for schools to get tooled up to train here. Eventually I am sure the demand is going to be there for people trained like this in health care alliances, health plans, insurance plans, government, private sector, to help everybody learn how to do cost effective medical care.

Right now we just do not have the number of people available who know how to do this really well.

Senator ROCKEFELLER. So that you see, for example, in the health care alliances we always talk about somebody will be looking at cost and somebody will be looking at quality, and you see the quality people as having the kind of training that the three of you have been talking about? It is not just how good is it, but how necessary is it.

Dr. PHELPS. I see the issue as the same set of people knowing how to measure the quality, knowing how people value the services and the output, and understanding the cost implications of providing it; and having people, research team or individuals that contain all of those ideas together and combine and blend them.

Senator ROCKEFELLER. Should we write that into the legislation or would that be very threatening to the medical community?

Dr. PHELPS. I think yes to both your questions, sir. I think it would be important to introduce the idea of cost effectiveness into legislation. I think much of the medical community would find it threatening.

Let me just close briefly, if I may. I have done these research projects for the Institute of Medicine evaluating the potential outcomes benefit from this type of outcomes related research. For all of the current interventions in the health care system, the returns

to this type of investment are tens and hundreds and sometimes thousands of times larger than the costs of providing this information.

The yields to this kind of research are just astronomical as compared to the costs of doing it. If we had these kinds of research opportunities in other areas, we would leap to them instantly. And despite the increased funding in the AHCPR and the shift in emphasis in some cases in the NIH and the Centers for Disease Control, I still think there remains huge untapped opportunities to improve the quality and the cost outcomes in our health care system by bringing cost effectiveness considerations clearly into the discussion.

Let me stop at that. Thank you, sir.

Senator ROCKEFELLER. Thank you very much, Dr. Phelps.

[The prepared statement of Dr. Phelps appears in the appendix.]

Senator ROCKEFELLER. Senator Grassley, do you have some questions you would like to put forward?

Senator GRASSLEY. Yes. My first question would be just for Dr. Wennberg, and my next two questions and that's all the questions I have, could be for anybody who wants to answer, appropriate for any of you to answer.

Many physicians tell us that they prescribe certain procedures only because their patients insist upon them. You are taking the opposite view, as I understand it, that the patient is, and will be, more conservative than the physician.

Can you disentangle this for us? Are patients likely to demand the non-threatening diagnostic procedures as contrasted with a potentially—we have a vote coming up.

Senator ROCKEFELLER. It is not meant to be until 11:30. What did your beeper say?

Senator GRASSLEY. Well, I turned it off.

Senator ROCKEFELLER. Good.

Senator GRASSLEY. 11:45.

Are the patients likely to demand the non-threatening diagnostic procedures as contrasted with the potentially dangerous surgical interventions or are the physicians who make this argument just being kind of self-serving from your point of view?

Dr. WENNBERG. I think the problem is one of a sea change in perceptions about the role of the physician. Traditionally, everyone has believed that physicians were competent to prescribe. In essence, they could understand the patient's values through their clinical experience and we all left the choice up to the doctor. What shall I have, doctor?

But as technology has become more powerful to intervene and more options are available, common problems are treated in many different ways. And it turns out that the different treatment have different outcomes. It is not as if the same outcome is achieved, for example, by watchful waiting in the case of prostate disease as is achieved by surgery. It is just a different outcome, different set of outcomes.

And it turns out that because it matters how the patient weighs the different benefits and risks of these outcomes it requires us as physicians to ask patients what they want. In our studies, for example, we could find nothing in the physical examination or in the

history or in the laboratory tests that would let the physician prescribe surgery for patients who had severe symptoms. We had to ask them.

And when you asked them, it turned out that only one out of five actually chose surgery. See? The point here is that in order to learn what people want, we need to reorient the doctor/patient relationship to one in which we actually ask the patient.

It is not that physicians have willfully controlled demand. They have been taught that way. When I went to medical school, I was always told that I knew what I was doing. It was a very important thing to believe, that the ideologies of medical practice were part of the initiation that you went through. And yet how different those initiations as we looked at the variations in practice from area to area. It is clear that there is not a homogeneous pattern of practice. Otherwise we would not be here today talking about it.

Senator GRASSLEY. Are patients more demanding today?

Dr. WENNBERG. A very interesting question about demand. Now, remember, I mentioned to you that if you live in New Haven the chances are twice as great that you will have a bypass surgery than if you live in Boston. But just the flip opposite in terms of hip replacement—the chances higher in Boston than in New Haven.

These steep gradients in surgical patterns are not following any kind of a distribution you would expect if patient preferences were doing this. You would have to say people are very strange in Boston to want twice as much surgery for that condition, but half as much for the other.

But the most telling evidence for supplier-induced demand is when practice patterns change radically. Tonsillectomy rates, for example, in one Vermont community dropped from a point where 65 percent of children were getting their tonsils out by the time they reached age 10 down to virtually zero. Nobody went out of town to get it.

In what other market could one see such a rapid change in supply without anyone going to the next city saying I need this because I cannot get it locally anymore because the suppliers have changed their attitudes.

I see very little evidence that consumer demand at the point of decision for surgery or invasive treatment is the source of the variability. I am not saying that some patients do not come in the office demanding it, but it is not the reason for the variation.

The most interesting point perhaps on that all is, my statements about the HMO experience in Seattle and Denver. When we took the pain to provide information in a structured way, we saw a radical shift downward in demand when shared decisionmaking replaced the old model of doctor-prescribed treatments.

Senator GRASSLEY. Dr. Brooks in the statement that he prepared for this hearing, and I want to quote these words, "Changing incentives from fee-for-service to managed care or to global budgeting is likely to lead to tremendous pressures to reduce the use of services." He goes on to say, and then I want to quote again, "It is critical that a strategy be developed and implemented to assess under utilization of appropriate and necessary services for those people who are at risk of not receiving needed care."

And hence, question. Would it be your view that appropriateness of care measures can be helpful in protecting against under utilization in the managed competition system?

Dr. CHASSIN. I will take a crack at that. I think that they can be and I think that systems to monitor for underuse are very important as reimbursement shifts from an emphasis or from an incentive to overuse, as you pointed out earlier, Senator, to an incentive to underuse.

But I think we need to look at reimbursement incentives that reward good quality and that are not perverse with respect to quality either by placing a premium on overuse or by placing a premium on underuse. We have not explored that possibility very well.

I think the other important point to make is that as it becomes more and more imperative to reduce cost, it will be much more attractive to resort to administrative solutions that reduce utilization. We know how to do that. Increasing cost sharing will reduce utilization. Decreasing payments will reduce expenditures and will reduce utilization.

But we also know that those methods reduce appropriate care about as much as they reduce inappropriate care. And without a selective approach to reducing inappropriate care selectively, thereby preserving access to needed and effective services, we are likely to do some harm along the way instead of focusing on removing the inappropriate care first.

I certainly agree with both of my colleagues that we need much more effectiveness research. We need to perfect the information that we give patients as well. But right now we expend resources and do harm. It seems to me ameliorating that problem requires a very high priority.

Senator GRASSLEY. Senator Durenberger touched on this a little bit. But should the development of quality measures be a public function and should the adoption of publicly developed quality measures for the evaluation of health plans be required by law?

Dr. PHELPS. Perhaps I could at least open the discussion on that. There are some types of measures and information that not only can but should be developed locally—health plan levels, alliance level, hospital staff level. These are important things to continue and carry out.

There are also certain types of things where the costs of doing the basic information gathering are so large that they just overwhelm even very large insurance plans. Senator Durenberger asked about the capability of producing some of this information locally. In order to study the effectiveness of a treatment for a disease that occurs fairly rarely, you must go beyond the confines of single regions.

Large randomized controlled trials to study health outcomes commonly enroll patients for 2 or 3 or 4 years in sometimes 20 or 30 different communities in order to capture enough patients to make valid scientific inference.

No single health plan, even an alliance in a large state can be expected to do that sort of thing. They just do not have the money. They do not have the time. And in some cases they do not have the capability. If you do those studies once nationally, everybody

can benefit from the results so that type of research in particular, is where the Federal role looms largest.

Dr. WENNBERG. I would just like to add to that science privatized and made proprietary is one of the problems that is on the horizon if guidelines become properties of managed care companies. Sometimes doctors don't even know what they are. And certainly I do not think anyone would argue for that strategy. We need an openness to our science. We need a peer review strategy for it and it must be available essentially to everyone.

There is a right way to doing science. It is something that should be public knowledge and not private knowledge. Certainly, however, the idea of developing strategies for improving quality will have to be a continuum that goes all the way from a national strategy to a regional strategy to the plan level strategy.

Let me give you an example of why that is an important distinction. The economy of scale necessary, for example, to understand the process by which bypass surgery is produced probably requires a population of at least 2 or 3 million people—I would say 3 million people—in order to pool enough experience to get the benchmarking that you need to compare Practice A, Practice B, to Practice C to find out what is really going on.

A very simple example of that is some of our own work in northern New England conducted by Dr. Plume, Dr. O'Connor and their colleagues. It turned out that one of the most important reasons for variations in surgeon specific mortality rates was where the incision was being made. One incision resulted in an infection rate of about 2 percent or 3 percent. And it was only when they pooled the data together that they could see that this presumably arbitrary piece of the process for producing a bypass operation, i.e., how the incision was made, was in fact causing excess mortality in those patients.

Learning this required the pooling of about 3,000 cases and the experience of about 21 surgeons in an environment where they could examine what they were doing and learn what they were doing to get at the result.

Senator ROCKEFELLER. Senator Grassley, can I just interrupt on that? Which is a very difficult thing for physicians to do, is it not? In other words, we do everything else in life. We have data. For business, the business schools they teach that. In other words, you aggregate procedures as a substitute for incisions. And then you get a flow chart from that and you make a decision as to what is best.

But in medicine physicians partly because of the way that you say you are trained and partly the way that we were also brought up to accept what it is that you say, that is not well received, is it?

Dr. WENNBERG. It is not well received, and moreover it is different. If it were a mechanical process and a machine that is doing it, it is easy to retool or get rid of the machine. But when it is a professional that is doing these "mistakes"—and they are not mistakes, they are basically doing it differently—it requires a workplace free of fear and a learning environment in order to actually get to the sources of variation.

I hope the committee will pay attention to that as you go along, because one of my great concerns is that we will create such fear in the workplace that we will make learning more difficult rather than easier.

Dr. CHASSIN. Let me expand just a moment on that and come to your question as well. I think that most physicians would, in fact, accept that method of learning. But we were trained that way. As Jack has indicated, we are mostly trained to learn from our own individual experience, which is not terrifically helpful in improving the quality of the care we provide.

We do have a couple of obstacles. One is the punitive way in which quality assurance has been implemented in the past and, indeed, is still implemented today. The shift from a punitive quality assurance environment to a positive, supportive quality improvement environment is one of the fundamental shifts that has to occur before physicians can become actively engaged in the kind of quality improvement that Jack has suggested and I think everyone would support.

Whether quality measures can be developed locally, I think the answer is largely no. There are some measures that are of importance to patients, plans, and payers that can be developed locally. Those are largely issues of patient satisfaction, waiting times, and characteristics of the particular setting in which health care is delivered.

But I entirely agree with Jack, the fundamental issues of quality, of the processes of care, and of the way care is delivered in a clinical fashion must be developed nationally. There is no difference in whether a coronary bypass operation is appropriate in Des Moines or New York City or Rochester. Those are judgments that arise from clinical science and experience and should be national.

The investment needs to be made at a national level as well. But I think it's important also, I have to emphasize, that clinical research on effectiveness is critical. We need to expand it. But we do not need to wait for that clinical research to approach the problem of inappropriate care.

And, in fact, I would argue further that it is unreasonable to expect clinical research to clarify many issues in inappropriate care. Think for a moment what it would take to clarify that what we think is an inappropriate reason for doing coronary surgery is inappropriate. You would have to get a group of patients who have this reason for doing a procedure that everybody agrees is inappropriate, that will produce more harm than good, and then convince some research committee that this is a good idea to do a research study on. That will never happen.

So inevitably, defining the bounds of acceptable practice and what is on the other side, what is inappropriate, will always be a job that will rely on expert opinion and will be based very little on rigorous research.

Senator GRASSLEY. In our discussion when Senator Moynihan was still here, when you were giving your opening statement, you stated that with more resources into developing these guidelines, we could probably get them brought on board in a fairly short period of time.

Would keeping them up-to-date be a major problem? You know, once they are in place.

Dr. WENNBERG. It is a significant problem. Depending on what particular service or procedure one is talking about, it may be a major or a minor problem. Some things change rapidly, other things do not change much at all over time.

But you have identified a very important feature of this effort, that many developers of guidelines and quality criteria do not attend to well enough. Another reason why a national effort focused and prioritized is, I think, essential.

Senator GRASSLEY. Thank you, Mr. Chairman.

Dr. PHELPS. Senator Grassley, let me touch briefly on this also if I may. The question of continually updating our knowledge is really important here. Continuous improvement on the research side is just as important in some sense as continuous improvement in the final production of the services by hospitals and doctors and other providers.

I guess I would like to be a little more optimistic than Dr. Chassin about the eventual role of science in providing information to construct these guidelines and bring them into fruition. There will always be areas that are clearly right and clearly wrong. There will always be gray areas as well.

These gray areas are, in fact, right where we should be conducting our scientifically based research. And because they are gray, and because there is ambiguity, and because there is valid scientific and professional disagreement, they represent areas where it is perfectly legitimate to carry out randomized trials.

So in the long run, with proper training of researchers and a change in shift and mentality, I think we can actually put a strong science base in most of these types of decisions. But we do not have that now and we must move ahead without that using qualified expert opinion and other mechanisms.

Senator ROCKEFELLER. Thank you very much.

Senator Daschle? You need to say something.

OPENING STATEMENT OF HON. THOMAS A. DASCHLE, A U.S. SENATOR FROM SOUTH DAKOTA

Senator DASCHLE. Well, Mr. Chairman, I am delighted to be here. I have a conflict and so I will excuse myself. But I appreciate the opportunity to welcome our witnesses. I understand from my staff that they have done a remarkable job in talking about this critical issue. I hope we have an opportunity to bring them back as we talk about this as it relates to health care.

Senator ROCKEFELLER. Henry Shakespeare would have been jealous.

Senator DASCHLE. Thank you.

Senator ROCKEFELLER. This is to anybody. You talk about how other doctors are trained and other health care providers. There are some who sort of have this concentric theory, concentric circle theory, that when you go to medical school that this becomes your learning, everything which is in the concentric circle. Then you do residency. You practice a bit. Have some children. And then there's another concentric circle, which is the concentric circle of your cur-

rent—the way you practice medicine, what you believe, what you have learned since medical school.

There is a very interesting fact, that is that the concentric circles, whereas they overlap, they are by no means the same, which implies if I am correct, which I guess I had better get you to validate or not validate, there is a difference between what people learn in medical school and what it is they might practice 10 or 15 years later, part of which would be necessary and good because of the change of medicine.

But would you disagreement with?

Dr. WENNBERG. No.

Dr. CHASSIN. No.

Dr. PHELPS. No.

Senator ROCKEFELLER. No, okay.

Then how is it that we influence physicians about this question of appropriate and inappropriate care? In other words, let us supposing AHCPR brings out these practice guidelines. They mean nothing unless they are adopted.

Then the question is, what is it that physicians listen to. Do they listen to their peers? Do they take it out of the New England Journal or JAMA? Is it collegial talk in a hallway or in a dressing room? How is it that physicians themselves best learn, most comfortably to themselves learn, about matters which would then be likely to change their behavior which they have cherished and which they have seen themselves as unique?

Dr. CHASSIN. You have raised, I think, one of the critical issues in actually effecting some of this change. Now we actually know a fair bit about how to change physician behavior in ways to improve quality. We know a fair bit about what works and what does not work.

What does not work is mailing information out in envelopes that physicians receive in their offices and typically toss in the waste basket. And yet that is the way in which most new information is communicated. It is either written in journals or sent out by agencies that develop guidelines.

We do know that from an educational standpoint, focused interventions that represent the communication of clear and well-developed guidelines work. One of the problems with past efforts is that guidelines have been so ambiguously worded and so general that they are not helpful clinically. So you have to start with a well constructed message.

If that is amplified by data on performance so that physicians can see how well they are doing compared with the guideline—that approval has been shown to change physician practice for the better.

So the well constructed guideline along with data on performance, followed up by feedback as practice changes, followed up by a monitoring that then looks 6, 8, 12 months later to see if practice is actually improved, that is the method that has been shown to demonstrate improvement as this has been studied in the past.

Senator ROCKEFELLER. Let me ask one more, Dr. Wennberg, one more part to that question. If this health care bill passes, and health alliances are a part of them, and if consumer information is going to be as big a part of that as I hope it is, it will not just

be in liability cases where the national practitioner data bank will be available.

But an enormous amount of information about individual physicians or groups of physicians in a plan or in hospitals, and the success rate, which can be different than the quality, can be different. The success rate they have had. The cost of procedures which they have made available either as physicians or as hospitals doing these procedures will be available in a more and more refined state to consumers as they choose health plans.

I just want to introduce that possibility in terms of doctors and the way they might look at that, react to that, et cetera.

Dr. WENNBERG. My experience, Senator Rockefeller, is that the most important missing ingredient in the public debate so far—to ease right off your statement—is the assumption that when patients choose between plans we therefore rationalize the health care system. In other words, that that act in itself assures quality is a supposition which should be carefully examined.

I tried to emphasize in my testimony that the critical choice for patient is not between plan, but rather between treatments when they have a particular condition. What is important is that a plan or a system of care assures that patients are informed in ways they understand about the uncertainties as well as what is known about the treatments that are available to them.

And secondly, that the workplace, that is to say the environment in which the doctor/patient relationship occurs is a non-coerced one in which patients can choose according to their preferences among the elements of the benefit package that we as a society have agreed are part of it.

I am not saying that they should choose inappropriate care or care that does not work, but between prostatectomy and watchful waiting, both of which work, that choice is critically important and will depend on what the patient's values are.

So the question of how to realize the shared decision making environment becomes to me one of the most important tasks before the Congress. If you can achieve an environment that promotes this sense of the ethic of the doctor/patient relationship, many of these issues about unnecessary care and inappropriate care—can be solved. The key problem is appropriate care used inappropriately, something that works in general, but does not work for me because I do not want it. I believe that is the key to the economic as well as the ethical problems that we are trying to grapple with here.

Senator ROCKEFELLER. Take me a step further here, how it is we can construct that in something called health reform.

Dr. WENNBERG. One of the things I would recommend is that certain ethical principles be established by Congress about what alliances and plans should do about shared decisionmaking. Plans should state their strategies for informing patients about options. How the represent options that encompass the current state-of-the-art with regard to appropriateness research or outcomes research, where the field is right now.

Let me give you an example of the link between appropriateness research, outcome research and shared decisionmaking. In the case of prostate disease, we know an awful lot now about what is at

stake in that decision, which was not known 7 or 8 years ago before this research took place. We know that one of the major theories that was being used in the community was wrong. That was that early surgery would make you live longer because it would prevent bad events like kidney failure and so forth and it does not do that.

There is not a life expectancy benefit for most people with prostate disease, even though many surgeons were operating under that principle. Care under this theory is inappropriate and should not be offered by the plan. That got clarified through outcomes research. And the probability estimates for the different outcomes that matter to patients, were identified and estimated. Then the value side of the equation enters. What the patient wants is the responsibility of the patient, not the doctor. Participating in this equation leads to the right choice.

Another example would be—and I would say——

Senator ROCKEFELLER. But then you talk about the responsibility of the patient with great goodness. I mean, the patient after all is not trained. The patient is making an intuitive immediate, probably millisecond judgment or maybe he goes home and talks with his family about it overnight.

Dr. WENNBERG. Now remember we are talking about decisions that you can take some time over. This is not an emergency situation. And most of the surgery in the United States is not emergency, it is elective.

Senator ROCKEFELLER. One has time for it, yes.

Dr. WENNBERG. The thing, you see, what is so fascinating about the work that Dr. Mulley, Dr. Wagner, Dr. Fowler, and Dr. Barry have been doing is that when patients become involved it is no longer the objective level of their symptoms as I mentioned that matter in their decisions. It was the degree to which they were actually bothered by their symptoms, which is different than their objective level, and the degree to which they were concerned about the risks of sexual dysfunction from the surgery, that predicted choice, you see.

That becomes then the fundamental route to rational choice. I do not believe this particular condition is different from any of the others on that list in table one.

Senator ROCKEFELLER. But then in that case, and this is addressed to any of the three of you, if you are a meeting of the American Medical Association it is best not to bring up the name of Dr. Lawrence Weed. But one of the things that he suggests, I think wisely, is that the physician and the patient engage early with each other, side-by-side so to speak, facing the computer, as they begin to make not just choices, but make an assessment of what the patient's health might be.

The whole concept of it is—this interests me very much—not just the physician making the choice, but also as you indicated the patient making the choice because that will often determine, particularly as patients become more cost conscience, more quality conscious, which I think they are bound to be, even in the next year as we debate all of this, much less pass any of it, and also as physicians see themselves more in cooperation with patients and as physicians are seen to be less god-like, so to speak, as you indicated, which I think is a phenomenon, which is happening.

Therefore, the education of the patient becomes incredibly important. And, therefore, you really do not want a patient to be making just a judgment based upon, let us say, am I going to have prostate surgery or am I not. Does it not have to be a longer process?

If I can suggest the approach of Dr. Weede—I have a first cousin who has done a lot of work with him and therefore I follow that—that the patient and the doctor sitting together, analyzing, talking to a computer, each of them side-by-side and the symbolism of that, talking to a computer and analyzing the patient situation and then making decisions virtually together based upon the information that the computer—the choices that the computer rules out and rules in, et cetera.

Dr. WENNBERG. This is not the actual strategies that we use. We use interactive video with the computer to provide the probability estimates to the patient and film clips of patients who had different experiences. So the patients' vicarious understanding of questions of the future, do not depend on the storytelling skill of the individual physician. But rather you could lay it out in a consistent way.

The thing that was most fascinating about this is that the patients almost to a person—I guess I can say to a man in this case—actively wanted to choose. In other words, it was not as if this was something foreign to them. Once they understood what the structure of this problem of choice was, they had an opinion. They knew what they wanted.

The other thing is that they can always change their minds. See, in other words it is an ongoing process, at least those who take watchful waiting.

Dr. CHASSIN. I think there are a very important class of decisions that are assisted materially by that kind of process. They are largely situations in which a number of alternatives are available and there is no one that clearly dominates, that clearly in the vast majority of circumstances produces better outcomes.

Then which set of outcomes you prefer, as Jack has suggested, is a very important ingredient in the decisionmaking. That is not inappropriate care.

Inappropriate care is something I think patients need to be protected from. And, in fact, mechanisms that we set up in government, either at the Federal level or the state level—or that the health plan sets up—must, I think, take into account that aspect of inappropriate care. Patients, because they will be harmed by it, need to be protected from inappropriate care.

I think consumer information also plays an important role, but not so much in the prevention of inappropriate care. There are three basic dimensions of quality. One is appropriateness. Did you pick the right thing to do?

The second is, technical quality. Did you do it well? Well enough to realize its maximum potential benefit.

And the third, and quite separate from those two, is patient satisfaction. Patient satisfaction can be measured at the local level. We have talked about that. I think there is no disagreement.

Technical quality can also be measured. And, in fact, we have measured in New York now for the last 5 years risk adjusted operative mortality for coronary bypass surgery, published that infor-

mation by hospital and by physician. And I think that by itself has been an important step.

But we have observed a dramatic fall in risk adjusted operative mortality, not because the information was used by the public to change doctors, but rather because it was an incentive to the physicians and hospitals to pay attention to this problem and to look behind the data, to look very carefully at their processes of care, and to improve their outcomes.

The peer pressure of the data being public was a very important incentive in that process. We should not lose sight of that. I think that, in fact, that will be largely the way in which public release of information will have its impact on quality.

Measures of appropriateness will not be very readily available to consumers as they are making their choices. That is going to have to be addressed by a different mechanism.

Senator ROCKEFELLER. Did you not a moment ago indicate that physicians are not readily subject to influence by journals or published data or did I hear that wrongly?

Dr. CHASSIN. It has been very difficult to show that published data by themselves have influenced physicians.

Senator ROCKEFELLER. So how are we going to do this? Look, I mean, we have this enormous problem and it is called a trillion Republic and private health care, wrapped all of us in the nation for the first time, turning its attention as a body politic, 260 million people, to something called health care. Never before discussed on such a broad level.

So you have to bring the cost of it down. You absolutely have to do that. As soon as you talk about that then people immediately start assuming that the quality is going to degenerate. You are saying not necessarily so. On the other hand, that is a very big sale as far as the public is concerned.

In fact, I would suggest probably as far as the medical profession is concerned. Because once the Congress or something called the government, particularly if it happens to be done by a government run by Democrats who start bringing the cost down, the word "quality" immediately gets doctors very, very nervous.

I mean, this is sort of a very large question. But, I mean, how do we do this in health care reform, in the discussion of it and in the creation of policy.

John Chafee, I apologize. I did not even know you were here, sir. I will make this my last question.

That you bring down cost, that you address appropriateness, that you have health alliances. They pummel out information to their consumers about these things called accountable health plans which have to do with doctors and hospitals, that the physician community is newly energized, both negatively and positively because of this thing called health reform coming up which they know needs to happen but they are not sure they want to have it happen because of something called Washington, D.C.

I mean, how do we get our hands on appropriateness, bringing down the cost, making sure of quality, making sure that patients have a degree of comfort in the direction of this, and making sure that physicians and other providers have that same degree of comfort?

Dr. PHELPS. Well, I guess the first thing to understand is, there is no single magic bullet here.

Senator ROCKEFELLER. We have certainly agreed on that.

Dr. PHELPS. We have to work on a lot of little nibbling away issues. Provision of information, such as Dr. Chassin has suggested about quality of severity—adjusted outcomes is very important, very useful.

I am an economist. I think about incentives to do things as my livelihood and you had asked earlier, how do we get doctors and patients to begin to use this information, these guidelines and so on once we produce them and we get them right.

There are ways to do that. For example, the legislation, as I understand it, proposes modifying malpractice law so that behaving according to guidelines essentially creates a de facto malpractice defense. That is a good incentive to get doctors to pay attention to what those guidelines say.

We also have to think about the incentives for disseminating information. These patient information systems that Dr. Wennberg and his colleagues are working with, making sure that those get paid for from the health alliances to the health plans is really important.

If those get cut out of the reimbursement system or there is no special incentive to put them in place, they are not going to emerge as rapidly. Drug companies know how to disseminate information about new drugs really well. They send out drug detail men and women around the country and there are these armies of people in the interstices of the health care system describing the effects of these new drug treatments to doctors.

Senator ROCKEFELLER. Along with lots of samples.

Dr. PHELPS. There has been some work suggesting that maybe we need academic-detail men out there spending time, minimizing, reducing the cost, easing the access of doctors to this kind of information. There are a lot of ways to do this.

We need to learn a lot how to do them better. But we need to provide incentives to distribute and acquire information, as well as incentives to do and not to do surgical and medical interventions and diagnostic tests. We need to think about the incentive structure, both on the procedures and on the information side of this question.

Dr. WENNBERG. Senator Rockefeller, your colleague and mine, Dr. C. Everett Koop, Dr. Keller, and I, have suggested that we need to have a focus of responsibility among the professions at the regional level to move beyond the punitive PRO kinds of strategies that have plagued the country in many ways.

We suggest consortia of academic institutions, practicing physicians who are given responsibility for the quality of care in the region. We suggest they be called "Regional Professional Foundations." They have a second function, namely that they integrate life time learning for the professions into the problem of managing the work force for quality, these two things being so overlapping.

Some of our problems are that providers simply do not know anything about epidemiology or statistics because they have no training. They cannot understand to use these tools to evaluate the literature because they were never taught that in medical school.

There is this whole new field of outcomes research, effectiveness research coming out, which is foreign language to many practitioners.

One solution can be sabbaticals back to learn. We have such excess capacity in the supply of specialists from my point of view that we can easily allocate some away from doing procedures to learn new skills. I hope this committee will look very carefully at the infrastructure for quality, and the infrastructure for supporting the profession over its lifetime, so it remains up-to-date, and can modify its skills as technology and societal demands change.

So I hope you will look carefully at that suggestion that Chick has made. It is a good one.

Dr. CHASSIN. Let me try and tackle that also. I have had now a couple of different opportunities to look at the health care community from an oversight vantage point. I think it is important to distinguish three different kinds of circumstances that one needs to address in order to induce change.

There are many individuals and institutions—physicians, hospitals, medical groups—who are exemplary and who need the right tools, who need to be freed from burdensome and intrusive regulation and they will do a good job. And it is our job, I think, to provide those tools and to allow them to be exemplary. It is a small number on a percentage basis.

There is also a percentage on the other side of the distribution that are dangers to the public health. They need to be dealt with by the most aggressively punitive regulatory enforcement that we have. We do not do a very good job on either side. Nor do we do a good job with the majority in the middle who require, I think, neither approach, but rather a mix of incentives.

As I suggested, we need to look very directly at the way reimbursement can be used to reward excellence. That is one way of getting their attention. Publishing data on performance is another way of getting their attention. Providing very specific technical assistance can also help. One of the problems is inertia. We do not really know how to look behind what we are doing. We have never been taught to do it. We do not have the wherewithal or the real enthusiasm to do it ourselves. How do we do it? What is step one? What is step two? What is step three?

That kind of technical assistance, literally how-to manuals, can be very important. We started to develop that years ago in the old PSRO Program. We have gotten very far away from that in the PRO Program. So it seems to me there are three different sets of circumstances.

For most of them we require a significant investment. I think that investment is most properly—we are starting to do it in New York, a few other States are starting to do it—but it is most properly done at a national level.

A physician, for example, who is faced with a patient who has read Reader's Digest and has read the American Cancer Society's recommendation that a prostate specific antigen test should be done on every man over 50 coming to his internist when he is 54 and saying, gee, you did not order that for me last year. How come?

Well, the internist could say with complete truth, there is not a shred of data that using this test as a screening device improves health. So I do not use it. But the patient has Reader's Digest and

wants to know why if this prestigious organization has recommended this test, the internist is not doing it.

That kind of pressure is very difficult to resist. Now, it would help considerably, if the physician were to say, well, but our society, our state, this agency has done a complete review of this and has found, in this circumstance it is not useful. In fact, it might even lead to harm because the most common reason for a positive test is going to be a false positive in your circumstance. That is why I am not doing it.

That is the kind of counter balance that I think we need. It is only investment in the development of guidelines, the development of criteria, of quality improvement, of technical assistance that we need to see.

Senator ROCKEFELLER. Senator Chafee?

Thank you, Dr. Chassin. Thank you.

Senator CHAFEE. Thank you, Mr. Chairman.

One of the problems that seems to be in here—I have some anecdotal information, so you can contradict me—is that patients frequently demand something, that the physician following your suggestion does not provide it, says it is not useful at all, and then the patient can subsequently sue.

Example: A woman with breast cancer comes in, requests that she be given a bone marrow transplant. As I understand it, there is no scientific evidence that a bone marrow transplant will do anything about breast cancer. But I understand this particular case that the physician refused. It was in an HMO. The woman subsequently died of breast cancer. Her family sued and won.

Now that has a chilling effect on the HMO. And, you know, should she have won? Well, I suppose her poor family and the HMO had deep pockets and so forth. It seems to me that is a very disturbing fact in here.

In other instances you find that the person demands it, the patient demands it. If you have a bad knee, the doctor says, well, take some exercise. No, I do not want exercise. He says, I want to be able to run next spring in the marathon. I have run every other marathon. My knee is bothering me. I want some orthoscopic surgery. I do not want to go through these exercises. What is the doctor to do?

I must say the statistics, the difference between New Haven and Boston in various procedures, Dr. Wennberg, were astonishing. But I do not know the solution. What do you do? I think the publicizing—it seems to me if I were in New Haven and I read that the chances for cardiac bypass surgery were twice—as I understand it, the chances for surgery were twice as high in New Haven than in Boston, I might go up to Boston and see what gives, whether I really need this surgery.

So I would think the publicity would be a useful factor. That I take it is what you are recommending. Is that right?

Dr. WENNBERG. Actually, the publicity has been around for quite awhile and it still stays the same. Maybe we should try some statistics for Rhode Island and maybe we would get some better response down there. But it is amazing how difficult it is to translate such an inconsistency at the intellectual level into behavioral changes.

Before you came in I mentioned that what these differences in rates represent are essentially two problems and they are of equal importance. One problem is that the scientific basis for asserting that one treatment has a benefit over the other is weak enough in many cases to allow lots of individual interpretations among the professions.

So it makes it easy for a person trained in bypass surgery to advocate a bypass surgical solution and someone trained in angioplasty to advocate an angioplasty solution. And it used to be there were people advocating drug solutions. But they kind of got pushed aside and it is not happening all that much in this country right now. Although in England that is the central tendency, just to give you an example.

There is a second set of problems. Oh, just before I go into the second set of problems, we can see clearly that the profiles in surgery represent that unique confluence of a particular aggregate supply of surgeons recommending that and the preferences they have for a particular procedure.

Senator CHAFEE. Let me see if I understand. In other words, if you have a lot of surgeons who specialize in heart bypass surgery you will have a lot of bypass surgery done.

Dr. WENNBERG. Yes, that is what is happening right now and that explains it. It does not explain all the variations because some surgeons do a lot more surgery than others do. But generally speaking there is a definite relationship between the per capita supply of neurosurgeons and the per capita supply of back surgery, if you wish, and bypass surgery, bypass surgeons.

So that is one problem. But the other problem is that under the current system of informing patients——

Senator ROCKEFELLER. And we see that enormously now in neonatology, do you not?

Dr. WENNBERG. Yes, it is everywhere. It is just a fundamental fact and behavior in the system. That is something that is terribly important to understand. That is strictly a supply side problem based on manpower training and ultimately how it gets dispersed.

But the other side of the problem, and here is where I think the solutions may rest, is that it is not just a matter of the recommendation, it is the fact that the doctor now prescribes the treatment and the patient accepts it.

When you begin to upset the supply and utilization, disequilibrium with information, in an environment where the patient can actually choose actively, then we see some radical changes in procedure rates.

I mentioned in my testimony the differences that occurred in prostate surgery once patients were informed in an environment where decisionmaking could be shared. In other words, the physicians were committed to it, and the patients were committed to it. And in that environment, the surgical rates dropped 50 percent as preferences went from the supplier side to the patient side in terms of how that decision would be made.

I believe that in that kind of behavior comes a solution to the variation problem. Your job is to figure out how to structure the work place so this can happen.

Senator CHAFEE. Do you think that from your experience, Dr. Chassin's experience, and Dr. Phelps', too, that there is a greater chance for such patients to come to an HMO where their objective is not to spend a lot of money? but their objective is to keep the person healthy so they do not come around too much; and, two, to keep down the cost.

Is there any difference between the HMO and a fee-for-service set-up?

Dr. CHASSIN. Well, we know a little bit about that. We do not know as much as we need to. We certainly know, for example, that HMO care is much less costly, predominantly because of a decreased tendency to use hospital care.

We also know that costs rise at about the same rate in the HMO sector as in the fee-for-service sector.

Senator CHAFEE. From a much lower base though.

Dr. CHASSIN. From a lower base. Absolutely. But the same process of adding new tests and treatments in occurs at about the same rate, but at a lower base. So it does not solve the rising cost problem.

We know there is inappropriate care in HMO's. It is not just a problem, I think, of lack of information. It is a problem also of enthusiasm. It is not just that physicians do not know, cannot choose between two or three strategies. But some physicians are enthusiasts for a particular choice and often are very persuasive.

So if you look, for example, at some of the work we have done in looking at variations, it is not the total number of physicians who might be capable of doing a procedure that is related to the rate of use of a procedure, but rather the number that actually do it, that are enthusiasts about it.

And if there are more enthusiasts, there will be more use of procedures. So it is not just information alone. That has to be girded with other incentives to dampen the enthusiasm.

The example that you raised of drug detail menace is a good example. They are not just purveyors of information. They are enthusiasts for their companies' products and that needs to be dampened if we are going to achieve change by some of the mechanisms I suggested.

Dr. PHELPS. The lack of scientific basis for many of these modern medical decisions hits the HMO just as much as the fee-for-service sector. Even in a fairly large HMO, there are probably only a few neurosurgeons that are actively busy in that group.

Their opinions about what are the best things to do, the best ways to do back surgery, are likely to predominate that HMO's decisions. And if they have a policy guidance manual, they are likely to predominate that policy guide.

So there may be a few differences, but there are still a lot of things going on in the HMO that can be made better by better information.

Senator CHAFEE. So what is your solution then?

Senator ROCKEFELLER. John, can I ask——

Senator CHAFEE. Who knows how to solve this problem?

Dr. WENNBERG. I want to make one point about the HMO's, if I may.

Senator CHAFEE. We have a vote now.

Dr. WENNBERG. Well, let me tell you the most important thing about the HMO from the perspective of the shared decision model, was that the salary structure allowed the physician to uncouple income from his advice giving. Unlike fee-for-service physicians, he did not mind if the rates dropped 50 percent, as they did, because the income to the institution and the income to the physician were not influenced.

So that the HMO provides a good environment for implementing this strategy because the rates can change without the system loosing income.

Senator CHAFEE. I need to have somebody explain that to me.

Senator ROCKEFELLER. Well, let me end on this one. You bring up a very important point. In the health care debate, everybody agrees, Republicans and Democrats, John and Jay, when we talk about that we have to make everybody more conscious of preventive health care. One of the ways you do that—and everything is glory and light when one talks about this. You get somebody to go to an HMO because they are paid for keeping you well, not paid for curing you when you have a disease or if they do it comes out of their salary and their profit margin, et cetera.

People say, well, that is a good thing because preventive care is important. Americans have got to learn how to be more sensible.

Then you introduce something called an HMO into health care reform and all of a sudden these thunderbolts come out of the sky saying, ah-ha, what they are going to do because this is not going to be an all fee-for-service system, they are going to dump—see the word "dump" has an enormous pejorative, emotional content. They are going to dump. They are going to force people to be dumped into HMO's in order to save money.

So that in other words suddenly the argument about preventive prevention and people making choices about staying well as opposed to getting sick and I do not think I have heard any of you necessarily to bulk an HMO's ability to perform surgery or to make critical medical choices wisely. All of a sudden that becomes a rather bad thing.

I mean, after all you get dumped into an HMO. What worse thing can the government do to you? Do you gentlemen buy that? This is something that John Chafee and I are anxiously awaiting your answer.

Dr. CHASSIN. A couple of points. I think there is ample evidence that HMO's can provide high quality care, just as high quality care as fee-for-service. There are, however, some other problems.

One of which is the notion that by using preventive care and primary care we actually save money. I think it turns out demonstrably that we do not. Use of preventive care and primary care is good quality because it prevents disease from occurring later and it enhances health.

But in almost every instance with the exception, I think, of only two services—immunizations and prenatal care—it ends up costing us money. When we use screening mammography, it costs us money to save a woman's life and detect breast cancer early.

When we screen for cholesterol and hypertension, it costs us money to save lives. So I think the idea that prevention and primary care will save us money is misplaced. And, in fact, we should

not burden the legitimate good quality argument for doing prevention in primary care with the requirement that it demonstrate that it save us money. We will be disappointed.

Now, the other problem though that I think is legitimate in talking about——

Senator CHAFEE. Those two exceptions, one being prenatal care and what was the other one?

Dr. CHASSIN. Immunization, childhood immunization, and possibly pneumococcal immunization for elderly, though the data, I think, are questionable.

But I think it is legitimate to worry because the fee-for-service method of reimbursement has been so effective at producing overuse and inappropriate care it is legitimate to worry that turning that incentive on its head, that is providing a direct incentive to produce too few services, could cause the same magnitude of problem with underuse that we have now with overuse.

Now when you couple that worry—and as I say, I think it is legitimate—we do not really have a good demonstration of the magnitude of that problem today because we have very few HMO's today, frankly. When you combine that worry with the knowledge that we have large underserved populations who do not need incentives to use fewer services, but need education, outreach and aggressive provision of principally primary care, but also the care that follows when you do have primary care and find disease, that raises more concerns about whether this is the appropriate vehicle to put all of our passengers in as we ride to health reform.

Senator ROCKEFELLER. Dr. Wennberg, do you want to comment?

Dr. WENNBERG. I want to say I agree with Dr. Chassin's remarks. I will say that—maybe I can help with this confusion that I left you with on this other issue—the HMO's come in several varieties. There is one I call the classic HMO, which is the prepaid group practice which has a defined population and the physicians essentially are on salary for that organization, has some remarkably good features for allowing quality to be built into the system.

Number one, the workforce in that firm can allocate itself according to problems. It is not constrained by the fees. So if the surgery rate once you ask patients what they want drops 50 percent, it can be allowed to drop 50 percent without having to make the firm go broke, which is what happens in fee-for-service systems.

Secondly, the HMO in its mature form can take a long term interest, both in the population that it is serving because they will be around for awhile——

Senator CHAFEE. Doctor, this has got to be a very swift answer. We are now in our final 7½ minutes, about 5 minutes for us to get over there and vote.

Dr. WENNBERG. All right. Secondly, they can take care of the professional profession over its lifetime. Lifetime learning is built into it.

Senator ROCKEFELLER. Do you not think that one of the phenomenon of the next 10 years is going to be an HMO which avoids the pitfall that you are talking about? Because after all, outcomes are going to be much more known by the public. In other words, the incentive to provide less care need not be the HMO of the future, indeed, even if there are some HMO's like that today.

I mean it is possible to do an HMO with emphasis on prevention. Forty-five percent of physicians are on salary. That is an enormous statement already. That means a lot of people have made the decision that for whatever reason they are going to take a more predictable lifestyle, that they are going to be less worried about——

Dr. CHASSIN. Oh, I think that it is possible. I would urge a little caution for two reasons. One, we have some experience with the diffusion of HMO technology in California in the early 1970s in the Medicaid program which was a disaster because there was no oversight.

Large numbers of organizations sprang up to provide capitated care to that population, but did not provide service. And as soon as that became clear over a period of years, it was a disaster.

The second problem is that monitoring for underuse is much more difficult than monitoring for overuse. If you are trying to find underuse, you are looking for events that should have happened but did not. There is precious little documentation of that.

Overuse is much easier. Technical quality, patient satisfaction is much easier to develop systems for. We have practically no systems, even theoretically, to even monitor well for underuse. That is going to take awhile to develop.

Senator ROCKEFELLER. Senator Chafee is right. We have to go. I have about 25 questions here I did not even get close to. You have been a wonderful panel. This is the kind of thing that should go on for days. This should be required by law to be on C-SPAN. I thank you all very, very much.

Dr. CHASSIN. Thank you.

Dr. WENNBERG. Thank you.

Dr. PHELPS. Thank you.

[Whereupon, at 12:04 p.m., the hearing in the above-entitled matter was adjourned.]

APPENDIX

ADDITIONAL MATERIAL SUBMITTED

PREPARED STATEMENT OF MARK R. CHASSIN

The elimination of inappropriate care must play a central role in health reform. Indeed, I believe it is the key to controlling costs safely, without jeopardizing health. It is clear that substantial savings will be required from funds currently spent on health care in order to provide universal access to care for all Americans. Such savings can be realized by curtailing inappropriate care.

Research indicates that a substantial amount of health care is provided for clinically inappropriate reasons in the United States. Whether in the use of medications, days spent in hospitals, diagnostic tests, or surgical procedures, inappropriate care has been found throughout the American health care system.[1 2 3]

The magnitude of the problem is somewhat difficult to assess, because studies have not evaluated representative populations or health services. In its study of quality of care in the Medicare program, the Institute of Medicine concluded that "evidence of overuse of health care services is substantial."[4] My own reading of this literature and my experience in various positions of oversight in the health care system suggests that we could safely eliminate about 20% of what we do in health care, and quality would actually improve because patients would be spared the risk that attends to inappropriate care.

It is important to be clear about what constitutes inappropriate care. I do not include in this category the clinical decisions that physicians and families agonize over because many courses of action are available and reasonable physicians may reasonably disagree over which one to select. In these circumstances risks and benefits are about equal, and personal patient preferences about the relative merits of one or another kind of risk or benefit are crucial to decisionmaking. Inappropriate care refers to circumstances in which health care services are provided when their risk exceeds their benefit. There are at least two different kinds of inappropriate care that are useful to distinguish. The first involves the overuse of commonly provided health services.

The original RAND studies that assessed inappropriateness in 1981 found that 14% of coronary artery bypass graft surgeries in one western state were performed for inappropriate reasons, and among Medicare beneficiaries, 17% inappropriate upper gastrointestinal endoscopies and coronary angiographies and 32% inappropriate carotid endarterectomies.[5 6] A similar study assessing care in 1987 showed 20% inappropriate coronary angiographies in the Medicare population,[7] and a recent RAND study found 16% inappropriate hysterectomies among selected HMOs.[8]

The methods that were applied in these studies allowed us to distinguish between care provided when risks exceeded benefits from circumstances when risks and benefits were about equal. We called the latter circumstances "equivocal" indications for providing services. The proportion of cases in which care was provided for equivocal reasons ranged from 9–32%.

Distinguishing inappropriate from equivocal reasons for providing health care is important for more than theoretical or research purposes. I would argue strongly that we must seek vigorously to eliminate inappropriate care before attempting schemes to ration effective care or to intervene in the difficult circumstances that often characterize choices about receiving care for equivocal indications.

Another kind of inappropriate care occurs when new tests and treatments enter the health care system. Many analyses suggest that the principal factor that explains why health care costs have been rising faster than costs in the rest of the economy for a half-century is all the new tests and treatments that we invent.[9] We are especially adept at creating these innovations. We are notably deficient in evalu-

ating when they produce good outcomes for patients and when they don't. We are even less effective at making sure they are used in the former circumstances and avoided in the latter.

This is not a problem only of medical technology, of big machines like MRI scanners and lithotripters. It is also a problem of new antibiotics and tranquilizers, of new blood tests like the prostate specific antigen test, and of new diagnostic and therapeutic procedures. Nearly every new test or treatment in recent memory has followed the same pattern of diffusion. The innovation is proved effective for a narrowly defined group of patients but when it gets out into the health care system it is used for much broader groups of patients in circumstances where benefit has not been established.

The Prostate Specific Antigen (PSA) test is just the latest incarnation of this problem. By one estimate, the cost of implementing the American Cancer Society's recommendation that every man over 50 have an annual PSA test is $28 billion, taking into account the further diagnostic and therapeutic interventions that would ensue.[10] While the use of the test is well-accepted as an important adjunct in managing patients with known prostate cancer,[11] no data document its effectiveness as a screening test.[12]

What is the harm, one might ask? Even if a patient with a headache has one chance in a thousand of having prostate cancer, demonstrating with a PSA test that cancer is not present is a tangible benefit. The problem is that everything we do in medicine has risk. For diagnostic tests, the principal risk is not immediate death or injury, it is the risk of a falsely positive finding and the further assessment that must follow it.

An example will serve to illustrate the potential magnitude of this problem. Let's say we have a diagnostic test that is 95% sensitive. That means 95% of people with the condition we seek to diagnose will test positive. Let's say the test is also 95% specific, meaning that only 5% of people without the condition will test positive. (The PSA test is much less efficient.) If we use the test in circumstances when the frequency of the condition we seek is 1%, then fully 84% of all positive tests will be falsely positive, leading to a cycle of more testing and more risk. If the true frequency of the condition we are chasing is more like 1 in 1000, as with the hypothetical example given above, the frequency of falsely positive tests rises to an astonishing 98%.

The pattern of incomplete evaluation that emerges from even the most cursory review of diagnostic tests such as CT scans, MRS, and PSA appears also in the assessment of new therapeutic advances such as percutaneous transluminal coronary angioplasty (PTCA), lithotripsy, laparoscopic surgery, and cancer chemotherapy, to mention only a few among many examples.[13] This kind of inappropriate care compounds the first. Today's commonly done, inappropriately provided health care service was yesterday's poorly assessed innovation. The less well we evaluate today's innovations, the more inappropriate use we build into the health care system in years to come.

To eliminate inappropriate care, we must understand what causes it. Several causes are apparent. Fee for service reimbursement plays an important role but is not the whole story. I believe that the enthusiasm with which physicians and other interested parties advocate a particular treatment or procedure plays a significant part. As new tests and treatments enter the health care system, enthusiasts promote their use in situations where benefit has not been demontrated but where some clinical theory suggests it might be present. These physicians are often supported by companies with an economic stake in the adoption of the innovation. Gradually, these enthusiasts often convince their colleagues and a new indication for the innovation is accepted, without any data documenting its efficacy. The data from the RAND appropriateness studies are consistent with this "enthusiasm hypothesis."[14]

This process is abetted by the way in which patients are referred from primary care physician to specialist in the fee for service sector. For fear of losing a source of referrals, a specialist may be reluctant to tell a primary care physician that the procedure he requested is unneeded. And primary care physicians too often abdicate decisionmaking responsibility when a referral is made. Fear of liability plays a role in this process, but I believe it is a small one.

Other factors outside the control of physicians are also important. As a society, we are infatuated with technology and innovation. The media often write uncritical stories about new tests and treatments that wildly exaggerate (or ignore) what is known about their effectiveness. Patients bring these expectations to their physicians and want to receive the latest new drug or procedure. Physicians often find such demands difficult to resist.

How then might inappropriate care be eliminated? Because its causes are multiple and reinforcing, it will not be successfully dealt with by relying on any one approach. Altering physician reimbursement by changing from fee for service to capitation or salary will not be successful by itself. HMOs are prone to inappropriate care, because their physicians are prone to the same enthusiasm as other physicians. IPAs are often susceptible to the same referral problems as fee for service medicine.

We must use many tools. For some services, limiting the capacity of the health care system may work. I believe that this mechanism has allowed New York State to contain the inappropriate use of invasive cardiac procedures. Through a strict certificate of need process, the number of hospitals permitted to offer these services is very limited. Only 31 hospitals in New York perform open heart surgery, compared with about four times that number in California. As a direct result of this planning process, New York experiences very little inappropriate care for these procedures. Recent RAND studies showed that in 1990 New York State exhibited 2.4% inappropriate coronary bypass surgeries, 4% inappropriate coronary angioplasties, and 4% inappropriate coronary angiographies.[15] [16] [17] These are the lowest figures for inappropriate care ever documented by this method.

Not all services can be regulated by this approach. Another important approach to reducing inappropriate care relies on the fact that the inappropriate use of health care services is a serious quality problem, because patients are exposed to the risk of unneeded services without the promise of commensurate benefit. Focusing the quality improvement activities of local hospitals and physicians on problems of inappropriate care can have positive results. One recent study documented dramatic reductions in inappropriate carotid endarterectomies following a voluntary quality improvement program, based on the development of practice guidelines, at a community hospital.[18]

I doubt that voluntary programs of quality improvement will be sufficiently effective to eliminate inappropriate care by themselves. They should be encouraged and rewarded, but they are unlikely to proliferate without additional incentives. Even the institutions most ardently committed to continuous quality improvement today rarely address major problems of inappropriate care.

Other incentives should be used, such as those proposed in health reform legislation submitted last spring by Governor Cuomo in New York. We proposed that reimbursement be linked to performance on quality measures. At present, reimbursement is largely neutral with respect to quality. We believe that it is possible to construct reimbursement incentives that reward excellence. One way excellence is measured is in very little inappropriate care. Also, regulation such as utilization review might be eliminated if performance indicates low levels of inappropriate care.

In addition, experience in New York State with the public release of information about quality suggests strongly that, if the data are reliable and valid measures of quality, physicians and hospitals will respond by energetically pursuing local quality improvement activities. We have witnessed a 36% decline in risk-adjusted operative mortality from coronary artery bypass surgery in the first three years of operation of just such a program in New York State.[19]

Finally, a program of public education is needed to dispel the widespread belief that more health care is always better health care. We need to educate the public about the risks of health care services. We need to work with the media, both print and electronic, so they understand the adverse effects of exaggerated claims of effectiveness of the latest advances in medical science.

Governor Cuomo proposed a wide-ranging Quality Initiative in his Special Message to the Legislature on Reforming the Health Care System earlier this year. This Message was followed by a detailed legislative proposal that described the Initiative. This program calls for a partnership between government and all individuals and institutions in the health care delivery system. We propose to use all of the available quality improvement tools, including those described here, to reduce inappropriate care, as well as to ensure that needed care is provided with a high degree of technical skill.

The combination of all of these approaches can be effective in curtailing inappropriate care. Any one alone is likely to fail. If we fail, we will have little choice but to resort to means of containing costs that will have deleterious effects on health. Rationing of effective health care services, drastic restrictions on hospital or physician payments, and substantial increases in consumer cost sharing are all likely to have this effect. The effort to eliminate inappropriate care must come first.

REFERENCES

1. Brook RH, Kamberg CJ, Mayer-Oakes A, Beers M, Raube K, and Steiner A. Appropriateness of acute medical care for the elderly: An analysis of the literature.

The RAND Corporation, Santa Monica CA, (RAND/R-3717-AARP/HF/RWJ/RC), September 1989.

2. Steill IG, Greenbery GH, McKnight RD, et al. Decision rules for the use of radiography in acute ankle injuries: Refinement and prospective validation. JAMA 1993; 269:1127-1132.

3. Greenspan, AM, Kay HR, Berger BC, et al. Incidence of Unwarranted Implantation of Permanent Cardiac Pacemakers in a Large Medical Population. N Engl J Med 1988; 318:15863.

4. Lohr KN (ed). Medicare: A strategy for quality assurance. Institute of Medicine, Washington D.C., 1990, volume 1, p. 230.

5. Winslow CM, Kosecoff JB, Chassin M, Kanouse DE, Brook RH. The appropriateness of performing coronary artery bypass surgery. JAMA 1988; 260:505-509.

6. Chassin MR, Kosecoff J, Park RE, Winslow C, Kahn K, Merrick N, Keesey J, Fink A, Solomon DH, Brook RH. Does Inappropriate Use Explain Geographic Variations in the Use of Health Services?: A Study of Three Procedures. Journal of the American Medical Association 1987; 253(18):2533-2537.

7. United States General Accounting Office. Medicare: Improvements Needed in the Identification of Inappropriate Hospital Care. Washington D.C.; December 1989; GAO/PEMD-90-7.

8. Bernstein SJ, McGlynn EA, Siu AL, et al. The appropriateness of hysterectomy: A comparison of care in seven health plans. JAMA 1993; 269:2398-2402.

9. Newhouse J. An iconoclastic view of health cost containment. Health Affairs 1993 (supplement); 12:152-171.

10. Burton TM. Blood test for prostate cancer is raising issues of reliability, drug-industry role. Wall Street Journal; February 16, 1993, p. B1.

11. Lange PH. The next era for prostate cancer: Controlled clinical trials. JAMA 1993;269:95-96.

12. Garnick MB. Prostate cancer: Screening, diagnosis, and management. Annals of Internal Medicine 1993; 118:804-818.

13. Grimes DA. Technology follies: The uncritical acceptance of medical innovation. JAMA 1993; 269:3030-3033.

14. Chassin MR. Explaining geographic variations: The enthusiasm hypothesis. Medical Care 1993; 31(5 suppl):YS37-44.

15. Leape LL, Hilborne LH, Park RE, Bernstein SJ, et al. The Appropriateness of Use of Coronary Artery Bypass Graft Surgery in New York State. JAMA 1993; 269:753-760.

16. Hilborne LH, Leape LL, Bernstein SJ, Park RE, et al. The Appropriateness of Use of Percutaneous Transluminal Coronary Angioplasty in New York State. JAMA 1993; 269:761-765.

17. Bernstein SJ, Hilborne LH, Leape LL, Fiske ME, et al. The Appropriateness of Use of Coronary Angiography in New York State. JAMA 1993; 269:766-769.

18. Asaph JW, Janoff K, Wayson K, Kilberg L, and Graham M. Carotid endarterectomy in a community hospital: A change in physicians' practice patterns. Am J Surg 1991; 161:616-618.

19. New York State Department of Health. Coronary artery bypass graft surgery in New York State: 1989-1991. Albany, NY; New York State Department of Health; 1992.

RESPONSE OF DR. CHASSIN TO A QUESTION SUBMITTED BY SENATOR PACKWOOD

Question. Oregon is the first state in the Union to courageously face the issue of whether or not we can afford to provide every known medical procedure to every person who wants such a procedure at taxpayer expense. Oregon decided we could not.

So Oregon put together a medical priority list of 696 treatments for its Medicaid program. At the top of the list are those treatments that are effective in treating or preventing illnesses. At the bottom of the list are treatments that have marginal benefit.

You say in your testimony that inappropriate utilization must be eliminated before rationing should be attempted.

In Oregon, we have already done what you recommend—made a list of treatment priorities according to what we know is effective. Do you support this kind of effort?

Answer. I have one major concern with Oregon's rationing scheme. It does not attack the problem of inappropriate care at all effectively. By judging the merit of individual health services globally, the approach fails to recognize that a large number of the services judged useful are often provided inappropriately in specific clinical circumstances.

Hysterectomy (or myomectomy) for leiomyoma of the uterus (line 479) is a covered service. However, for small leiomyomas no surgical treatment is required. Research has indicated that hysterectomy is often performed inappropriately; one of the more common inappropriate indications is small leiomyomas. This is one of many examples in which the Oregon plan fails to identify inappropriate uses of common health services and provides coverage for them.

On the other hand, many of the services the Oregon plan lists as unworthy of coverage provide clear benefits to patients under the right circumstances. The plan excludes coverage for medical treatment for atopic dermatitis, chronic bronchitis, and allergic rhinitis. While mild versions of these conditions may require minimal if any active intervention, severe cases can be quite disabling.

In a nutshell, my problem with the Oregon plan is that it would permit inappropriate hysterectomies for small leiomyomas but prohibit coverage for treatment of severe chronic bronchitis. I would prefer an approach that focuses clearly on eliminating all the inappropriate uses of health services, which under some specific circumstances are very beneficial but under others actually produce harm. Adopting such an approach would, I believe, permit us to afford to provide necessary and effective care to everyone.

PREPARED STATEMENT OF SENATOR DAVE DURENBERGER

Mr. Chairman, I welcome the opportunity to discuss medical practice patterns and the appropriateness of care with this distinguished panel. If we are to reach our stated goal of universal access to superior quality, cost-effective care through universal coverage of financial risk our health care system must become more productive.

I have believed for a long time that productivity is achieved by changing the practice of medicine—and by that I mean how physicians and other health providers organize themselves and inform themselves to deliver cost-effective services.

Frankly, I have been quite frustrated with the seeming fixation that the present debate seems to have with alliances, HIPCs, and cooperative purchasing arrangements. While buyers are important, the really important institution in the reform discussion is the AHP, or accountable health plan. It is the health *plans,* managed by the providers, that will change the practice of medicine and get costs under control.

I'd like to extend my welcome to Dr. Eddy and Dr. Wennberg. I know both of these gentlemen and am very familiar with their research.

Jack has taught us all a lot about utilization of care. He has made "prostate" a household word. David has helped many practitioners evaluate what works and what doesn't in technology. They are both very clear and cogent thinkers—maybe that is due in part to the fact that both of them spend most of their time in Jackson Hole, Wyoming.

They both practice what they preach—that is that better decision making which leads to more appropriate care comes from evaluation. Finding the appropriate role for government and adequate support for the evaluative sciences has been a major concern of mine, and of this committee. We made a stab at that when George Mitchell and I sponsored a bill to establish the Agency for Health Care Policy and Research (AHCPR) back in 1989.

Our witnesses today speak very strongly about the need to improve resources for the science of evaluation. I agree. In the Managed Competition Act of 1993 (S. 1579), which Senator Breaux and I introduced last week, we have a very detailed set of provisions to improve evaluation. We create an Agency for Clinical Evaluation (ACE) to coordinate the evaluative sciences now scattered among many federal agencies and to link up with the private sector. I would appreciate it if this knowledgeable panel would study those provisions and give us their reactions.

I must temper my enthusiasm slightly having read the testimony submitted today. The witnesses have shown us that there is a lot of inappropriate and unnecessary care in our present system by any measure. However, I take issue with their conclusions that global caps are the vehicle to eliminate waste in the system.

Caps are a political solution that will likely create a lot of mischief. They will lead to market distortions that will defeat the purpose of cost containment. How a cap produces better utilization and more appropriate care simply escapes me. I will submit some questions for the witnesses to pursue this issue further.

PREPARED STATEMENT OF SENATOR CHARLES E. GRASSLEY

I believe that the subject to be considered today, the appropriateness of care, is very important. It is even a critical one, Mr. Chairman.

It is critical because Americans have a right to expect that the health care they receive be of the highest quality.

It is also critical because in a managed competition system the consumer must be able to evaluate the quality of the medical care provided through their health plan.

There appears to be some skepticism as to whether we actually have such measures available. I hope our witnesses today will be able to give us some assurances that appropriate quality of care measures will be available in a timely fashion.

Our citizens also should expect that they will actually receive the health care services that they need.

The incentives in a managed competition plan are liable to encourage underutilization. In fact, if one of the core problems of a fee-for-service, third-party payer system is overutilization, then one of the core problems of a capitated, managed competition system is liable to be underutilization.

This is a point made by two of our earlier witnesses, Karen Davis and Stuart Altman. It is a point made also in the written testimony of Robert Brook, who was unable to be with us today. I understand that that testimony will be included in the record.

I am looking forward to learning whether our witnesses believe that appropriateness of care measures can be designed that will help us track not just the quality of care and the provision of unnecessary care, but also those situations in which care may not be given when it should be.

Thank you, Mr. Chairman. I am looking forward to the testimony.

PREPARED STATEMENT OF SENATOR ORRIN G. HATCH

Mr. Chairman, I am delighted that this hearing is taking place today. With the rapidly rising high cost of health care, I have had a long interest in finding ways to make the health care system more efficient and to improve the appropriateness of care.

The Rand Corporation and others have reported that between 10 and 30 percent of medical treatments are either unnecessary or ineffective. Clearly, we as a nation are spending billions of dollars that are not improving the health of the citizenry.

The federal government's Agency for Health Care Policy and Research also has disclosed in its medical effectiveness research projects a lack of good evidence of effectiveness for far too many medical services and procedures currently in wide use. Because AHCPR is the only agency changed with the broad responsibility for conducting this type of research, I was disappointed when it was not fully funded for FY '94 at the level proposed in the President's budget with increased funding of $30 million as I and other Senators had urged.

I think that spending millions of taxpayer dollars on effectiveness research can lead to savings in the range of billions of dollars. Undoubtedly, this research is a good investment for all of us.

I also welcome the development of clinical practice guidelines. I am certain that the availability of these will have a significant impact on the reduction of costs associated with medical liability cases and the practice of defensive medicine.

I look forward to the testimony of the witnesses today. And I have a special interest in their observations on what should be the respective roles of the government and the private sector in the conduct of additional effectiveness research and the development of clinical practice guidelines.

PREPARED STATEMENT OF CHARLES E. PHELPS

The Administration has proposed changes that could markedly alter our health care system, with the goals of universal insurance coverage, improving the quality of care, and--very high on the agenda—reducing health care costs considerably. These goals together, however, create a confining box from which it may be difficult to escape, because expanded insurance coverage and higher quality inevitably add to costs, despite much wishful thinking to the contrary. However, a major new emphasis on the cost-effectiveness of health care interventions does provide a way out of this box.

Much has been written, and Dr. Chassin's testimony today discusses issues associated with elimination of "inappropriate" care, defined as that for which the medical

risks exceed the medical benefits. Removing such care from the health care system represents a laudable goal, one with which every American would likely agree if given the chance, although there may be some disagreement about the extent of savings that might emerge from such an effort, even if 100% successful.

Today I wish to impress upon you an equally important goal. We must not only eliminate care where the medical risks exceed the benefits, but also care that creates little improvement in patients' well being per dollar of resources expended. We must introduce cost-effectiveness criteria into the policy making process about health care financing, health care organizations, investment in health care technology, and the training decisions we make about providers of health care. It simply makes no sense to introduce a budget process that ultimately limits per-person spending on medical care and then to ignore costs in our considerations of appropriateness. Unfortunately, we currently do not possess the knowledge of how to do this well. Thus, we must make major new investments to produce and disseminate information about how to carry out cost-effective medical care.

The Administration's proposed health care plans ultimately place a cap on the expenditure of resources within the health care sector. This cap grows with the economy, but nevertheless represents a cap. Once such caps are in place, it is *fundamentally necessary* to introduce considerations of cost as well as medical risk into the analysis of what is "appropriate" and "inappropriate" medical care *in order to achieve the best possible health for the people of the United States.* This point is so essential that I will restate it, to be sure that the issue is well-understood: Any approach to implementing the Administration's or any other health care plan that ultimately limits spending on medical care will not achieve the best possible health outcomes for each and every person in the country unless costs as well as benefits are considered. We *must* move to a cost-effective medical care system to avoid harm to the health of the public.

How can we achieve this goal? Newly developed methods provide specific ways of measuring the benefits to people's health using the same "yardstick" no matter what the medical intervention. This "yardstick"—called Quality Adjusted Life Years (QALY, pronounced "kwa-lee")—combines both gains in length of life and in the quality of life. With this "yardstick" in hand, we can—in concept—estimate the cost-effectiveness of almost any medical intervention, measured as added dollars spent per added QALYs obtained. When properly used, this QALY yardstick completely incorporates not only the medical benefits that health care providers can see and measure, but also the value that individuals place upon these outcomes. Dr. Wennberg's talk and written comments emphasize the importance of both medically defined events and their value to people in judging the "best" outcome.

It is easy to prove that—for a given budget, such as proposed by the Clinton plan—the best health outcomes for a defined population emerge if and only if the extent and intensity of use of every intervention in the system is adjusted so that the added benefit from one more "unit" of care, divided by its cost, is the same for every intervention in the health care system. Learning how to do this represents an enormous but immensely valuable research effort.

Let me provide the intuition behind this result. People commonly think about the cost-effectiveness of an intervention as something fixed, cast in concrete like a statue in the park. Thus, people will say that "CABG" surgery costs $30,000 per life year saved, while mammography for woman costs only $10,000 per life year saved. Alas, neither statement is correct in any meaningful sense. CABG surgery is incredibly cost effective, under $8000 per QALY, when performed on high-risk patients with substantial occlusion of the left main coronary artery. However, these are not the only patients receiving CABG surgery. CABG operations on individuals with single vessel disease and only moderate angina (a group commonly receiving the intervention) costs $65,000 per QALY.

Similarly, mammography for women clearly produces health benefits, but not for free. Annual breast examination for women 55–64 years old has a cost per QALY of about $11,000; adding mammography to the same group brings the cost per QALY above $30,000. Screening 45–54 year old women further increases the cost per QALY. Increasing the frequency of screening to every 6 months would add greater costs per QALY, and reducing the frequency to every 2 years would reduce the cost per QALY.

Another example in the realm of heart disease provides an even more striking example. A cholesterol-reducing drug (lovastatin, in low doses) creates improved health outcomes at only $1,700 per QALY for relatively older, male heart attack survivors. For younger female hypertensive non-smokers (a group with mild risk of heart attack), the cost per QALY rises to over $700,000. For the same female non-smokers who are not hypertensive, the cost per QALY reaches an astronomical $1,500,000 per QALY.

Now let me return to the basic idea. Suppose we have a fixed budget for health care (as the health alliances proposed by the Administration will lead to), and that mammography and CABG surgery are the only two medical interventions. (The idea is equally valid when you consider many interventions.) Suppose that current practice only uses breast examination as a screening tool for women over age 55, and that CABG surgery is done on patients with a wide variety of indications, including those with left-main artery disease and those with moderate angina. These interventions will each produce some additional quality adjusted life years, each at different costs per QALY. *All of these interventions* would pass tests of "appropriateness" as defined medically (benefits exceed risks). Now consider what happens if we eliminate one CABG surgery on patients with moderate angina, costing (say) $20,000, and producing (on average) 0.3 quality adjusted life years of added benefit. Take this $20,000 and invest in mammograms for women, costing (say) $200 each. (The price data are fictitious to make the illustration easy; the same result holds true when using real prices or costs.) We can pay for 100 mammograms with the $20,000, and they will create a total of about .7 QALYS of improvement in health, on average, for those 100 women. (In this case, most will not benefit, and a few will receive considerable extensions in their life from early detection of breast cancer.) Thus, shifting resources from CABG to mammography in this example improves the overall health of the public.

This is not just an issue of acute therapy vs. prevention. The same idea works equally well if we started with the premise that all women had been receiving mammography as well as breast examination, but that CABG was performed only on persons with left main artery disease. In that case, since the incremental cost per QALY for CABG is only $6,500, but the cost per QALY for mammography plus breast examination screening is $30,000 per QALY, the proper allocation of resources would be to shift *away* from mammography on the lowest risk women and begin to do some additional CABG surgery, perhaps on a group at lower risk than those with major left-main artery occlusion, but still not descending to the "moderate angina" group. In either case, the health of the public is improved by shifting away from the procedures with the highest cost per QALY towards those with the lowest cost per QALY. The problems come from trying to identify which is which, and finding ways to change the ways we use resources in the health care system at present.

These examples given earlier only hint at a pervasive phenomenon: The incremental cost per QALY varies hugely across medical interventions used in the U.S., both across interventions and—equally importantly—depending on how these interventions are used, for whom, how often, and for varying indications. A few interventions actually save money as well as improving health. However, most interventions, including most "preventive care," add costs while adding QALYs. Some medical care interventions add QALYS for several hundred or several thousand dollars each. However, as my previous examples hinted, many others add QALYs at costs of hundreds of thousands and sometimes millions of dollars per QALY. The ways in which we use these types of interventions will and should come under increasing scrutiny in this new cost-conscious environment.

What does this mean for managed competition and other forms of cost-conscious health care? For almost every medical intervention imaginable, there exists *some* set of patients for whom the intervention produces health improvements at relatively low cost per QALY. As the intervention spreads to a wider group, the "bang for the buck" invariably declines. We do not need to mandate the use of such criteria in a managed competition setting, because every health plan in such an environment, competing on both costs and quality of care, will automatically seek to use their health care budget in ways similar to the mechanism I described earlier, shifting away from the most costly towards the least costly interventions, in terms of their cost per QALY standing. Unfortunately, we do not currently have a good understanding of the costs per QALY for many interventions (and groups receiving those interventions). Thus, the limiting feature of managed competition appears to be a new form of knowledge that our society does not in general possess: an understanding of the cost effectiveness of medical interventions as they are used for different populations, and with different intensity and frequency.

Introducing cost into the "appropriateness" calculation requires a considerable investment in new information. Our country made this investment in the basic science of medicine, beginning in the 1950s and 1960s, through the creation and expansion of the National Institutes of Health. The research flowing from these investments has created the ability to cure disease and improve well being in ways that we never could have imagined in the 1950s, including diagnostic capabilities from CT and MRI scanners, genetic tests for rare and common diseases, and soon even the ability to use gene-therapy to cure even genetically encoded disorders. We continue to in-

vest in the production of basic biomedical knowledge at a high rate, with the NIH budget for this year of approximately $9 billion. By contrast, the investment in knowledge about *how to use the health care interventions that we now have,* even after a large increase this past year, is not quite $150 million, about $1 for each $60 spent at the NIH.

Studying the relationship between medical interventions (both diagnosis and treatment) and subsequent outcomes provides the best way to determine the cost effectiveness of various types of medical care, including an understanding of how cost-effectiveness ratios change when one changes the intensity of treatment, the frequency of screening, or the number and types of people receiving these interventions. Studies to learn this—just as basic biomedical research—require funding. Immediately, more funding is needed to carry out the most important studies. Also of urgent need now are greatly increased funds for training programs for people to carry out this research program. At the Agency for Health Care Policy and Research, research in the Medical Treatment Effectiveness Program (MEDTEP) provides vital new information on just these issues. Additional research in general health services research studies other aspects of the health care delivery system. Both are needed to improve the effectiveness of health care delivery in the United States.

Although funding for studies of this type has increased importantly in the past year, the potential gains from further investments still exceed the costs of carrying out such studies greatly. In research I conducted for the Institute of Medicine (Phelps and Parente, 1990, Phelps and Mooney 1992), I have shown that the returns from studying the proper ways of using existing medical interventions exceed the costs of carrying out those studies by factors of tens, hundreds, and in a few cases, even thousands. This research, extended by the Institute of Medicine last year (Sox and Donaldson, 1992) has developed a method to prioritize such research, so that the returns to investment in knowledge are as large as we can plan for. The overall message is quite clear: we have invested greatly in basic biomedical knowledge, but now we need an important new investment to learn how to use the medical interventions that currently exist.

Without this investment in new knowledge, we cannot achieve a truly cost effective and therefore humane system of health care and health insurance. No matter what the mechanism for universal insurance—employer mandate, individual mandate, or any other system—and no matter what the mechanisms of cost control—managed competition, single payer, or other systems—none of the relevant "actors" in the health care system now possess the knowledge they need to deliver health services in the cost-effective way I have discussed. The investment in new knowledge of "how, for whom, and when" will pay immense dividends no matter what the system of financing, or no matter what mechanisms of cost control emerge.

You must also understand that no single private entity, no health alliance, no health plan, no HMO, no hospital, no doctor's group, and certainly no single doctor, has sufficient incentive to provide the full amount of research in this area that our country should choose. The federal government has a clear and crucial role here in the production and dissemination of information. The NIH has done this before for information about the biomedical sciences, and the logic of investment in research there stands unchallenged. We are now on the brink of an opportunity to make a similarly important investment in a new form of knowledge, allowing us to build a cost effective health care system that offers the *only* way to provide the best possible health outcomes to our population.

Even after this research program is in full blossom, we need to find ways to get doctors and their patients to seek out and use this information. This requires much new thinking about ways to disseminate this information. Dr. Wennberg and his colleagues in the Maine Medical Assessment Program have provided one vision about how to accomplish this. Other approaches include "academic detailing," mimicking the efforts of drug "detail" men who live in the interstices of the health care system, transmitting knowledge about drugs to doctors, on behalf of the manufacturers of those drugs. Much more is needed.

We also need to consider incentives and mechanisms to increase the use of information about the best uses of medical interventions. These may include payment mechanisms, hospital or insurance-plan based educational efforts, licensure examination modifications, or many other areas. Put simply, we have little knowledge about how to change doctors' and patients' behavior, even when we have the "right" answer. This represents another important area of research.

In summary, the *only* way to achieve the Clinton Administration's and others' goals of universal access to health care, reduced costs, and improved quality of care is to strengthen greatly the role of cost-effectiveness analysis in the organization and delivery of health care. This will require both new research (to characterize the

cost/QALY for various medical interventions and groups receiving them) *and* a major rethinking of how we decide on the benefits our health insurance provides, the resources we devote to health care, and the ways we use those resources.

PREPARED STATEMENT OF JOHN E. WENNBERG

Thank you for the opportunity to testify on the problem of practice variations and the question of unnecessary care. The basic facts are simple; the amounts and kinds of medical care patients consume depend more on where they live than on the diseases they have. In health care, geography is destiny.

VARIATIONS IN THE RATES OF TREATMENT

Virtually every medical condition can be treated in more than one way. For many conditions, there are medical as well as surgical treatments that are appropriate. Watchful waiting—living with symptoms in order to avoid the risks of more invasive treatment—is often a reasonable alternative.

Physicians have different opinions about the outcomes and different preferences for the risks and benefits for these treatments. In a given community, the per capita numbers and specialty distribution of local physicians as well as individual physicians' own predilections affect the chances for undergoing a particular treatment. This uncertainty about the best choice of treatment, and the tendency of physicians to choose treatments according to their own preferences, rather than those of patients, has created a health care economy driven by supplier-induced demand.

Nine conditions (see Table One) account for well over half of the major surgery done in the United States. For each condition, there are other, non-surgical, options. The rates at which these various treatments are performed vary substantially from one community to another.

In some parts of Maine, the rates for prostate surgery in the 1980s were so high that we predicted that more than half of men would have prostate surgery for benign prostate disease (BPH) by the time they reached age 85, while in other parts of Maine, where most cases of BPH were managed with watchful waiting, the chance for surgery was only 15 percent. Communities served by the nation's most prestigious academic medical centers are not immune from the supplier-induced demand that follows from uncertainty about outcomes and entanglement of physicians' preferences for treatment with those of their patients. Residents of New Haven, for example, have twice the risk for cardiac bypass surgery as do Bostonians (whose clinicians favor non-surgical interventions more often); New Haven women have about twice the risk for hysterectomy as women in Boston; but for hip surgery and surgery on the arteries of the neck, the risks for surgery are much higher for residents of Boston than for New Havenites. For these conditions, New Haven clinicians prefer the more conservative medical management.

Similar patterns of treatment variation exist for the other conditions listed in Table One.

Table One

Angina (chest pain due to clogged arteries in the heart)	Bypass surgery vs. angioplasty vs. drugs vs. dietary and life-style modification
Gallstones	Surgery vs. stone crushing vs. medical management vs. watchful waiting
Peripheral Vascular Disease	Bypass surgery vs. angioplasty vs. medical management
Cataracts	Lens extraction (by type) vs. watchful waiting
Arthritis of the Knee or Hip	Joint replacement vs. medical management
Prostatism (benign prostatic hyperplasia, or BPH)	Surgery (by type) vs. balloon dilation vs. drugs micro-wave diathermy vs. watchful waiting
Back Pain due to Disc Disease	Surgery (by type) vs. various medical management strategies
Atherosclerosis of Carotid Artery with Threat of Stroke	Surgery vs. aspirin

WHICH RATE IS RIGHT?

Let me try to help the committee grapple with the question of how much surgery is "unnecessary." Clearly, surgery that doesn't produce a benefit that patients want is "unnecessary," and I certainly agree with my colleague Dr. Brook that if experts agree that there is no benefit, it would be a good idea not to use that care. Unfortunately, even the experts can be wrong, and until adequate studies are performed,

we really won't know what works. Outcomes research is crucial to any national strategy to learn the extent of unnecessary care; but there is yet another dimension to the question. Even care that works is inappropriate if it is not wanted by patients. *The true demand for care becomes apparent only when patients are informed about options, about what is known and not known about the outcomes of the various treatments, and when they are free to choose according to their own preferences.* To learn which rate is right, the doctor-patient relationship must be changed from one in which physicians prescribe care based on the assumption that they know their patients' preferences, to one where patients actively share in making decisions.

I learned about the importance of outcomes research and reform of the doctor-patient relationship through a series of studies my colleagues and I undertook to investigate the cause of variations in treatment for prostate disease. We learned that for this condition, the amount of resources needed to provide care that works and that patients actually want is substantially less than the amount now invested. This research shows what can be done to obtain working answers to questions about unnecessary care.

Learning what works—The first task in our research agenda was to uncover the differences in theory that motivated the treatment decisions of physicians. Some urologists, we learned, believed that surgery should be prescribed early in the course of BPH because they believed that the condition, left untreated, often becomes life-threatening. If one waits, these physicians believed, the patient will be older and sicker and more likely to die when surgery is needed to relieve bladder or kidney obstruction; therefore early surgery enhances life expectancy. Other urologists, more optimistic about the natural history of untreated BPH, argued that surgery was warranted in most men because of its ability to reduce symptoms and improve the quality of life.

The outcomes research we undertook showed that the preventive theory was incorrect. Early surgery appears to lead to a slight decrease in life expectancy, because for most men BPH does not progress to life-threatening obstruction. Those without evidence of such obstruction are better off with watchful waiting, if the goal of treatment is to increase life expectancy. If prostate surgery has a place for men with symptoms, it is in accordance with the quality of life theory.

Learning What Patients Want—The uncertainty about which rate is right is more profound than the failure to understand the theoretical basis for clinical decisionmaking or to measure the outcomes that matter to patients. The urologists we met, including those who believed in the quality of life hypothesis, practiced within the delegated decisionmaking tradition. They understood that they bore a special responsibility as the patient's agent to interpret for him what he needed and to convince him, for reasons of his own best interest, to accept their prescription. Yet what patients want cannot be predicted from objective information available to the physician. Data gained from the physical examination, laboratory tests measuring such factors as urine flow, or even answers to questions about the severity of symptoms or impairment of quality of life did not predict what the individual patient wanted. Patients who by all such objective measures are similar differ in their preferences for treatment. *To know what patients want, physicians must ask them.*

Evidence for Excess Capacity—When offered a choice, nearly 80 percent of men with severe symptoms preferred to live with their symptoms rather than undergo the risk of operation. The degree to which patients were bothered by their symptoms and how much they feared impotence or other sexual complications were the important determinants of choice. When decisionmaking changed from a model that relied on the doctor to prescribe treatment to one in which decisions were shared, the link between supply and utilization was broken, and the per capita rates of surgery declined. The trend toward conservative treatment choice was evident even in HMOs, where the rates of surgery are already relatively low and where patients face no cost barriers at the point of delivery. The rates in Kaiser Permanente in Denver and at Group Health Cooperative in Seattle fell about 50 percent when shared decisionmaking was instituted. It became evident that even the HMOs had invested more resources in the treatment of prostate disease than patients wanted.

THE PROBLEM OF VARIATIONS IN PLACE OF TREATMENT

There is another form of practice variation of equal, if not greater, importance: the variations in rates of hospitalization that follow from differences in per capita supply of hospital resources. The effect that the supply of hospital beds exerts is almost entirely on physicians' decisions to treat patients with conditions such as congestive heart failure, pneumonia or low back pain in the hospital, rather than elsewhere—at home, in hospices, nursing homes or clinics. Again, Boston and New Haven provide an example. In 1989, the number of staffed hospital beds invested

in the health of Bostonians was 3.8 per 1,000 residents; for New Haven, the rate was 2.6 per thousand. The number of employees required to maintain those beds for Bostonians was 23 per 1,000 residents. By virtue of this incremental investment, Bostonians experience nearly 60 percent higher rates of hospitalization for a host of chronic and acute medical conditions that do the residents of New Haven, who are much more likely to be treated elsewhere in the community.

Variations such as these are typical of the entire U.S. health care economy. The per capita numbers of hospital beds among local markets within a state typically range from fewer than 2.5 beds per 1,000 to well over 5.0 beds; the per capita numbers of hospital employees and expenditures show two-fold or greater differences.

Evidence for Excess Capacity— As the numbers of beds increase, more resources are invested in the care of the chronically ill, as measured by the proportion admitted to the hospital and the frequency of readmission. More is invested in the last year of life and in terminal care. The quality of death is affected; residents of communities with more hospital beds per capita experience a greater probability that when death occurs, it will occur in a hospital. This effect on the place of death is a constant and near linear function of per capita bed use, ranging from about 30 percent of deaths occurring in hospitals in areas with low per capita bed supply areas to 60 percent in high per capita areas. (Exhibit One)

We found no evidence that increasing the investment in acute hospital care above the level in New Haven produced better health care outcomes. Mortality rates for Bostonians were about the same. When we extended the study to take all areas into account, the evidence was that more is not better: mortality rates tended to be higher in high rate areas, even when differences in demographic characteristics and age structure were taken into account.

THE PROBLEM OF EXCESS CAPACITY

Why, indeed, should greater spending be expected to bring better results? Hospital capacity is not based on explicit theories about what works in medicine. The optimal number of beds is unknown. The number actually built in a community or made available in an HMO has no theoretic or empirical basis related to health outcomes. One looks in vain in medical texts to learn how many beds are needed for treating a population's burden of illness. The number of beds is the result of the way the hospital industry has been planned and regulated. Per capita rates are arbitrary, the product of imperatives of institutions, communities, managed care companies and regulators--not the needs of patients or dictates of medical science.

The number of physicians who are trained is governed by equally arbitrary policies, many of which were set in the 1960s when there was great concern about medical scarcity. The number of specialists trained is the product of administrative and political choices, not the numbers required to produce services that are known to work or that patients want. In the case of procedure-oriented specialties, supply is well in excess of the number of practitioners required to produce the treatments that physicians agree are efficacious. For example, when neurosurgeons enter medical markets, they almost invariably find that the available supply has already taken care of the demand for surgical management of brain tumors and head trauma, which are the neurosurgical procedures that all physicians agree are needed. As a consequence neurosurgeons must invest most of their efforts in treating conditions for which there are valid non-surgical options. The most readily available opportunities are the surgical management of carotid artery disease and back pain due to disc disease, conditions found in Table One. For these conditions, the rates of surgery show large variations--determined by the numbers of neurosurgeons in the population—among neighboring communities. In fee-for-service markets, all physician specialists find employment.

The movement to capitation will change this dynamic. The numbers of specialists available to the U.S. health care economy are well in excess of the numbers that pre-paid group practices such as Kaiser Permanente believe are needed. (Exhibit Two). If the interpretation of population need for specialist care exemplified by the hiring practices of these organizations becomes the norm, large numbers of specialists in the U.S. will face unemployment. The irony is that the numbers of specialists employed per capita by these HMOs is arbitrary from the point of view of patient demand in a reformed health care micro-economy. We learned that even the relatively low per capita number of urologists employed by HMOs was in excess of the numbers required when patients were free to choose their treatments for prostate disease.

The strategies for health care reform that emphasize capitation free the professional workplace from the incentives of fee-for-service medicine and create an opportunity to build a micro-economy where patient choice determines the rates of utiliza-

tion of treatments. Learning what works and what patients want for the conditions listed in Table One has enormous importance for the health care economy: more than half of the major surgery performed in the United States is devoted to these conditions. I believe that for most, if not all, of these conditions, reform along the lines I am calling for will lead to a decline in demand, perhaps to a level well below the rates of surgery now provided in HMOs. Patients, I believe, will prove more risk averse than physicians, and many who are now recommended for invasive treatments will choose more conservative approaches, when they are offered a choice. If so, then large amounts of resources will be available to reallocate. If not, then policy makers will at least know that medical progress and the wants of patients rather than supplier-induced demand drive costs upward. In the meantime, however, the evidence we have shows that more is not better.

POLICIES OF REFORM

I want to urge that the Clinton Health Plan and any alternative proposals that come before the Congress be evaluated in terms of their programs for improving the scientific and ethical basis of clinical decisionmaking and their ability to set limits and deal with the problems of excess capacity.

Certain principles and guidelines that find their empirical justification in the study of geographic variations and outcomes research may help the Congress with this task.

PRINCIPLES FOR SETTING OF LIMITS

The first concerns the general welfare of the public: *It is safe for patients and in the public interest to place global restrictions on growth.* The excesses in capacity that exist in our health care system mean that the amount spent on health care can be directly limited. Studies of the geographic variations in services in this country provide solid evidence that the capacities of the hospital industry and of the physician specialty workforce are now well in excess of what is required to provide services that are efficacious and that patients actually want. Most medical resources are allocated for treatments for which the theoretical basis for allocation is implicitly associated with the supply of resources and for which there is no empirical evidence that more is better. The nation can and should deal directly with the forces of inflationary growth in the health care sector—with policies that determine the numbers and distribution of manpower, the size of the hospital industry, and the quantities of technology—without fear that such actions will induce rationing of services that are known to be valuable. A health care system can be achieved that is in equilibrium with other sectors of the national economy without fear that valuable services must be rationed.

The second principle concerns the welfare of those who do not now have access to care because they lack insurance: *Full entitlement of all Americans to health care can be instituted without increases in the proportion of GNP invested in health and without a loss of welfare to those now insured.* The fear that policies that extend health care entitlement to all citizens must necessarily exacerbate the cost crisis is unwarranted. The dynamics that determine the capacity and costs of health care markets are to a large extent independent of illness rates and the demands of patients. Fewer than 15 percent of Americans are uninsured. An understanding of the epidemiology of medical care leads to the prediction that their entitlement would permit them to be absorbed into the health care system without loss of benefit to those now in and without any special increase in aggregate expenditures. The capacity to treat the uninsured is already there; the trick is to make it possible for the uninsured to compete on an equal basis for the attention of the health care system. In a steady state situation, the increases in costs for treating the uninsured will be offset by savings realized by reduced utilization among those now insured.

The third principle concerns unmet needs such as long-term and community-based care, as well as the special interests of patients for whom expensive medical care is effective: *The resources required to meet unmet needs should be obtained by reallocation of excess capacity and not by rationing effective care.* The excess resources now invested in acute hospital care should be reallocated to meet unmet needs for community-based care or long-term care. Reallocated excess capacity, rather than rationing of effective high cost treatments that patients want, should be the resource for meeting unmet needs. Every state has its own Bostons and New Havens, and large quantities of resources are available for reallocation. Let me give you a sense of the magnitude of the opportunity. If, in 1989, the utilization patterns of Boston had been like those of New Haven, 1000 hospital beds in Boston would have gone unused; 7,800 health care workers and $500 million would have been

made available for reallocation to other medical needs. Similar opportunities exist throughout the country.

The issue of the physician workforce policy merits special Congressional attention. The federal government's subsidies to graduate medical education have played an important role in stimulating the excess supply of specialists, and federal reimbursement policy through the Medicare program makes the situation worse. The current imbalance will not be easily redressed; even with Draconian cuts in the numbers, it would take years to reduce the supply of specialists towards the numbers per capita required by managed care organizations such as Kaiser Permanente. For example, the supply of radiologists is so far in excess of the per capita numbers required by HMOs that even if no more radiologists were produced for the next 20 years, the numbers would still be in excess.

New thinking is required. One of our unmet needs is to learn what works and how to produce care of high quality. A dynamic policy will make it possible for physicians and other health care workers to allocate time to the complex tasks of managing quality in modern systems of care. The complexities of modern technology require a flexible, lifetime approach to professional education. A dynamic policy will provide the opportunities for professional renewal, and even the adoption of new specialties. The requirements for innovation suggest new roles for academic medical centers in fostering outcomes research, in promoting networks of quality, and in providing lifetime learning.

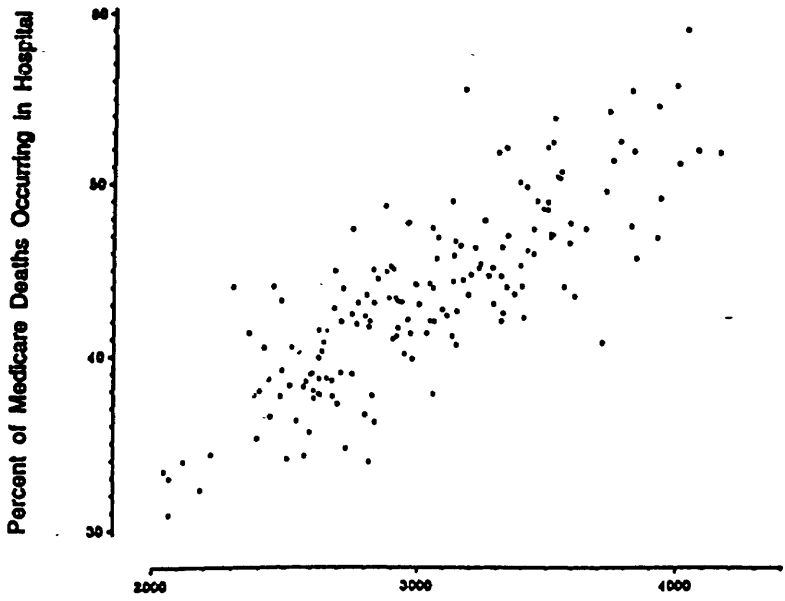
PRINCIPLES FOR REFORM OF THE DOCTOR-PATIENT RELATIONSHIP

Whatever the shape the Congress gives to the new American health care economy, I urge that the historic opportunity to promote reform of the scientific and ethical basis of clinical decisionmaking not be missed.

The essential base for this reform is a strong, well-funded federal science policy for the evaluative clinical sciences. In an age of increasing technological complexity and choice, as well as increasing public sector involvement in health care, it is essential for public policy to support the needed improvements in the scientific and ethical basis of clinical medicine made possible by the evaluative sciences. The overall goals should be to: (1) establish the evaluative sciences as mainstream disciplines in the nation's professional schools, and as an expected competency for the practice of medicine, on equal footing with the biomedical sciences; (2) establish the ethic of evaluation as a defining characteristic of the competent health care professional; and (3) provide the focus for empowering the health professions to take charge of the multiple tasks required to assure quality, reduce supplier- induced demand, and promote lifetime learning.

It is also essential that federal oversight be dedicated to promoting reform of the doctor-patient relationship. State health plans, health alliances and accountable health plans should each be evaluated in terms of how well they set in motion the processes to meet the following guiding principles:

1. Patients should be fully informed about what is known and not known about the outcomes of the relevant treatment options;
2. Patient preferences should determine the choice of interventions among available options;
3. The quality of care should be continuously monitored and improved; and
4. The outcomes of new as well as conventional treatment theories should be continuously evaluated and re-evaluated.



Investment in Hospital Care Measured in Patient Days of Care Delivered (per 1,000 Medicare Enrollees, all New England Areas)

Adjusted for Sociodemographic Characteristics

Exhibit One

Shows the relationship between the level of investment in hospital care measured in patient days of care per 1,000 Medicare enrollees (horizontal axis) and the proportion of Medicare deaths that occur while the patient is hospitalized rather than at home, in a nursing home, clinic or hospice. Each dot represents the experience of one New England Hospital Service area. As hospital capacity increases, the intensity of inpatient investment in the care of the chronically ill and terminally ill patients increases. The incremental investments do not pay off in terms of longer survival of the population. Indeed, paradoxically, greater investment is associated with higher population-based death rates, suggesting that the iatrogenic effects of hospitalization are not outweighed by the benefits.

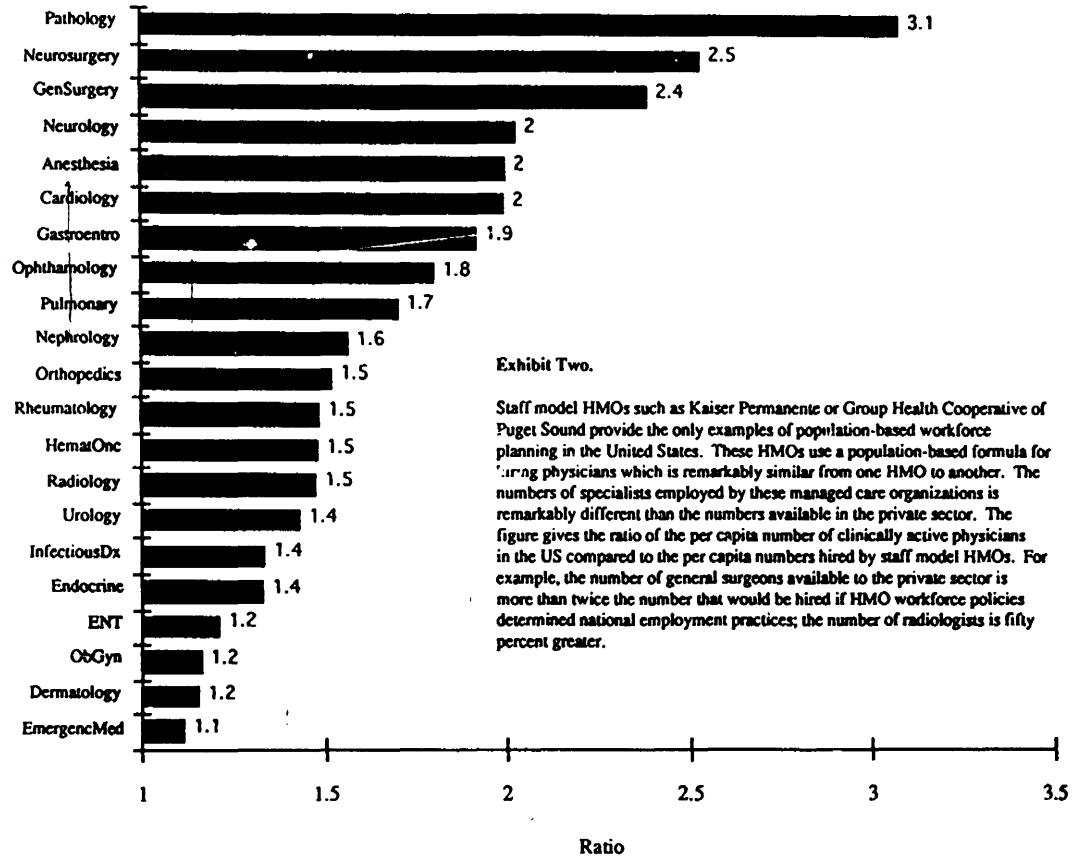


Exhibit Two.

Staff model HMOs such as Kaiser Permanente or Group Health Cooperative of Puget Sound provide the only examples of population-based workforce planning in the United States. These HMOs use a population-based formula for hiring physicians which is remarkably similar from one HMO to another. The numbers of specialists employed by these managed care organizations is remarkably different than the numbers available in the private sector. The figure gives the ratio of the per capita number of clinically active physicians in the US compared to the per capita numbers hired by staff model HMOs. For example, the number of general surgeons available to the private sector is more than twice the number that would be hired if HMO workforce policies determined national employment practices; the number of radiologists is fifty percent greater.

RESPONSE OF DR. WENNBERG TO A QUESTION SUBMITTED BY SENATOR PACKWOOD

Question. You say that if we eliminate inappropriate care in the health system we can:

1. Cover everyone
2. Not have to ration services; and
3. Keep the rate of growth in health care spending to the consumer price index (CPI).

The Clinton plan places a cap on health care spending which equals population growth plus CPI by the year 1999. Do you believe that we can lower the rate of growth in health care spending to the CPI in just 5 years? If we do this, what can we expect to see in terms of access and quality during this five year period?

Answer. Thank you for your question. In my written testimony I do indeed state my belief that we can cover everyone without rationing care. The savings I am talking about go far beyond curtailing administrative costs and fraud and abuse. I propose that we should reallocate resources that are now used inefficiently, in the sense that they produce care with no recognized benefit, or care that is not wanted by patients. Let me pinpoint the sources of these savings:

Reallocate excess capacity in acute hospital care. Local health care markets vary dramatically in their aggregate rate of use of health care. For example, the proportion of the GDP now spent on health care for New Haven residents is substantially lower than the amount spent on Bostonians, *without evidence that more is better.* Much of the difference in expenditures between the two communities is explained by the level of investment in acute hospital care for patients with chronic illnesses. As I indicated in my oral testimony, if in 1989 the residents of Boston used hospital resources at the per capita level in New Haven, Boston would need 1,000 fewer hospital beds, and 7,800 health care workers and $500 million would be available for reallocation to meet other needs.

There are no strong theories as to what benefits are accruing to the citizens of communities with higher investments in health care resources. Mortality rates are not lower in areas with greater investments, despite much more intensive investment in acute care technologies. Clinicians in low rate areas do not believe they are rationing care; nor do they recognize the theoretical advantages of per capita investments greater than they themselves are making. The epidemiological data make the case that is safe for patients and in the public interest to use these investments in care more efficiently.

Introduce shared decisionmaking for "high variation" surgical conditions. In my written testimony, I listed nine conditions that account for more than 50 percent of the elective surgery in the United States. For each of these conditions, there are appropriate alternative treatments. Under the current model of clinical decisionmaking, which delegates decisionmaking to physicians, the supply of medical personnel and their own preferences with regard to these options is a major factor in determining the "demand" for care. Under shared decisionmaking, the link between supply and utilization is broken. The epidemiological data available so far make the case that patients are more risk-averse than physicians, choosing less invasive care more often than when utilization depends on professional prescriptions. We have seen this for prostate disease, where surgery rates dropped by up to 50 percent (even in an HMO) after shared decisionmaking was adopted. On the basis of what I have seen, I am willing to estimate substantial savings once patients understand the risks and benefits of modern treatments.

Install Outcomes Research as a Barrier to Early Diffusion of Untested Technology. Early adoption of unproven treatments is a pervasive problem in the conduct of medical practice. Outcomes research, through its emphasis on the iterative, long-term evaluation of all alternatives for treating common conditions (such as those in Table One of my written presentation), offers a focus for organizing the professional leadership to adopt the ethic of evaluation. This ethic should include participation in prospective clinical trials and other outcomes research projects as part of what is expected of a good professional. Among academic leaders and leaders of specialty societies, there have been some remarkable shifts in opinion about the responsibilities of the profession for learning what works. The amount of money we spend on inappropriate use of technologies is enormous; fostering a change in the profession's attitude toward the rapid adoption of untested technology can unquestionably save money.

Deal with Excess Capacity in the Supply of Specialists. The numbers of specialists available in the U.S. as a whole greatly exceeds the numbers that would be needed if HMOs like Kaiser-Permanente or Group Health Cooperative of Puget Sound became the model for organizing health care in the United States. We can and should

provide opportunities for current supply to find equilibrium with demand as the health care system shifts toward shared decisionmaking. The need is not for more resources, but for better use of those we have. Dr. C. Everett Koop, Dr. Robert Keller and I have suggested the formation of Regional Professional Foundations to provide the profession with some of the tools it needs to deal with problems in the quality of care and changing social demands under health care reform. We would be pleased to provide you with further details about this proposal.

Can we lower the growth in health care spending to the CPI in just five years? Yes, but it must become a national goal to do so without harming patients. The incentives that are established, and how well they match with local and regional realities in terms of current distribution of supply and social and economic organization of providers, will be decisive. Since the Congress will set the rules, I must respectfully redirect this question to the Committee.

RESPONSE OF DR. WENNBERG TO A QUESTION SUBMITTED BY SENATOR DURENBERGER

Question. Dr. Wennberg, your testimony on excess capacity is quite compelling. In 1990, the number of staffed hospital beds in Minnesota was 3.2 per 1,000 residents, while the Twin Cities metro area was only 2.7 per 1,000. In addition, preliminary data for more recent estimates is showing that beds are decreasing while the population is increasing. In other words, judging against the range you cited of 2.5 beds per 1,000 to over 5.0 beds, Minnesota has squeezed much of the excess capacity out of the system. Do you think government should be in charge of this process or the market?

As an example, I cite the closing of the Mt. Sinai Hospital in Minneapolis. When government tried to close this hospital, there was a loud hew and cry about the need for this type of hospital and fingers pointed at public officials. As a result, the hospital was not closed. However, several years later, when it simply could not compete in the marketplace, the hospital closed with nary a whimper. Doesn't that teach us something about the superiority of the marketplace to the political process?

Answer. Thank you for your question about whether the hospital or the market should be in charge of the process of getting rid of excess hospital capacity. The closing of Mt. Sinai is a good example of what the market can do in Minneapolis, and I am sure we will hear of more cases like it in California and in other places where competition is well situated to deal with excess capacity by creating incentives for health plans to choose the more efficient hospital. But we must not assume that models that work in one jurisdiction will work in another.

I hope the Congress will consider carefully the complementary roles that competition and cooperation play in the theory of continuous quality improvement. Within an organization, stability, cooperation, self-study and systems building are the key to a better product. Competition decides which product is better at what price. But in many parts of the country, population density is too low to support more than one, or possibly two, efficient hospitals, and it is difficult to see how a hospital can compete effectively with itself to downsize.

In some parts of the country, government or quasi-government strategies for containing acute hospital sector costs seem to work well. Rochester, N.Y., and New Haven, Ct., have per capita hospital costs that are about as low as Minneapolis. These are the counter-examples to Minneapolis, stories of the success of government regulation and regional planning in containing capacity. In other parts of the country, regulation doesn't do very well; Massachusetts is as highly regulated as Connecticut, but with very different results. The devil is in the details.

COMMUNICATIONS

American Medical Association
Physicians dedicated to the health of America

James S. Todd, MD
Executive Vice President

515 North State Street
Chicago, Illinois 60610

312 464-5000
312 464-4184 Fax

November 8, 1993

The Honorable Daniel P. Moynihan
Chairman
Committee on Finance
United States Senate
205 Dirksen Senate Office Building
Washington, DC 20510

RE: Medical Practice Patterns and Appropriateness of Care

Dear Chairman Moynihan:

We appreciate the Committee's interest in the issues of medical care quality, utilization, and costs, especially in view of research conducted in the 1970s and the 1980s on the appropriate use of medical services for cardiac pacemaker implants, carotid endarterectomy, coronary artery bypass surgery, coronary angiography, and upper gastrointestinal endoscopy. The findings reported by John T. Kelly, MD, PhD, and Shirley E. Kellie, MD, MSc, on more recent data from Medicare peer review organizations indicate lower levels of unnecessary hospital admissions and medical services than the significant levels reported earlier.

As the health system reform dialogue continues, the American Medical Association (AMA) welcomes the discussion of appropriate use of medical services with respect to quality, utilization, and cost of medical care. We further believe that clinical practice guidelines that are properly developed, disseminated, and implemented will help to ensure that patients receive only appropriate, effective, and necessary medical care. The AMA and over 45 physician organizations have taken a leading role in their development and continues to actively participate with the Agency for Health Care Policy and Research (AHCPR) in this important effort.

The AMA is pleased to submit for the record of the Senate Finance Committee hearing of October 26, 1993, Medical Practice Patterns and Appropriateness of Care, the attached study reported in the Archives of Pathology and Laboratory Medicine in November 1990.

We would be pleased to provide any additional assistance as the Senate Finance Committee continues to examine this important issue.

Sincerely,

James S. Todd MD

James S. Todd, MD

Attachment

Reprinted from the Archives of Pathology and Laboratory Medicine
November 1990 Volume 114

Appropriateness of Medical Care

Findings, Strategies

John T. Kelly, MD, PhD, Shirley E. Kellie, MD, MSc

• **Concerns regarding significant levels of inappropriate medical services—as high as 20% or more—continue to influence discussions regarding medical care quality, utilization, and costs. The basis of these concerns are findings from a series of studies of the appropriateness of use of several medical and surgical services provided in the late 1970s and early 1980s—cardiac pacemaker implants, carotid endarterectomy, coronary artery bypass surgery, coronary angiography, and upper gastrointestinal endoscopy. More recent data from Medicare peer review organizations, however, indicate lower levels of unnecessary hospital admissions and medical services. Despite uncertainties regarding the extent of inappropriate care, additional efforts are required to better define appropriate medical care. A promising effort to meet this need is the development of practice parameters, which include practice guidelines and standards.**

(*Arch Pathol Lab Med.* 1990;114: 1119-1121)

Concerns that a significant portion—as much as 20% or more—of all medical services provided in the United States is unnecessary, or performed for inappropriate indications, have influenced current public policy discussions regarding quality, utilization, and cost of medical care.[1,2] These statements are based largely on findings from a series of studies of appropriateness of use of several medical and surgical services provided in the late 1970s and early 1980s—cardiac pacemaker implants,[3] carotid endarterectomy,[4,5] coronary artery bypass surgery,[6] coronary angiography[7,8] and upper gastrointestinal endoscopy.[9] Data from Medicare Peer Review Organizations (PROs)[10] and carriers[11,12] indicate lower levels of unnecessary hospital admissions and medical services. Despite uncertainties regarding the extent of inappropriate or unnecessary medical care, additional efforts are needed to better define appropriate care by specifying appropriate use of procedures and management for specific medical conditions. A promising effort to meet this need is the development of practice parameters. Practice parameters, which include practice guidelines and practice standards, are strategies for patient management developed to assist physicians in clinical decision making.[13]

APPROPRIATENESS STUDIES: METHODS, FINDINGS

Greenspan et al[3] reviewed retrospectively the appropriateness of indications for Medicare-reimbursed permanent cardiac pacemaker implants performed at 30 hospitals in Philadelphia, Pa, during the first 6 months of 1983. Initially, diagnoses recorded on medical record face-sheets were reviewed. Subsequently, in-depth medical record reviews were carried out to determine the adequacy of evidence to support the face-sheet diagnoses, as reflected in the diagnostic evaluations and their documentation in the records. Criteria were developed and used by a review panel to assess the appropriateness of the indications for the pacemaker implants. Subsequently, hospitals were given an opportunity to review and provide feedback to the panel with regard to determinations of the appropriateness of the reviewed implants. Based on their review, the panel determined that 44% of implants were definitely indicated, 36% possibly indicated, and 20% not indicated.

This study reviewed pacemaker implants in 1983; these implants were performed prior to the publication in 1984 of guidelines for use of permanent cardiac pacemaker implantations by the Joint Task Force of the American Heart Association, Dallas, Tex, and the American College of Cardiology, Bethesda, Md.[14] During the 1980s, physicians have significantly modified their use of pacemakers,[15] as reflected in a 27.9% reduction in the rate of permanent pacemaker implants observed between 1983 and 1986.[16]

Winslow et al,[4] Merrick et al,[5] Winslow et al,[6] Chassin et al,[7,8] and Kahn et al[9] studied the appropriateness of use of coronary artery bypass graft surgery, carotid endarterectomy, coronary angiography, and upper gastrointestinal endoscopy. The procedures reviewed were performed between 1979 and 1982. These researchers used a modified Delphi technique to develop appropriateness criteria to assess the appropriateness of indications for which the procedures were performed.[17] Appropriate health care was defined by these investigators ". . . to mean that the expected health benefit (including both quality of life and/or longevity) exceeded the expected negative consequences by a sufficiently wide margin so that the procedure was worth doing."[18] The ap-

Accepted for publication September 5, 1990
From the Office of Quality Assurance, American Medical Association, Chicago, Ill.
Presented at the College of American Pathologists Conference XVII Quality Assurance in Pathology and Laboratory Medicine: Perspectives for the 1990s, April 30, 1990
Reprint requests to the Office of Quality Assurance, American Medical Association, 515 N State St, Chicago, IL 60610 (Dr Kelly).

propriateness criteria were subsequently used to assess information abstracted from medical records to determine the appropriateness of use of each of the four studied procedures. Based on findings in this process, the investigators report that 14% of coronary artery bypass surgeries, 32% of carotid endarterectomies, 17% of coronary angiographies, and 17% of upper gastrointestinal tract endoscopies were performed for inappropriate indications.

Findings reported by these investigators[4,5] are based on review of procedures performed in the late 1970s and the early 1980s. During the 7 to 11 years that have elapsed since the studied procedures were performed, several physician organizations have published practice parameters for the appropriate use of three of these four procedures: the American College of Cardiology and the American Heart Association for use of coronary angiography;[7] the American Society of Gastrointestinal Endoscopy, Manchester, Mass,[8] and the American College of Physicians, Philadelphia, Pa,[9] for use of upper gastrointestinal tract endoscopy; and carotid endarterectomy.[10,11] Some evidence is present to suggest that use of carotid endarterectomy is declining in the United States.[12]

OTHER DATA SOURCES

The Health Care Financing Administration (HCFA), through contracts with PROs, retrospectively reviewed the appropriateness of 7 213 265 Medicare-reimbursed hospital admissions (26% of 27 397 688 total admissions) for the period from 1986 through 1988.[13] Based on retrospective reviews of admissions in which the reasons for admission were assessed against preestablished PRO criteria for appropriate admission, approximately 2% of the hospital admissions were identified as unnecessary, because the care could have been provided at a lower level (eg, in the ambulatory setting), and/or because the services were medically unnecessary.

During the 1988 calendar year, Medicare carriers reviewed 383.1 million Medicare Part B claims submitted for payment. These claims were for physicians' services, outpatient hospital and diagnostic services, and a wide variety of other medical services.[14] Review of these claims resulted in full or partial denial of 17.1% of all claims, payment was denied for 14.1% of all billed charges. Payment determinations were based on findings as to whether the services were ". . . reasonable and necessary for the diagnosis or treatment of illness or injury or to improve functioning of a malformed body member."[15] Of the total disallowed charges, 13.8% were due to determinations that the services were medically unnecessary. Thus, approximately 1.9% of the Part B charges submitted in 1988 were determined to be medically unnecessary.

EVALUATION OF FINDINGS

Comparison of levels of reported unnecessary use of cardiac pacemakers, coronary angiography, coronary artery bypass surgery, carotid endarterectomy, and upper gastrointestinal tract endoscopy with lower levels of unnecessary hospital admissions and other medical services identified by Medicare PROs and carriers raises issues regarding the basis for determining what constitutes appropriate medical care. To a significant extent, experts may differ with regard to their interpretation of published research findings and collective clinical experience, both of which serve as a basis for establishing appropriate clinical management. Wennberg[2] and Wennberg et al[3] attribute their findings of geographic variation in medical care utilization rates to physicians' uncertainties regarding the clinical outcomes associated with specific medical practices. Such uncertainties regarding optimal patient management strategies contribute to the difficulty of establishing criteria for appropriate medical care, as well as determining existing levels of inappropriate care.

COMMENT

Efforts to assure that necessary medical care is provided and unnecessary medical services are avoided are evident in the activities of physician organizations to develop practice parameters,[7-9] and outlined in recent federal legislation to increase funding for clinical outcome research and to facilitate the development of practice guidelines.[16]

As of 1980, eight medical societies had developed practice parameters; by 1990, at least 26 physician organizations had already developed practice parameters, and more than 10 other physician organizations had plans for future development of practice parameters.[9] In 1990, over 700 practice parameters had already been developed by physician organizations, and over 120 practice parameters were under development by physician organizations.[8,9] Many additional practice parameters are anticipated to be developed during the next decade.

Analysis of available practice parameters indicates substantial variation in their method of development, format, content, purpose, and application. Some practice parameters are based primarily on a systematic review and synthesis of published scientific studies; other parameters are based principally on expert consensus.[17] To facilitate the scientific validity and clinical utility of practice parameters, various attributes to guide the development of practice parameters have been proposed.[18-20] The Institute of Medicine of the National Academy of Sciences, Washington, DC, has proposed attributes of practice guidelines that address the following: validity, reliability/reproducibility, clinical applicability, clinical flexibility, clarity, multidisciplinary process, scheduled review, and documentation.[18] The American Medical Association/Specialty Society Practice Parameters Partnership and Forum have proposed attributes that specify the following: practice parameters should be developed by or in conjunction with physician organizations, reliable methodologies that integrate relevant research findings and appropriate clinical expertise should be used, practice parameters should be as comprehensive and specific as possible, practice parameters should be based on current information, and practice parameters should be widely disseminated.[19]

Evaluations of the extent of the impact of practice parameters on clinical practice have been limited. Reductions in the utilization rates of services such as permanent cardiac pacemakers, cesarean-section rates, and pelvimetry

have been associated with the development of practice parameters.[8-10] Decreases in the incidence of adverse anesthesia events have been attributed to the development of clinical standards or anesthesiology.[8,9,11] Improvements in the performance of sampling techniques for Papanicolaou smears and reductions in the average length of stay in a hospital intensive care unit have been attributed to the use of practice guidelines.[8,12] However, practice guidelines had little effect on cesarean-section rates in Ontario, Canada.[8] Additional efforts are needed to assess the impact of practice parameters on clinical practice and to identify the factors that influence this impact. Issues of particular interest include physician education, quality assurance activities, professional liability, reimbursement decisions, and patient education.[7,8,13]

CONCLUSION

Properly developed practice parameters present a promising strategy to define appropriate medical care, provide a rational foundation for examining the appropriateness of medical care provided, and establish a means to assure appropriate utilization of health care services.

References

1 Sullivan LW. Shattuck lecture: the health care priorities of the Bush administration. *N Engl J Med* 1989;321:125-12b.

2 *Annual Report to Congress, 1989*. Washington DC: Physician Payment Review Commission; 1989.

3 Greenspan AM, Kay HR, Berger BC, Greenberg RM, Greenspan AJ, Gaughan MJS. The incidence of unwarranted implantation of permanent cardiac pacemakers in a large medical population. *N Engl J Med* 1988;318:158-163.

4 Winslow CM, Solomon DH, Chassin MR, et al. The appropriateness of carotid endarterectomy. *N Engl J Med* 1988;318:721-727.

5 Merrick NJ, Brook RH, Fink A, Solomon DH. Use of carotid endarterectomy in five California Veterans Administration Medical Centers. *JAMA* 1986;256:2531-2535.

6 Winslow CM, Kosecoff JB, Chassin M, Kanouse DE, Brook RH. The appropriateness of performing coronary artery bypass surgery. *JAMA* 1988;260:505-509.

7 Chassin MR, Kosecoff J, Park RE, et al. Does inappropriate use explain geographic variations in the use of health care services? a study of three procedures. *JAMA* 1987;258:2533-2537.

8 Chassin MR, Brook RH, Park RE, et al. Variations in the use of medical and surgical services by the Medicare population. *N Engl J Med* 1986;314:285-290.

9 Kahn KL, Chassin MR, Solomon DH, Brook RH, Kosecoff J. Upper gastrointestinal endoscopy. *Ann Intern Med* 1989;110:245-246.

10 *Utilization and Quality Control Peer Review Organizations. Executive Data Summary*. Baltimore, Md: Health Care Financing Administration; 1989.

11. *Medicare Carriers' Manual*. Baltimore, Md: Health Care Financing Administration; 1989, part 3, section 2000.

12. *Carrier Reasonable Charge and Denial Activity Report: January-March 1989*. Baltimore, Md: Health Care Financing Administration; 1989.

13. AMA/Specialty Society Practice Parameters Partnership. *Attributes to Guide the Development of Practice Parameters*. Chicago, Ill: American Medical Association; 1990.

14. Guidelines for permanent cardiac pacemaker implantation: a report of the Joint American College of Cardiology/American Heart Association Task Force on Assessment of Cardiovascular Procedures (Subcommittee on Pacemaker Implantation). *J Am Coll Cardiol* 1984;4:434-442.

15 Parsonnet VP, Bernstein AD, Galasso D. Pacing practices in the USA: a 1985 survey. *PACE* 1987;10:726. Abstract.

16. Mitchell JB, Wedig G, Cromwell J. The Medicare physician fee freeze: What really happened? *Health Aff (Millwood)*. 1989;8:21-33.

17. Brooks, RH, Chassin, MR, Fink A, Solomon DH, Yoseloff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Tech Assess Health Care* 1986;2:53-63.

18 Guidelines for coronary angiography: a report of the American College of Cardiology/American Heart Association Task Force on the Assessment of Diagnostic and Therapeutic Cardiovascular Angiography (Subcommittee on Coronary Angiography). *J Am Coll Cardiol* 1987;10:935-950.

19 American Society for Gastrointestinal Endoscopy. *The Role of Endoscopy in the Surveillance of Premalignant Conditions of the Upper Gastrointestinal Tract. The Role of Endoscopy in the Management of Esophagitis. The Role of Endoscopy in the Management of the Patient With Peptic Ulcer Disease. The Role of Endoscopy in the Management of Upper Gastrointestinal Hemorrhage. Guidelines for Clinical Application*. Manchester, Mass: American Society for Gastrointestinal Endoscopy; 1986; publications Nos. 1002, 1006-1006.

20. Feldman RE. Upper gastrointestinal endoscopy. *Ann Intern Med*. 1989;110:244.

21. Cebul RD, Whisnant JP. Carotid endarterectomy. *Ann Intern Med* 1989;111:660-670.

22. American College of Physicians. Indications for carotid endarterectomy. *Ann Intern Med*. 1989;111:673-677.

23. Pokras R, Dyken ML. Dramatic changes in the performance of endarterectomy for diseases of the extracranial arteries of the head. *Stroke* 1988;19:1289-1290.

24 Wennberg JE. Dealing with medical practice variations: a proposal for action. *Health Aff (Millwood)* 1984; 3:6-32.

25 Wennberg JR, Mulley AG, Hanley D. An assessment of prostatectomy for benign urinary tract obstruction: geographic variations and the evaluation of medical care outcomes. *JAMA* 1988;259:3027-3030.

26. Kelly JT. *Conference on Practice Guidelines*. Washington, DC: Physician Payment Review Commission, October 11, 1988.

27. *Comments on The Medical Care and Quality Improvement Act of 1989 (HR 1695)*. Chicago, Ill: American Medical Association; 1989.

28. *Board of Trustees Report R. A-89*. Chicago, Ill: American Medical Association; 1989.

29. *Board of Trustees Report EE. I-89*. Chicago, Ill: American Medical Association, 1989.

30. Kelly JT, Swartwout JE. Development of practice parameters by physician organizations. *QRB* 1990;16:54-57.

31 Pierce EC. The development of anesthesia guidelines and standards. *QRB* 1990;16:61-64.

32. Sommer A, Weiner JP, Gamble L. Developing specialty-wide standards of practice: the experience of ophthalmology. *QRB* 1990;16:65-70.

33 Speicher CE. Practice parameters: an opportunity for pathologists to take a leadership role in patient care. *Arch Pathol Lab Med* 1990;114:823-824.

34 US House of Representatives. *Omnibus Budget Reconciliation Act of 1989: Conference Report to Accompany HR 3299*. November 21, 1989.

35 Swartwout JE. *Listing of Practice Parameters, Guidelines, and Technology Assessments*. Chicago, Ill: American Medical Association; 1989.

36 *Practice Parameters Update*. Chicago, Ill: American Medical Association; 1990.

37 Johnson KB, Hirshfeld EB, Ile ML, Kelly JT. *Legal Implications of Practice Parameters*. Chicago, Ill: American Medical Association; 1990.

38. Eddy DM. Practice policies-guidelines for methods. *JAMA* 1990;263:1839-1841.

39. Eddy DM. Guidelines for policy statements: the explicit approach. *JAMA*. 1990;263:2239-2243.

40 Leape LL. Practice guidelines and standards: an overview. *QRB* 1990;16:42-49.

41 National Academy of Sciences-Institute of Medicine. Committee to Advise the Public Health Services on Clinical Practice Guidelines. In Field MJ, Lohr KN, eds. *Clinical Practice Guidelines: Directions for a New Program*. Washington, DC: National Academy Press; 1990.

42. Todd JS. Do practice guidelines guide practice? *N Engl J Med* 1990;322:1822.

43. Myers SA, Gleicher N. A successful program to lower cesarean-section rates. *N Engl J Med* 1988;319:1511-1516.

44 Chassin MR, McCue SM. A randomized trial of medical quality assurance: improving physicians' use of pelvimetry. *JAMA*. 1986;256:1012-1016.

45 Holzer JF. The advent of clinical standards for professional liability. *QRB* 1990;16:71-79.

46 Howard HR. HMO improves pap smear screening. *QA Rev*. 1989;1:2-3.

47. Eagle KA, Mulley AG, Skates SJ, et al. Length of stay in the intensive care unit: effects of practice guidelines and feedback. *JAMA*. 1990;264:992-997.

48. Lomas J, Anderson GM, Domnick-Pierre K, et al. Do practice guidelines guide practice? the effect of a consensus statement on the practice of physicians. *N Engl J Med* 1989;321:1306-1311.

49. Wennberg JE, Blowers L, Parker R, Gittelsohn AM. Changes in tonsillectomy rates associated with feedback and review. *Pediatrics* 1977;59:821-824.

50. Kelly JT. Practice parameters called key to effective QA programs. *QA Rev*. 1990;2:1-2.

51. *Council on Medical Service Report D. A-89*. Chicago, Ill: American Medical Association; 1989.

52 White LJ, Ball JR. Integrating practice guidelines with financial incentives. *QRB*. 1990; 16:50-53.

53 Barry MJ, Mulley AG, Fowler FJ, et al. Watchful waiting vs immediate transurethral resection for symptomatic prostatism: the importance of patients' preferences. *JAMA*. 1988;259:3010-3017.

Printed and Published in the United States of America

STATEMENT OF ROBERT H. BROOK, M.D., SC.D. F.A.C.P. PROFESSOR OF MEDICINE AND HEALTH SERVICES, UCLA CENTER FOR THE HEALTH SCIENCES, DIRECTOR, HEALTH SCIENCES PROGRAM, THE RAND CORPORATION *

This testimony consists of a series of facts, statements, and concerns that need to be addressed in any effort to reform the current health care system. The connecting theme is the concern that economic reform, alone, will not be sufficient to ensure that the U.S. will have a truly better health care system than it has today. Economic reforms, such as increasing competition, changing regulations, changing coinsurance rates, using global budgets, or reducing administrative costs, can all have useful outcomes from the point of view of society. But economic reforms are not enough. Other health care reforms must be implemented as well if we want to provide all American people with all necessary care at a price society is willing to afford. Much of what is covered in this testimony is based on work done at RAND/UCLA over the past twenty-five years, although work of other people and organizations is also included.

Health Care Reform Always Has Unexpected Results. For instance, in California requiring medically indigent adults to obtain all their care from designated county hospitals resulted in people dying prematurely for lack of access to medical care.[1] Other forms of cost containment can produce mixed results in terms of quality. RAND's clinical evaluation of the implementation of Medicare's prospective payment DRG program for hospital care demonstrated that, overall, there were few adverse effects on health, even though this program cut length of stay of hospitalized patients for some conditions by 25 percent.[2] That study demonstrated that changing financial incentives can greatly affect what doctors do, without great harm. However, that same study demonstrated that some elderly patients were being discharged from the hospital quicker and sicker than they were before the implementation of the DRG prospective payment program.[3] We also demonstrated that these people died at a rate that was higher than would have been expected if they were discharged medically stable. We have written an editorial in the Journal of the American Medical Association calling physicians to evaluate their elderly patients more carefully before discharging them to return home in such a clinically unstable condition.[4] We have recommended at time of hospital discharge that certain history and physical findings be collected, and that patients either remain in the hospital until they are clinically stable or that, at least, a comprehensive discharge plan be written in the medical record and communicated with the patient and his or her family regarding follow-up of potentially dangerous conditions. We have also urged that the federal government, which funded the original evaluation study of the impact of the DRG program, consider funding a follow-up study to determine whether the situation has gotten better or worse in the ensuing five years. These two examples, the cessation of Medi-Cal coverage for medically indigent adults and the implementation of DRG prospective payment, illustrate the need for objective evaluation of the consequences of reform on both health and cost of medical care. Every reform has unexpected effects. Thus, each reform needs to be carefully evaluated. Such evaluation should include detailed clinical data. It is not sufficient to rely on administrative data or global measures of health, such as infant mortality rates or life expectancy, because these are not affected much by marginal changes in the investment in the health care system.

Economic Reform alone will not improve health care. It is clear, based on work done at RAND and elsewhere, that economic incentives dramatically alter the volume of care received. The Health Insurance Experiment demonstrated that when co-insurance rates increased, people used substantially less care.[5] But changing the volume of care did not result in the selective elimination of those services we do not want and keeping those we do.[6] In particular, we have demonstrated in Israel,[7] the United Kingdom,[8] and for managed care organizations in the United States,[9] that volume of services can be controlled, but that discretionary and inappropriate care is not selectively eliminated. For instance, the Trent region of the U.K., which has global hospital budgets and salaried doctors, does one-seventh the number of coronary angiographies and coronary artery bypass surgeries that we do in the United States.[8] But, at the same time that care for people with serious potentially life threatening conditions such as left main disease or three vessel disease was being rationed, i.e., such people were being placed on waiting lists; 45 percent of coronary angiographies and bypass surgeries rendered in this region were not appropriate. In other words, underuse and overuse of care existed simultaneously. Similarly, in Israel, where care is Provided by competing HMOs, at the same time

*The views and conclusions expressed are those of the author and should not be interpreted as representing those of RAND or any of the agencies sponsoring its research.

that resources were so scarce that closure of one of the country's four medical schools was being considered, 29 percent of the gallbladder operations done in four of that country's hospitals were being performed for reasons that were less than appropriate.[7] Recently, we studied the use of hysterectomies by seven well established and respected managed care plans in the United States. In 16 percent of the cases reviewed the procedure was rated clinically inappropriate and in another 26 percent the procedure was of uncertain clinical benefit.[9] Studies conducted in the 1980s at RAND also found that there was no relationship between the volume of procedures performed in a geographic area and the clinical appropriateness of those procedures.[10]

These and other studies demonstrate that volume of care can be controlled by changing the economic incentives affecting plans, physicians, and hospitals. They also show that we will not selectively eliminate inappropriate procedures and retain appropriate ones if the U.S. does not implement clinical reform at the same time it implements economic reform. It does not matter whether the U.S. adopts a single payor system, competition based on managed care, global budgets, salaried doctors, or whatever. Unless attention is paid at the clinical level to what is being done, it is likely that any system that is adopted will contain a mix of procedures and services that government or society wishes to pay for as well as those for which it does not wish to pay. If we are to have a better health system for all Americans, we must make major changes at both the clinical and economic levels. We can do both simultaneously.

Appropriateness Guidelines Could Be Used To Improve The Practice Of Medicine If Physicians Were Required To Work With The Patient To Complete A Checklist About Whether The Procedure Was Necessary and Appropriate. It is feasible to examine health care reform at a clinical level. Using methods we have developed, it is possible, for the 75 to 100 procedures that make up most of what physicians do, to develop and place in the public domain detailed guidelines or protocols specifying under what clinical circumstances and for what types of patients a particular procedure represents care that is necessary, appropriate, equivocal, or inappropriate. We at RAND, with many collaborators, have done this for eleven procedures. If the federal government wanted to do this, it would be entirely feasible to develop similar guidelines for 75 to 100 procedures within a two- to three-year time period.

To produce such guidelines requires examining the results of scientific studies and obtaining expert judgement. Neither is sufficient alone. We have developed a method that combines the two and produces useful guidelines about appropriateness of care.[11] We have applied our method regarding appropriateness of care to a number of procedures. In general, we find that about one-third to one-quarter of procedures are performed for less than appropriate reasons.[12] By less than appropriate I mean reasons for which the benefit and the risk to the patient are about equal or the risk exceeds the benefit. This finding leads me, as a physician, to conclude that before rationing health care on the basis of cost, it would be worthwhile to develop a health care reform strategy that selectively eliminates payment for those procedures that are discretionary but continues to pay for those that are needed but expensive.[13] If we had available guidelines that precisely defined under what clinical conditions procedures, tests, and services were appropriate, equivocal, or inappropriate, then perhaps such a policy could be implemented.

Let me illustrate with a study RAND conducted on a random sample of patients, who underwent angioplasty for coronary artery disease in the state of New York.[14] Using guidelines based on scientific studies and expert judgement, we divided patients into three groups. About 35 percent of all angioplasties in the state of New York were medically necessary, i.e., the benefit exceeded the risk, the benefit was substantial, most people obtained the benefit, and doctors felt that, if they did not offer this procedure to their patients, they should be held liable. On the other hand, 4 percent of the angioplasties fell into the category where the risk clearly exceeded the benefit, and were deemed inappropriate. The remaining 61 percent of angioplasties could not be clearly categorized as necessary or inappropriate. In a reformed health care system, what percentage of these angioplasties should be covered? Should such a system cover all angioplasties other than those in which the risk clearly exceeds the benefit? Or should it cover only those angioplasties that are clearly medically necessary? It is unlikely that there will be sufficient funds to cover everything that is not clearly inappropriate. But it is possible to cover care that is necessary. I urge the adoption of a health care reform package that explicitly limits coverage of care to those procedures that are necessary and excludes those that are less than necessary. This does not mean that angioplasty should be withheld from those people who want it for what are considered discretionary reasons. If after full disclosure by their physicians, such people really want the procedure (i.e., in tech-

nical jargon the patient's utility for having the procedure exceeds that for not having it), she should be allowed to have it. But such procedures should not be paid for out of public funds.

The use of appropriateness guidelines could also improve the practice of medicine. Would the practice of medicine be improved if, before anybody had one of the 75 to 100 procedures that are dangerous, costly, or possibly very beneficial, the doctor and the patient were required to work together, probably with the aid of a computer, to complete a checklist regarding why the procedure was being performed? This would ensure that, before a patient had one of these procedures, all necessary questions were asked and the patient knew that providing accurate answers to these questions was crucial and could lead to a decision to have or not to have a procedure. For instance, an accurate answer to a question concerning the level of exertion that causes chest pain is crucial to the decision to undergo coronary artery revascularization. It would also ensure that both the patient and the physician knew which diagnostic tests should be performed and/or what alternate therapies should be tried prior to recommending a procedure. In particular, RAND research has shown that for almost one-half of NIH consensus conference recommendations, concerning subjects such as treatment of breast cancer and heart disease, over 50 percent of physicians did not conform with the recommendations. The same study showed that physicians wanted help to improve their performance. In particular, they wanted brief clinical facts at the time they were relevant to individual patients. In short, they wanted guidelines.[15] Thus, explicitly assessing appropriateness prior to performing a procedure would increase communication between the doctor and the patient, shift some power from the doctor to the patient, and result in much more informed consent. Once an appropriateness rating was calculated, the doctor and patient could discuss extenuating clinical circumstances that might make his or her case appropriate as opposed to uncertain, or might result in the patient deciding he or she would like to purchase a service that was considered to be equivocal at best. However, such a purchase would occur only after full disclosure. About $150M per year is needed to develop and maintain in the public domain appropriateness guidelines.

Quality Could Be Used As A Basis For Competition Among Plans, Doctors, And Hospitals, But Doing So Would Require The Development And Testing Of Tools To Measure It. Health care reform provides an opportunity to improve quality of care and thus to introduce competition based on quality. We can measure quality, but not with the currently available tools. There have been virtually no funds available from any public sector agency to develop adequate measures of quality that are sufficiently valid to allow public release of comparisons among plans, physicians, or hospitals. RAND has conducted a few studies and we and other researchers and health plans have developed a few measures, but there will need to be a major investment in the tools to measure quality if we are to pursue any health care reform policy based upon competition over quality.[16] Such an investment will need to be in the $50M per year range. It could result in our ability to give the public valid information about quality at the plan level in three years. Of course, substantially more money would be required to collect and distribute data to the public about quality.

The focus of the quality tools should be on those areas of medicine in which the greatest good or harm can be done. Many of the quality tools currently being talked about relate to things that are easily measured, such as, rates of immunizations, mammograms, pap smears, etc. While these preventive services are important, they are not the areas where most money is being spent and where differences in quality produce the greatest good or harm. On the other hand, quality is a major factor in coronary artery bypass surgery or care of the patient with a heart attack or heart failure. The outcome depends heavily on who performs the procedure or in which hospital the patient is hospitalized. Five extra deaths per hundred people operated on for coronary artery disease is not an unexpected finding, depending on the doctor and the hospital.[17] Likewise, in a national study RAND found that, depending on whether one goes to a hospital at the twenty-fifth percentile or the seventy-fifth percentile of technical process quality, the likelihood one would die from a heart attack or heart failure can increase 25 percent for heart attack and 77 percent for heart failure (e.g., for heart attack patients over sixty-five, the death rate within 30 days of hospital admission is 24 percent if admitted to a hospital in the top quarter of the quality distribution compared to a death rate of 30 percent for a hospital in the bottom quarter).[18] It is not unreasonable to assert that perhaps a quarter of hospital deaths from pneumonia, heart attack, stroke, or heart failure might be preventable if quality of care in American hospitals was better.[19] Thus, any serious competition over quality must focus on those areas in which differences in quality lead to major differences in the length or quality of life. If competition is to occur

over these areas of quality, public tools and standards to measure quality must be developed. I reiterate the words public tools and standards, because it is very important that any health care reform activity seriously commit itself to maintaining, in the public sector, tools and standards of quality. Developing and placing quality standards in the public domain will facilitate widespread adoption of these standards, assessment of the validity and reliability of tools to measure quality, and the assurance that these standards will be kept up-to-date.

Competition over quality also requires that a balance be struck regarding what measures of quality are to be used. There are three basic domains of quality: appropriateness of care, technical excellence, and patient satisfaction. When you or I go to a doctor, we would like to get a procedure if we need it and not get it if we do not need it. In other words, we would like the care to be appropriate. Changing incentives from fee-for-service to managed care or to global budgeting is likely to lead to tremendous pressure to reduce the use of services as illustrated in our DRG study described above. Thus, it is critical that a strategy be developed and implemented to assess underutilization of appropriate and necessary services for those people who are at risk of not receiving needed care. Do people who actually need an angioplasty get one? Do people who need a bone marrow transplant get one? Do people who need hospitalization for pneumonia or expensive antibiotics for a serious infection get them? It is not enough to just measure whether or not people get simple, preventive services, we must also make sure that underutilization is not increased as a function of any health care reform initiative.

Tools are also needed to measure technical excellence. If we get hospitalized for heart failure, we would like care to be good enough that the 30-day death rate was closer toll percent than to 19 percent. If we get vessels put into our heart, we would like them to be handled in a way that resulted in them staying open for eight years rather than four. Finally, measures of patient satisfaction are also needed. But it will be a mistake if, due to the expense of collecting information on appropriateness and technical excellence, competition focuses on patient satisfaction. We know that the style of care patients find satisfactory in a doctor does not necessarily predict whether that doctor uses procedures appropriately or delivers services in a technically excellent manner.[20] Failure to give adequate attention to appropriateness and technical excellence could seriously impair the health of the population.

Measures Of Quality Of Care Are Likely To Show Differences Across Location, Ethnicity, And Gender. This final point involves an ethical issue. Until now, little information has been available about how quality of care varies by ethnic or other patient characteristics. RAND and others are beginning to put some of this evidence together. For instance, we have found that quality of care rendered to hospitalized patients in academic centers is, on average, better than that rendered in nonacademic urban settings. And urban hospitals provide better quality of care than do rural hospitals.[21] What if research explicitly finds and reports that quality of care dramatically varies by plan and shows that life expectancy or quality of life is materially affected by the plan chosen? What if research shows that a low quality plan also has a disproportionate number of members in inner city areas who are minorities or women? How will the political pressure that will result when research demonstrates discrimination against ethnic or gender groups be handled? We may find these results because plans in such areas are poorly organized or managed, because doctors who practice there are less competent than those who practice elsewhere, or for a whole host of other system or professional reasons. Nonetheless, if the U.S. is preparing to embark upon a system of competition over quality, it must he willing to face the consequences that will occur when research finds that quality of care is unevenly distributed across states, areas in states, and racial and ethnic groups. Public release of information about quality and competition at the plan level may help to eliminate some of the variation.

If competition over quality is not an essential element of reform, then the above-noted research RAND has done suggests that economic reforms will result in the use of services decreasing dramatically. This will affect both people who need services and people who do not need services and, if the decline in use is great enough, will harm the average American's health. With this in mind, perhaps health care reform should begin with competition about quality at the plan level and not at the physician or hospital levels. Since so much needs to be done to develop valid measures of quality for external consumption, it would be enough if, within the next few years, the public was assured that they would have access, at the time they enroll in a health care plan, to data about the plan's willingness to provide appropriate care when they need it, about the technical excellence of care, especially acute care (e.g., quality of inpatient care for those with pneumonia) and about patient-centered care activities, such as patient satisfaction. Production of valid information about quality at the physician or hospital level is not currently feasible for most diagnoses

or procedures. (We can develop valid comparisons of quality at the physician level for coronary artery bypass surgery and at the hospital level for pneumonia, stroke, heart attack, and heart failure. The last four diseases make up about 30 percent of hospital deaths, but the patient's choice is likely to be in which plan he enrolls, not to which hospital he goes.) However, if public standards of quality are developed and a data collection system is implemented to measure quality at the plan level that can be demonstrated to be fair, unbiased, valid, and to adequately account for differences in the mix of patients by plan, that will go a long way toward making sure that health care reform does not increase the already large variation in quality that exists in the current system, or reduce the average level of quality for all Americans.

REFERENCES

1. Lurie N, Ward NB, Shapiro MF, et al. "Termination of Medi-Cal benefits: a follow-up study one year later." New England Journal of Medicine. 1986;314:1266–1268.

2. Kahn KL, Keeler EB, Sherwood MJ, et al. "Comparing outcomes of care pre- and post-implementation of the DRG-based prospective payment system." Journal of the American Medical Association. 1990;264:1984–1988.

3. Kosecoff J, Kahn KL, Rogers WH, et al. "Prospective payment system and impairment at discharge: The quicker-and-sicker story revisited." Journal of the American Medical Association. 1990;264:1980–1983.

4. Brook RH, Kahn KL, Kosecoff J, "Assessing clinical instability at discharge clinician's responsibility." Journal of the American Medical Association, 1992;268:1321–1322.

5. Newhouse JP, Manning WG, Morris CN, et al. "Some interim results from a controlled trial in health insurance." New England Journal of Medicine. 1981;305:1501–1507.

6. Lohr K, Brook RH, Kamberg C, et al. "Use of medical care in the RAND Health Insurance Experiment: diagnosis- and service-specific analyses in a randomized controlled trial." Medical Care. 24:9; Supplement, 1986.

7. Pilpel D, Fraser GM, Kosecoff J, et al. "Regional differences in appropriateness of cholecystectomy in a prepaid health insurance system." Public Health Reviews. 1992/93;20:61–74.

8. Bernstein SJ, Kosecoff J, Gray D, et al. "The appropriateness of the use of cardiovascular procedures: British versus U.S. perspectives." International Journal of Technology Assessment in Health Care. 1993;9:3–10.

9. Bernstein SJ, McGlynn EA, Siu AL, et al. "The appropriateness of hysterectomy: a comparison of care in seven health plans." Journal of the American Medical Association. 1993;269:2398–2402.

10. Chassin MR, Kosecoff J, Park RE, et al. "Does inappropriate use explain geographic variations in the use of health services?: a study of three procedures." Journal of the American Medical Association. 1987;258:2533–2537.

11. Brook RH, Chassin MR, Fink A. "A method for the detailed assessment of the appropriateness of medical technologies." International Journal of Technology Assessment in Health Care. 1986;2:53–63.

12. Brook RH, Kamberg CJ, Mayer-Oakes A. "Appropriateness of acute medical care for the elderly: An analysis of the literature." Health Policy. 1990;14:177–280.

13. Brook RH, Lohr KN. "Will we need to ration effective health care?" Issues In Science And Technology. 1986;3:68–77.

14. Hilborne LH, Leape LL., Bernstein SJ, et al. "The appropriateness of use of percutaneous transluminal coronary angioplasty in New York State." Journal of the Medical Association. 1993;269:761–765.

15. Kanouse DE, Brook RH, Winkler JD, et al. *Changing Medical Practice Through Technology Assessment: An Evaluation of the NIH Consensus Development Program.* Health Administration Press, Ann Arbor, MI, 1989.

16. Siu AL, McGlynn EA, Beers MH, et al. Choosing Quality-of-Care Measures Based on the Expected Impact of Improved Quality of Care for the Major Causes of Mortality and Morbidity. JR–03, The RAND Corporation, 1992.

17. Williams S, Nash DB, Goldfarb N. "Differences in mortality from coronary artery bypass graft surgery at five teaching hospitals." Journal of the American Medical Association. 1991;266:810–815.

18. Kahn KL, Rogers WH, Rubenstein LV, et al. "Measuring quality of care with explicit process criteria pre- and post-implementation of the DRG-based prospective payment system." Journal of the American Medical Association. 1990;264:1969–1973.

19. Dubois RW, Brook RH. "Preventable deaths: who, how often, and why?" Annals of Internal Medicine. 1988;109:582–589.
20. Hayward RA, AM Bernard, JS Rosevear, et al. "An Evaluation of Generic Screens for Poor Quality of Hospital Care on a General Medicine Service." Medical Care. 1993;31:394–402.
21. Keeler EB, Rubenstein LV, Kahn KL, et al. "Hospital characteristics and quality of care." Journal of the American Medical Association. 1992;268:1709–1714.

STATEMENT OF DAVID M. EDDY, MD, PH.D.

To put my remarks in context, I need to tell you that although I am a physician trained in surgery, I immediately left clinical practice to get a PhD in applied mathematics and economics. I have spent the last 20 years working with healthcare providers, insurers, businesses and governments assessing the value of medical practices and designing coverage policies and guidelines. In many ways I represent the type of person who will have to implement the mandates of healthcare reform at the level of health plans, providers and patients. I will describe what healthcare reform will look like from the bottom up, and will make recommendations based on my observations of how plans and providers will respond to various components of health reform.

I agree with the testimony of my colleagues this morning that *studies of variations in practice patterns and inappropriate care indicate that there are a lot of very questionable practices that could probably be cut without harming quality in any measurable way. These findings give us good reason to be optimistic that the three objectives of reform can be achieved: containing costs to the CPI, while expanding access to all Americans, without harming quality.* If we could just snap our fingers and immediately reach the "steady state" that Dr. Wennberg described, in which the excess capacity of practitioners and hospital beds is reallocated from ineffective to effective practices, in which all inappropriate uses of procedure are clearly identified, and in which practitioners who are using procedures inappropriately compliantly stop, we could avoid most of the agonizing debates we are currently engaged in.

Unfortunately, we cannot just snap our fingers. Making the changes suggested by the research on practice variations and inappropriate care will be very difficult and controversial. The transition will be marked by charges of intrusion into the patient-physician relationship, harm to quality, and rationing. If healthcare reform is to make the transition successfully, it must include the necessary principles, methods and resources.

The basic problem is that practice variations are a statistical observation of aggregate behavior. They are extremely important in telling us that some practices are probably overused, they motivate us to search for wasteful practices, and they encourage us to try to sort those practices that are ineffective or harmful from those that are essential. Unfortunately, they do not tell us which practice is which. They do not even assure us that the average use of a practice is correct.

Studies of inappropriate care do tell us the opinions of a panel of experts about which specific indications are inappropriate. Unfortunately, there are often other panels that have different opinions. And even if all the experts agreed, the individual practitioners who are using the practices that are labeled "inappropriate" by the experts (and by their patients, and their lawyers, and some reporters), will disagree.

It is not as though the breakfast conversation tomorrow morning will be, "Honey, you know all those unnecessary and inappropriate practices I've been doing for the last 20 years. I'll have to stop that when healthcare reform is passed. We'll switch to pasta on Thursdays." *The unfortunate fact is that medical practices are not color-coded by their appropriateness. Determinations of just what is appropriate and what is not are the result of difficult deliberations based on confusing and incomplete information and involving agonizing value judgments. For every practice someone would call "wasteful and inappropriate, " there are advocates who consider it essential.*

Examples appear in the news every day. The most recent is mammography screening in women younger than age 50. It was left out of the President's original proposal for a very good reason. A blue ribbon panel, this one convened by the government itself, reviewed all the evidence and concluded that despite decades of trials there was no evidence it provides any benefit to women in that age group. Given the chance of a false-positive result, with all the trauma and anxiety that causes, and given its high expense, that service would clearly be considered a very poor use of resources in any study of variations or inappropriate care. I assure you that if the practitioners in Boston screened younger women, but the practitioners

in New Haven did not, there would be no observable difference in mortality rates. Yet as you know very well, there are very strong constituents who argue otherwise. To balance the genders of my examples, I will point out that the debate over screening for prostate cancer (with Prostatic Specific Antigen) fits the same description, with the exception that there is virtually no evidence of effectiveness at all for this test, and it has a much higher probability of being falsely positive.

My point is not to open debates about these particular services, but to emphasize that *one person's "waste" is another person's "essential care."* Very defensible attempts to reallocate specialists and hospital beds, and to curtail inappropriate practices will be fiercely countered with charges that quality will suffer. They will argue, correctly in most cases, that the practice offers some benefit—hope or peace of mind, if nothing else—and that the only reason to exclude it from coverage is cost. It will be called "rationing." Whether or not quality actually *will* suffer, only Mother Nature knows, but we *do* know that it will *appear* to suffer, that the decision not to use it will have been influenced by its cost, and it *will* be called rationing by a lot of people. *If healthcare reform is to succeed, it must contain the principles, methods and resources for assessing the value of individual practices, for formally incorporating costs in decisions about practices, and for setting priorities among practices.*

Determining which practices are inappropriate, and defending those decisions, will require two things: good evidence about the benefits, harms and costs of the practice in question, and a judgment that its benefits are not worth its costs. Unfortunately, we face very serious problems with both steps. First, the available evidence for determining whether a technology has any benefit, much less the magnitude of its benefit, is extremely poor. This statement might seem surprising. On the one hand, we are flooded with information about medical technologies. A search of the literature on an important technology will identify hundreds, if not thousands, of articles On the other hand, exceedingly few if any of these articles actually provide clear answers to the questions that are clinically and economically important. *The reasons range from poor selection of topics, to poor research designs, poor analysis, and poor interpretation.* The result is gross gaps in the available evidence, and frank errors in the conclusions that are reached.

One example of a gap in information will illustrate this problem. In a recent analysis colleagues of mine and I did of treatments for unexplained infertility, the most obvious piece of information that a family and their physician would want to know is the chance that a particular popular treatment program would result in what the obstetricians call a "take-home baby." Imagine a husband and wife sitting with their physician, listening to a recommendation for a treatment that could take years, that has some risks, and that costs a lot of money, and asking, "Doctor, if we undergo this treatment, what is the chance that it will be successful in giving us a healthy baby?" The sad fact is, as obvious and important as this question is, and although dozens of studies have been conducted on this subject, there was no way to answer the question with any precision. This is true for virtually any treatment strategy currently in use for unexplained infertility. To be sure, when a physician is asked that question, he or she will give some answer. But that answer will come from their *beliefs*, not from solid evidence. Perhaps worse, different physicians will have different beliefs, make different recommendations, and give different answers, and there is no way to tell who is correct. The implications for informed consent, consensus development, and expert testimony, not to mention the quality and efficiency of care, are obvious.

The state of information about the costs of technologies is just as bad. Although it is relatively easy to determine the direct or immediate costs of a drug or a device, it is very difficult to determine the cost of a procedure that involves a package of services, such as the treatment of stage I breast cancer. It is even more difficult to determine the long-term financial effects of a technology.

But poor information is only the first of the problems we face. A second and in many ways more difficult one is that *even if we did know the benefits and costs of a technology, as a society we are quite schizophrenic about how we deal with that information.* On the one hand, the public cry is clear: "Healthcare costs must be contained!" On the other hand, another public cry is equally clear: "Thou shalt not consider costs when making decisions about medical technologies!" Even in a sophisticated audience, an honest and sympathetic attempt to discuss the need to balance cost vs quality can be immediately silenced by the charge, "But that's rationing." *The result* of this schizophrenia *is that we are quite unprepared to deal with this most central issue of healthcare reform.*

To appreciate the precise nature of the new demands for difficult judgments that healthcare reform will create, it is necessary to understand that for decades costs have always been considered implicitly in medical decisionmaking. The practice of medicine has always required drawing lines that imply comparisons between the ad-

ditional benefit of a practice to its additional cost. Examples are decisions about how long a patient should be kept in an ICU (which balances the probability of an emergency event vs the high cost of the bed), how frequently a screening test should be given (it is virtually always possible to squeeze out a bit more benefit by screening at a higher frequency but at a higher cost), the appropriate indications of a diagnostic test or treatment (we don't give CT scans to everyone with a headache, even though there is a small chance of finding a tumor). Even years ago when costs were not an issue, it would have been absolutely impossible to provide every service that could potentially provide any benefit. We have always understood that subconsciously, and were comfortable using words such as "practical" and "prudent" to draw the necessary lines. *The difference under reform will be that those decisions, which until now have been done so silently that even physicians hardly appreciated the tradeoffs they were making, will now have to be done more explicitly and with ever increasing deference to costs.* Those are two enormous changes. Although they don't introduce a new concept, they make an old one that was previously kept under the table highly visible, and they shift the perception from a "practical use of resources" to "rationing." *Without a coherent policy on how decisions about medical technologies should be made in the new era, they will be made in an inconsistent, unfair, and inefficient way that both harms quality and fails to control costs.* Decisions will be determined more by the technical appeal of particular technologies, or social views about the individuals who receive them, than by the true effect of the technology on the quality of care.

The debate about mammography screening for women younger than age 50 is also an excellent example of our ambivalence toward costs. There are as many negative trials (indicating that screened women might actually be *harmed*) as positive ones, and the combined results of all the trials show no effectiveness or benefit. Although there is a *chance* that it has benefit, that chance is very small, the best evidence is that it doesn't, and there is no question but that the resources could be used much better elsewhere. This example is about as close as we ever get to identifying a waste of resources. Despite this, the technology has a strong psychological and social appeal. It addresses a female disease, a cancer, and prevention. Even if the best evidence is that it has no effect, it *could* have benefit, and that small chance is difficult to deny. If mammography for women younger than age 50 is added to the benefit package, a social desire to seek out a potential benefit, no matter how small, will have squarely overruled the need to control costs. The New York Times reports Secretary Shalala as assuring the public that "costs will not be considered" when she and First Lady Clinton reassess this and other technologies that are important to women.

What needs to be done to correct these problems? Unfortunately, there are no easy solutions. *Solving these problems will require major changes in our perceptions, our expectations and our practices.* They will require a long time and a lot of compromise.

With respect to research, we must modify the quality of clinical research: what research is done, their designs, their interpretations, and the use of the results. Notice that I stress the quality of research, not the quantity. We already spend an enormous amount of money on clinical research. In addition to the billions of dollars spent at the federal level, a large amount of additional time is spent by investigators and practitioners on less formal research, funded from other sources or from cost-shifting. Unfortunately, most of this effort is simply wasted—it has little actual value for making clinical decisions. The evidence for this harsh conclusion is that despite thousands of medical journals publishing tens of thousands of articles every year, we still don't have the information needed to answer the most important clinical questions. Solving this problem will require more than additional money. *It will require a reform of the intellectual underpinnings of clinical and evaluation research—beginning with medical education and proceeding to address the incentives for research, the selection of topics, the quality of research designs, the quality of publications, and the ability of policymakers and practitioners to use the results properly.* I can provide more specific recommendations if desired.

With respect to improving our ability to make the difficult decisions that require tradeoffs between quality and cost, the greatest single need is for national leadership to resolve our schizophrenia. First, we must decide explicitly and finally whether we really want to hold the growth of healthcare costs to the CPI or not. If we do, then we should make that clear and stop saying that costs should be ignored when making clinical decisions. We should also be prepared to make some compromises and tradeoffs—yes, to take steps that will be called "rationing." If we do not choose to constrain healthcare costs, then we should admit that, and adjust to the consequences of having the nation's budget for health inexorably drain resources away from other social programs.

If we do choose to control healthcare costs, then we must agree on the principles and procedures that will be used to help us make the difficult decisions that will be required to achieve that objective. In my opinion, *in order to take advantage of the research on practice variations and appropriateness, in order to find the services that are truly inappropriate, and in order to successfully reduce their use and defend those decisions, healthcare reform should contain the following elements,* many of which are already in the President's plan.

1. *Sooner or later, the increase in healthcare costs must be kept to the CPI.* Even if we do not know the precise proportion of the GNP that should be spent on health care, we do know that healthcare costs cannot continue indefinitely to increase at a rate that exceeds the rate of inflation. If that were to occur, then regardless of what the "correct" percent of the GNP is, healthcare costs would eventually exceed that threshold, and would eventually wipe out other necessary programs.

2. *We do need a formal, global limit or cap on the allowable increase in costs.* If competition by itself can succeed in holding costs to the CPI, the formal limit will never need to be activated, and discussions over whether it was needed will have been settled by the marketplace. Today's debates about the pros and cons of market forces vs central controls will become moot. On the other hand, if competition fails to hold costs in line, the arguments caps are not needed because competition alone could do the job, will also have been settled by the marketplace. Again, the debate will be moot. We should keep the cap in the plan.

3. I believe *the caps described in the President's plan are structured properly.* The President's proposal is essentially to cap the total costs incurred by plans, adjusted for changes in population. To avoid micromanagement and the distortions that causes, the cap should be applied at as high a level as possible. A cap on per capita premiums covers all the services an individual might need, and avoids having to specify particular services, volumes, or prices. Those issues are left to plans and providers to determine. This is as it should be; the decisions that tradeoff quality vs cost must be made by those who have control over and are held responsible for both quality and cost. A cap at the level of plans will force plans to allocate their resources to maximize quality while staying within a budget—which is just the right incentive. A cap will not make plans happy and will be very difficult to comply with, but that pain will have to be accepted in order to contain costs.

4. In distinction to the first three recommendations, which are already in the Administration's proposal, the next item is not. To provide the leadership the country desperately needs to make the agonizing decisions, to incorporate costs in decisions and guidelines about medical activities, and to set priorities, I recommend that *the legislation include the following observations and principles:*

1. *The overall objective is to maximize the total benefit provided to the population being served, within the constraints of the available resources.*
2. *There are limits to the resources we can spend on health care.*
3. *Because resources are limited, it will not be possible to provide everyone with every service that might have some benefit.*
4. *Because resources are limited, priorities will have to be set.*
5. *When setting priorities among services, it is both appropriate and necessary to consider the costs of the services. It is appropriate to deny coverage for some services because of their costs.*
6. *The priority of a service should be determined by the amount of benefit it provides, for the resources it requires (the ratio of its benefits to its costs).*
7. Because there will be vigorous debates about whether particular services provide benefit, and how much benefit they provide, two "subprinciples" are needed:
   a. *To the greatest extent possible, estimates of benefits, harms and costs should be based on empirical evidence.*
   b. *When empirical evidence contradicts subjective judgments, empirical evidence takes priority.*
8. *Before any treatment is recommended for use, it should satisfy four criteria.*
   a. *There should be convincing evidence that the treatment is effective in improving health outcomes.*
   b. *The benefits of the treatment must be judged to outweigh its harms.*
   c. *The net benefit of the treatment (its benefits minus its harms) must be judged to be worth the financial cost of the treatment.*
   d. *Allocating resources to the treatment should serve the objective stated in the first principle, which is to maximize the health of the population served, subject to the available resources.*
9. *When making judgments about benefits, harms and costs, to the greatest extent possible the judgments should reflect the collective preferences of the people being served (eg, the members of a health plan).*

In conclusion, I believe that it *is* possible to achieve the three main goals of healthcare reform: universal access, and better quality, at a cost that stays within the CPI. However, achieving those goals will not happen unless the reform contains aggressive steps to improve the information base for medicine, and to help make the necessary tradeoffs between quality and cost.

○