

**Michael:** So please get yourself comfortable. People generally pick up, in terms of the materials, background materials, everything except the evaluation, and I'd really like you to take the opportunity to pick up an evaluation form and use it. This seminar is put on by the California Research Bureau and it is funded by the California Health Care Foundation.

It's the second one in a series. The first one was health information technology and electronic medical records. The third one will be a seminar on individual mandates. First I'd like to acknowledge the staff at CRB, the acting director, Charlene Simmons, who retires next week, had her retirement party yesterday afternoon.

And then we had Pam Morasada, who did all of this work, and it was a phenomenal amount. Then we had our colleagues, Jennie Putafut and Brian Salla participate as well, and then the editor, Carly Pulley. Then we'd like to thank Lucian Wilson and Adam Doherty, they prepared the background paper that's part of the materials that we have.

Then we have everybody's PowerPoints, and if you would like a copy of the PowerPoints, electronic, I can email it to you, so you would be able to have it that way. And so today

is health technology assessment, and we have three gentlemen who are well established within their peers in terms of their knowledge, expertise and participation in the discussions on health technology assessment.

And today what we're going to do is we will have Cliff Goodman, who's from the Lewin Group, is going to both be the presenter, and he's also going to be the moderator. And we have asked each of the speakers to take 30 to 35 minutes to make the presentation, and then after that we will have a roundtable, which Cliff will also moderate.

Then we'll have question and answers for anybody at that point. So we're looking forward to probably getting out of here pretty close to two. It's a little longer than most, but this is, since health care technology plays a significant role in the health care cost, the rate of growth of health care cost, this is one we want to spend more time on than we have the previous one.

Again, Cliff is senior vice president at the Lewin Group, and he's going to take over and moderate this, and he's going to ask the speakers to discuss this, and then take us through into the roundtable and into the question and answers.

**Cliff Goodman:** Thank you very much, Michael, very glad to be here. We're going to talk about something that you may not realize is one of the most important trends in the health care field, whether at an international level, national level or state level. I would say that if I were to organize a health technology assessment seminar anywhere in the world, and you asked me the top five or ten people I'd want to have speaking, two of them are here today, on your left.

So I think we're all very fortunate that we've been able to attract such fine speakers here, and I'll try to hold up my end of the bargain. I call my presentation "Health Technology Assessment, a Policymaker Starter Kit. I know that some of you are somewhat familiar with the field, but I'd like to give you a little baseline or grounding in it so you'll have the working vocabulary for it.

I'd like to point out some of the trends and issues that will be addressed by the subsequent speakers. First off, what do we mean by health care technology? It's not just hardware, it's not just gizmos, it's not just devices. As you can see, we're going to describe this in terms of the physical nature of clinical purpose and stage of diffusion.

Technology really is the practical application of knowledge in many ways. Let's look at physical nature first. When we talk about technology, we're talking about drugs of all types, biologics, which are things like blood and vaccines, and now these biotechnology-derived products, which by the way are of utmost importance here in the state of California, because you have in this state two of the world's leading biotech companies, definitely the leaders in the field, and important economically as well as from a biomedical research standpoint here in the state.

Device equipment and supplies, medical and surgical procedures, support systems. A clinical laboratory is a technology. A drug formulary is a technology. Even a patient's record system is a technology, a group of technologies. And finally, we even use organizational delivery managerial systems. A payment system is a form of applied knowledge, is a form of technology.

You can ask, "Is this thing safe, is it effective, is it cost-effective?" And so forth. So when we speak of technology, we mean it in the broadest sense. But what's interesting is that there's a unified approach to this. The kinds of ways

that we inquire about technology, its intended and unintended consequences, that form of inquiry really does apply across this broad spectrum.

Well, clinical purpose. I told you about what it is from a physical nature standpoint. When we talk about technology, it can be applied in many ways. To prevent, prevention, to screen for something, to look for it in asymptomatic populations, to diagnose someone with symptoms, to treat, to rehabilitate and to palliate. Palliation refers typically to the best care for people, to make them comfortable, often towards the end of life.

So, think about it across physical nature, clinical purpose, and stage of diffusion. Think of the life cycle of the technology. When you know what about a technology is going to depend on where it is in its lifecycle. So, future, it could be someone's mind's eye in drawings. Experimental is laboratory animal testing, typically.

Investigational usually means clinical studies, clinical means in folks, in people. Established, mainstream, standard approach, or perhaps obsolete. So, think of that sort of lifecycle as well. Now, one of the reasons the three of us are

here, and I bet a lot of you might be interested in the field, is that technology of course shows extraordinary promise, but we have all been burned, and I maintain a list, and this is my subset list of technologies that were determined to be ineffective or harmful after they were broadly diffused.

Now, not all of these technologies are bad for all indications. Some are good for some indications, but not others. But good gosh, autologous bone marrow transplantation with [inaudible] chemotherapy for breast cancer in women. Cox-2 inhibitors, you've heard about Vioxx recently, haven't you, the last couple years?

Drug-eluting coronary artery stents. Good in some cases, maybe not so good in others. Electronic fetal monitoring during labor without access to fetal scalp sampling, episiotomies, lit birth. There's a whole list of these things that basically got out of the bag too early before we really understood how well they worked.

Now, something does get out of the bag, and we need to track it and monitor it with what we call post-marketing surveillance to pick up these adverse effects. But listen, if you're in this field, you've got to consider the balance of

protecting the health of the public and providing timely access to innovative technology that may extend, save lives or improve quality of life.

By the way, the one at the bottom is the one that really triggered a major change in this country, as well as Western Europe, was thalidomide for sedation of pregnant women circa 1959/1960, in the UK, Germany and the United States. That was an important event that actually changed the way the Food and Drug Administration.

In any case, there's one list. Now, on the flipside, there are some good technologies out there that are even known to be cost-effective, and they are underused. What I suggest to people in this field is that not only are you responsible for protecting against the potentially harmful, you need to understand why we don't use the good stuff.

And there are many factors, and you, as policymakers need to think about, "What are the factors that limit patient access?" Is it geography, is it [inaudible], is it culture, is it payment? There are so many reasons why we don't use what we need to use, and lives are lost as a result of not doing this well.

Well, what is health technology assessment? This is my own definition that I put together that I consider to be the five or so main underpinnings of it. What are they? First of all, it involves a systematic evaluation of properties, effects or other impacts of technology. Second, the main purpose is to inform policymaking.

It doesn't constitute policy in and of itself, but it informs policymaking. It may address direct and intended consequences as well as indirect and unintended consequences. When we design a technology, a drug, some molecule, we think, "This is what we want it to do." And until the last few decades, we weren't very good about thinking about what might it do that we hadn't intended?

Now, what's fascinating is that some of the unintended consequences of technology are favorable, pleasant surprises. They allow us to expand indications and use it for different things. Quite interesting. It is conducted by interdisciplinary groups, not just a bunch of doctors, not just a bunch of economists by themselves, not just a bunch of epidemiologists by themselves.

Typically it's an interdisciplinary form of inquiry, and it uses explicit analytical frameworks and a variety of methods.

Those are the main elements of health technology assessment. Now, what about these properties and impacts assessed? What is it that we ask about in technology assessment?

Well, for one, technical properties. When you plug it in, does the light go on? In a diagnostic test or an x-ray or CT scan, what's the resolution? What's the accuracy? What's the sensitivity specificity? What's the positive and negative predictive value? How well does it function technically?

Not the same as, "Is it safe?" What's the chance of risk in various people? Not the same as efficacy and effectiveness, which is basically, "How well does it do what you meant for it to do?" Efficacy and effectiveness, by the way, are different, and we made that distinction very clearly in this field.

Think about it. If you hear about a new technology at Stanford, done by a leading physician who developed a new procedure, and surrounded by top staff and laboratory equipment and all this with carefully hand-selected patients,

it'll probably work pretty well compared to how well it's going to work in the community, when it's not at Stanford, not being run by the physicians who designed it, not with carefully selected patients, not with 24-hour, round the clock attention.

That's effectiveness, difference. You've got to ask both questions. Cost and other economic attributes, and there are a variety of them, I'll show them to you briefly later. Social, legal, ethical or political impacts, all these other things that flow from technology. Any technology assessment may address one, some or all of those types of impacts or properties.

Now, a little bit of history for you, and I'm going to tell you something I bet nobody in the room knows, maybe not even Alan or Jed, I'm not sure. This is sort of a timeline, I don't want to get into detail about related concepts, health technology assessment grew up in the 1970s. In the '80s we started saying, "Gosh, you know, politically focusing on a technology didn't go over very well with the doctors and the pharmaceutical companies and devices."

They said, "You're picking on us and our stuff." So we said, "Let's focus on what these things do to people, that should be the entry point," which was about outcomes effectiveness research. Pharmaco-economics got big. The first publication that used the term was in 1989, and that was applying economics to pharmacy.

That's become very big now because the pharmaceutical industry's putting a lot of money into helping people decide or to demonstrate value, i.e. economic value of their molecules. Evidence-based medicine took hold in 1990 with some colleagues of ours, David Eddy and others. And now comparative effectiveness research today is really hot nationally.

Alan's going to talk about that, I can't think of a better person to address it. You know I'm from the Washington, D.C. area, this is very hot in the Obama Administration, there's a lot of activity in this area. Now, the thing I'm going to tell you that nobody in this room knows, I bet, is my first point, because you're policymakers:

Technology assessment was first spoken in the U.S. House of Representatives by Congressman Amelio D'Addario from

New York. He wasn't talking about health care at all. Congressman D'Addario was, at the time, and those of you that were growing up in the '60s know this, when we first started thinking about, "What is DDT doing to our environment? What are supersonic transports doing to the air, quality of the air?"

It was all about how technology and life was having unintended consequences, and we first started thinking about this in an analytical way then. And health technology assessment was not done formally until the early to mid-1970s. So that's a little fact of the day. Technology assessment was a political term first spoken in a Congress.

Well, Alan's going to talk about comparative effectiveness research. The question going on today is, "Well, in the form of a drug let's say, we know how well it works against placebo because it wouldn't have gotten FDA approval to be in the market, but we don't use placebo, we use some other drug. Tell me how well the new drug works against the standard of care, the old drug.

"I want head-to-head comparisons, not just something new versus a placebo. How well does this work in clinical

practice?" Remember, effectiveness, comparative effectiveness research, not comparative efficacy research. Alan's is going to talk to you about that. It's very important these days. There's a lot of money and stakeholders lined up on either side of this issue.

A lot of the pharma medical device companies aren't very excited about this unless it's done well. And this can be done not well, believe me. So, the wrong ways of doing this that can lead to improper answers, but in general most stakeholders think it's a good idea. It's just how well will it be implemented?

HT is not just performed by any one organization or stakeholder group. It can be done for a variety of purposes by a variety of groups. So it may be to advise a regulatory agency, such as the FDA, about whether to approve something or not. It can be used to advise payers or health plans. Jed Weissberg is a key person at Kaiser.

He is in that particular role, among others. Both Alan and Jed are involved in advising clinicians and patients about the appropriate use of a technology. So that's a function of it, or role. It could help managers of hospitals and other health

care organizations make decisions about acquiring, “Are we going to get a new PET scanner this year?”

“Do we need a new clinical laboratory? How about a new medical record system?” Support decisions by companies. Companies do HTA. When they’re developing a new product, they’re saying, they have to make go/no-go decisions, and determine return on investment of products in the pipeline.

You better know they do their form of technology assessment. And supporting decisions by financial groups about investment. I, several times a year or more, get calls from investor capitalist types saying, “Cliff, we got this new device. It’s an Israeli stent. We’re thinking about investing in it. What do you know?”

So they care about it too. All right. Well, three main groups of methods, and we can spend a semester talking about this, and we won’t be able to do it, obviously. But just at the very high level, we’re looking at different kinds of information for supporting health technology assessment. The first one is primary data collection.

This is basically getting data in clinical trials, typically. What do we see in experiments or quasi-experiments? The best form of that is in RCT, randomized control to trial.

Secondary or integrative analyses, you may have heard “systematic review” or “meta-analysis.” These are ways of taking data from published or existing primary data studies, and combining them or synthesizing them or integrating and say, “What can we conclude across the evidence from multiple studies?”

And I’ve made economic analysis a third category where economic analysis weighs the costs and benefits, including outcomes or other results. And it’s interesting. Economic analyses can draw on primary studies, they can draw on meta-analyses and strategic review. So they’re not quite three distinct forms of information gathering, but there is some overlap.

Economic analyses are becoming better developed in drawing multiple methods, including, by the way, highly sophisticated mathematical modeling. You’ve heard about evidence-based medicine. I’m showing you what we call an evidence hierarchy. Here’s the deal: Until the last decade or

two, if Dr. So and So from Mayo Clinic said that he thought this procedure was swell and published an article of a series of 50 patients, that might've been enough.

It's not enough anymore, because a case series of patients is subject to various forms of bias that can lead you to the wrong conclusion about whether or not this thing works. So we're pushing up on the evidence hierarchy towards wanting that higher form of evidence near the top. That's important, and that's been good, I believe, for all of us, because there's a higher burden of proof and a higher standard.

I won't go into detail about this. This is what we call a clinical pathway or a causal pathway. It's actually kind of cool. When we say, "Is there good evidence for something working?" There might be multiple steps involved. I'll just tell you briefly how this one works; it's adapted from a highly evidence-oriented group called the U.S. Preventive Services Taskforce.

They look at screening of studies and try to look at the evidence for those. What they've said is, "Well, you've got a population at risk, and they get some screening tests," like

maybe a cholesterol test. And you want to say, "Is this good or not?" Well, it depends on what you mean. It might be very good at early detection of a target condition, and number one might give you some good evidence there, and by the way you might have some adverse effects.

But it might not stop there. You might say, "Well, does the early detection of that target condition, like maybe high cholesterol, lead you to make a choice about an alternative treatment, A versus B?" Maybe diet and exercise versus a statin, who knows? And then that alternative treatment might have an intermediate outcome.

It might knock down your cholesterol levels. Might do that. It might also have an adverse effect. But knocking down your cholesterol level isn't everything that you need to know, because what are you really trying to do here? You're trying to save people's lives, you're trying to diminish the incidence of heart attack and stroke and some other things that might follow from this, and that's mortality, basically about lifespan, morbidity, how sick are you? And quality of life.

So each one of those arrows is associated with a number that says, "What's the evidence on this one?" And the neat thing about this chart, this schematic, I love it, is that you could get from the population at risk getting a screen test all the way to what happens to people in a stepwise fashion from individual studies, or the crown, best study you can have would be the one that says, "Take these people, randomize them to get the test or not, follow them for life and find out what happens at the end, in one great big study."

Easy for me to say, hard to do. Many years, quite expensive. So, sometimes we have to make choices by doing the #7 kind of study or piecing it together that way. There are advantages and disadvantages either way. When you hear about things like cost-effectiveness analysis or cost analyses, various types of economic analyses, that terminology is often misused, and there's a lot of sloppiness in the terminology here.

There are, even at a high level, about a half a dozen different kinds of ways to do some kind of cost analysis, and Alan and Jed both are quite familiar with these and have done a lot of work in the area. One basically says, "What's the cost

of something? How much does colorectal cancer cost the State of California every year? How much?"

That'll tell you how big this problem is, but it's not asking about what happens to the people. Cost minimization analysis says, "I've got A versus B. There's nothing to tell me in some situations that A is any better or worse than B. If I can assume that they're equally effective and I'm a pair, I choose to pay the lowest price. That's cost minimization analysis," often used, called reference pricing around the world.

Cost-effectiveness is starting to weigh cost versus some natural units. Dollars per averted heart attack. Dollars per averted case of cervical cancer. Dollars per extending life by one year. There's something called cost consequence analysis. Cost utility's a very interesting one because it's cost per unit of patient utility from improved status of life.

Why is that interesting? That's interesting because it's not disease-specific. I can measure utility as zero, I'm dead, one, I'm perfect health, whether it's heart disease, diabetes or a bad knee, and we say, "What's our investment here in health

care? For how the patient senses utility of their improved or changed health.”

Cost-benefit analysis is actually the oldest one. It actually started in defense. Robert McNamara, of all people, actually used it in his informative analysis in the '60s. The tough thing about cost-benefit analysis is that it's not only how much you're spending for what you get out, but what you get out has to be quantified in dollar terms.

So if somebody extends life by a year, how much is that worth to you here in the state of California, and does everybody's extended year of life get valued the same way? Budget impact analysis, used in a few ways. One says, “You've got a fixed budget. What's the most efficient way to spend those resources on any given set of choices?”

Okay, so the point is to not memorize this. The point is to realize when you hear cost analysis or cost effectiveness analysis, there's a bunch of ways to consider how to do this. We have things called ICERs, which are Incremental Cost-Effectiveness Ratios. We say, “Well, how do you make a comparison? What's the difference in cost between some

new intervention and a comparator, divided by the difference in their effectiveness?”

That’s basically what it is, not that complicated. And as you can see, these are the kinds of answers you get. The last time I used this slide, I was in Europe, and I think I, as you can see, I’ve got the euro sign there for you. You need to think internationally here. When you’re on the Pacific Rim, you’ve got to think about a bunch of currencies, although euro’s probably not one, although euros aren’t bad these days, right? Okay.

Quickly now, when I want to think about cost effectiveness analysis, I try to simplify it, at least for me, in two dimensions. It kind of goes like this: Here’s my current technology. It has a certain level of effectiveness, and it has a certain cost. This is my comparator; this is where I am now. Then the question becomes where does the new technology fall on this grid? How does it compare?

Well, maybe like this. Let’s say I’ve got a new technology here. It is more effective and less costly. That’s an easy decision; I’ll adopt that one. But if it’s up here, it’s what? Less effective and more costly, you’re thinking to reject it.

Where life gets interesting, where these two fellows make some of their living is in these quadrants.

It is more effective, but it costs more. That's what happens with a lot of new drugs and biotech devices, and even sometimes it's less effective, but it costs less, we might want to think about that too. That's a decision as well. And then you start saying to yourself, "Well, what really happens is," and this happens a lot in health care these days, is it's just a little bit more effective, but a lot more costly, how much are you willing to spend for that?

And if it's a lot more effective, but a little more costly, you probably want to go with it. But these are tough choices, and can be done at the policy level. Think about time horizon. Over what period of time are you making these decisions? Short-term, medium-term, long-term, does it matter? Well, heck yes, it matters.

Think about this: Here's today, zero. Ten years, 20 years. When you start investing in health care intervention, unless it's removing an inflamed appendix right now, where you feel the benefit that day, if you're investing in smoking prevention or lowering the incidence of heart disease, you

may invest today and not feel the benefit for one year, five years, ten years, 20 years, 40 years.

So the time horizon will definitely affect how you perceive the balance of what, the money you're putting in with the benefit you're getting out. And if a technology assessment doesn't think about an appropriate policy time horizon for a clinical decision-maker or policymaker, may get the wrong answer.

This happens more often than you want to think. This is one of my favorite concepts here, I wish we had more time to talk about it: You might hear from the subsequent speakers the term "quality-adjusted life year." An important concept, ever more important these days. Here's how it works:

Here's my patient today. This patient's not doing so well, expected lifespan is five years. On a scale of zero, death, to one, perfect health, this person today is at 0.8, and there are ways of measuring this. So, the course of this person's decline is, as you see, bordered at the blue. Five years of life, quality of life is going to go down this way.

Now enter a new treatment. How might we assess how well this treatment works? Well, two ways. One, it might lengthen life, in this case by a year, and for any given year the patient's going to have a better quality of life. This is your investment decision. How much are you going to pay for the area in yellow?

How much are you going to pay for that? Is it worth it to you? What's the opportunity cost? You want to put your money someplace else or on this guy? So we make decisions about investment in health care, we're saying this is a good way to present how you're making a decision, what your investment choices are.

And you will hear things like "cost per quality-adjusted life year." How much money do I want to spend for one unit in that yellow? Here's where it got real controversial: In 1991, a colleague, Alan Maynard from the UK, and this actually has some historical significance when Alan published this, because it got a lot of people up in arms.

Alan went in the literature; he found all he could about cost per quality for various procedures. He had to do some adjustment standardization, and he put them into this thing

he called a league table. In this league table, he ranked them from what? Least cost to buy one quality-adjusted life year at the top, to, at the bottom, greatest cost to buy one quality-adjusted life year, one square unit of that yellow.

The equivalent, if you will, of one year of perfect health, although it might be spread out over multiple years. And what Alan said was this: If you have a fixed budget, and you're responsible for your population, think the population of California, think about the MediCal population, by the way this happened just up north in Oregon in the '90s, same thing, very close to it, I should say, and you don't have enough money to spend on all the procedures needed by everybody in the beneficiary population.

What do you have to do? Maybe make some choices. How do you make those choices? Well, if this were your complete menu, imagine if you will, of choices, and you couldn't pay for all these things for everybody that needed them, what would your decision paradigm be? Get in a lot of fights about this one.

And what Alan said was, an economist will tell you, "Well, what's the most efficient investment of your limited

resources?" You get that by saying the amount of money I have, how do I buy the most quallies. Remember? Because quallies are good across all kinds of conditions. So I'll just jump to the answer, and the answer was you start at the top, and you keep spending your money on the people that need the first one.

And then when you have money left, you go to the second one, and then you go to the third one, and you keep doing down the list until you run out of money. And if you're below the cut line, sorry, you're out. And a policymaker could say, "We use the state's money as efficiently as possible. We bought more quality-adjusted life years for our beneficiaries this way than any other way. Unfortunately, it may leave some of these people out in the cold."

You're policymakers, I'm sure you could deal with that. And by the way, there are paradigms for, how much money do you want to spend for a quality-adjusted life year? And some people say that, "Well, depending upon your economy," you know, here in the United States we often say, "Somewhere between \$50,000 and \$100,000 per QALY."

Basically, at \$100,000, if you've got anything in this area here to the right and below of red, that's a good investment. But maybe you don't have such a good economy; maybe you are a middle-income country someplace. Maybe for them they can't afford \$100,000 per QALY, maybe they'll go \$50,000 per QALY, or \$20,000 per QALY, or less.

In the UK there's a rough gray area sort of discussion about, in the UK, between 20,000 and 30,000 British pounds, above that you're starting to say, "This may not be cost-effective," and that's taken into account in policy decisions. Current trends in HT, and I'll finish the last couple slides here. I have zipped through stuff that you spend years studying.

I'm not going to apologize for it, but I will understand that you don't get it all right away, but here's some more.

Current trends. For one thing, there's greater demand for HT to support all these kinds of decisions we talked about.

There's more demand for this kind of inquiry. Number two, there's more government and private sector groups doing this stuff.

Here in the State of California, some nationally known, actually internationally known outfits are doing this stuff.

There's more transparent, systematic and consultative HT. In other words, more stakeholders involved in providing input. Pharma companies, medical device companies, organized medicine.

Instead of fighting this process, they're more involved with it. Higher standards of evidence, we talked about that. More evidence from real-world practice, like effectiveness and comparative effectiveness research. More emphasis on cost-effectiveness, and we're getting better at doing that.

More use of systematic reviews and integrative analyses. And this one, I like eight a lot. Number eight, more interest in tailoring evidence requirements and methods to particular kinds of technologies. It's not like there's one way to do these things for all kinds of technologies. If you're looking at a technology for patients who otherwise have no choices for life-threatening conditions, you may weigh cost and benefits a little bit differently, or if you're looking at what we call orphan diseases, where there's so few patients with a condition that you can't get enough to enroll in a clinical trial that's going to generate a statistically significant finding

or has the power to do that, I should say, you've got to maybe change your approaches a bit.

More closing the evidence gap by linking payment to new evidence generation. Sometimes, instead of an insurance company, the payer saying, "We're going to pay," or not pay, they'll say, "We're going to pay sometimes, but you need to collect data so we can figure out more." More specificity in HT findings.

Through subgroup analyses and other cases, there's more analytical breakdown. Much greater international collaboration in methods. Everyone knows what everybody's doing through the internet, international access. If somebody in France publishes an HGA, we all know about it in 22 seconds.

So there's a global network with regard to this stuff. Horizon scanning. What's coming over the horizon? Can we get ready for it before we get hit by it? Interesting. Rapid assessments, can you do this faster? More efforts to coordinate and align and harmonize regulatory decisions and clinical decisions and payment decisions.

And industry's more aware of this stuff and interested in HTA as opposed to opposing it. That was full-speed HTA for you. [Applause] Thank you. I'm still standing to prove it. I believe Alan's next.

Alan Garber, well, you already heard a bit of intro. Alan really is one of the nations and the world's leading authorities in this field. He's published widely in cost-effectiveness analysis. He's right at the center of the controversy and very knowledgeable about comparative effectiveness research, and I hope that my intro will help Alan kind of pursue this discussion. Alan?

**Alan Garber:**

Cliff just gave you a fabulous overview of health technology assessment, something that we normally teach over a course of days, you get in a half-hour. I'm actually going to reinforce a fair bit of what he said, but I would like to emphasize a few points that I think are particularly important and timely.

First, I want to settle the context by asking the question, basically, why is Washington so interested in comparative effectiveness now? Jed's going to be talking a little bit about questions and issues for California specifically, and I hope

that our discussion will focus a lot on, what about California? What kind of approach do we take in the state in this area?

But I would like to talk for a moment about Washington. Some of you may be comforted by this, some of you may be even more upset, because California is not the only government entity in the United States that has a budget problem. [Laughter] And this is a figure I like to show a lot.

It's a set of projections of how much we're going to be spending on Medicare as a percentage of our gross domestic product over the coming years. The source of these estimates is the Office of the Actuary and the Centers for Medicare and Medicaid Services. They've consistently been excessively conservative in their projections of growth expenditures.

Here, conservative means that they have consistently underestimated what the Medicare program will cost. What this figure shows you is, this top line part here is projected expenditures. Historically, and then the projected amount. The other parts of this figure are the sources of revenue. The bottom here, payroll taxes is what you see on your W-2

forms as the Medicare tax, that's percentage of payroll paid for the employer and the employee.

There's a very small tax on benefits, and then there are premium payments. So, I think probably none of you are collecting Medicare right now, but you have to pay a premium to get Medicare part B, the physician and ancillary service component, and you have to pay a premium to get the drug benefit.

So that's a source of revenue beneficiaries' pay. Then there are the state transfers back and forth between the state Medicaid programs, and then general revenue transfers. General revenue transfers refers to the unfunded part of Medicare that is due to the difference between what part B and part D cost, and what people pay in premiums, premiums covering maybe 20-25% of the cost of those components of Medicare.

So that's unfunded in the sense that it doesn't have a designated source of revenue. Now, there's this odd thing here called HI deficit. Do people know what HI stands for in the Medicare context? Hospital Insurance. In Medicare part A, when you pay the payroll tax, it goes into something

called the Hospital Insurance Trust Fund, which you might rightly believe is an accounting fiction.

But anyway, it has money in it now, and it's projected not to have any money in it by around 2018. And we're spending more now on Medicare part A payments that are coming in in payroll taxes. So that's this deficit, which will grow when the, even more when this trust fund disappears. The bottom line is this area here, this orange area at the top and the blue area here do not have designated sources of funding, and think of that as the budget gap.

Now, the reason these expenditures are growing are an interaction between, of course, the rise in the number of Medicare beneficiaries, particularly as baby boomers age, plus rises in the costs of technology, etc., which means that the average cost per beneficiary is anticipated to rise.

So this unfunded liability, by 2050, according to these projections, will be about six percent of gross domestic product. I'm not talking about health care percentage GDP or Medicare percentage GDP. This is only the fraction of Medicare that we don't have a source of funding for if we continue with current rules.

That will be in 2006 dollars, 2007 dollars. \$7,600 per working age adult. The reason I'm expressing this figure in those terms is it gives you an idea of what kind of tax you might need to cover this unfunded liability. That is several hundred dollars more than the current per capita personal income tax.

So in other words, think about doing the equivalent of more than doubling the personal income tax to cover this. The fact that it's called general revenues indicates it just means however the federal government gets money. But probably this would be a good little benchmark to use.

So, did I give you a moment of relief from California's budget problems? Anyway, we're all going to be hit by this as well as California's budget problems. This is going to hit us a little bit later, perhaps. But this has got to be in the backs of the minds of legislators in Washington, and certainly the Obama Administration as it contemplates what's going to need to be done health reform.

So, comparative effectiveness research is something that was sort of bleakly and mildly supported by candidate McCain, and very much supported, as Cliff mentioned, by President

Elect Obama. It seems like a common sense notion that you should find out what works. Cliff gave you what I think is the single best definition of what this is from the Agency for Health Care Research and Quality.

Cliff's colleagues, the Lewin Group, I'm sorry for this, but this is the name of the report, so... They came up with the estimate that over ten years you would save \$368 billion at the national level from establishing a center for medical effectiveness, which really means the national or the federal agency that would be sponsoring and conducting comparative effectiveness research.

Michael has given you copies of the CBO, the Congressional Budget Office report on comparative effectiveness. And for those of you who have the interest, I strongly urge you to read that report. It's a really very thoughtful and sophisticated piece of work there, continuing to look at what the savings might be.

This is their first foray into analyzing comparative effectiveness, but they did a terrific job. So it's probably the single best background source. So, comparative effectiveness research inevitably will involve the array of activities Cliff

described. These include just summarizing what's already in the literature, but really the more interesting question, the reason that legislators in D.C. are talking about so much money, putting so much money into it, is they really want to get new evidence, new data in one form or another.

So the question is if we make the investment, what will we gain from it? And that is completely tied up in the question of how we apply the information. I just listed a set of questions here, and this is... Incidentally, you should interrupt me at any time with questions or comments, but I would like to have further discussion during our panel discussion.

But these are what I think are key issues, and they are certainly subjects to debate in D.C. What should be the structure and funding of the agency that's responsible for the research? Cliff has very intimate knowledge of what happened with the Congressional Office of Technology Assessment, which was defunded in 1994/95, somewhere in there, in part because it was viewed as an impediment to what some interest groups wanted to accomplish.

The Agency For Health Care Research and Quality used to be called The Agency For Health Care Policy and Research, and had a near-death experience. You could say they were reincarnated perhaps. But they had a near-death experience in part because they had conducted research that offended some very powerful interest groups, particularly some back surgeons in Texas.

So, the issue of structure and funding really has to do with, how do you ensure adequate funding, and how do you ensure that they will not be punished for being too candid about what the results of the research are? Which question should be studied first? This is obviously crucial. Which questions are the most important?

How do you prioritize among the many, many questions that could be asked? What new information will be collected? I think it's inevitable that new information's going to be collected. Cliff raised some of the issues. I'll go into that a bit more. And then implementation. What in the world are you going to do with the research?

Cliff gave you an example of cost-effectiveness, which actually has much more, in some respects, natural series of

steps you might take if you're a government agency, to decide what to do with it. He described the algorithm; basically, in order to maximize value you first do the stuff that has the highest bang for the buck.

That's very obvious. Now, it may seem that it's obvious what to do with comparative effectiveness research that doesn't include cost. But I'll contend it is. And incidentally, Judd, Cliff and I have all served on the Medicare, what used to be called the Medicare Coverage Advisory Committee, now it's called the Medicare Evidence Development and Coverage Advisory Committee.

And that really does engage in a form of comparative effectiveness research, and trying to decide what Medicare should cover. You would be amazed at how much debate there can be over questions that you might think have very, very obvious answers. That is, is there enough evidence to say, "Expensive treatment A is any better than treatment B?"

And the people who are supporting comparative effectiveness in Washington say, "Don't consider costs," may have a somewhat naïve view of how easy it will be to gain consensus about some of these issues. But once you

have this evidence in place, then you're going to ask, "Is it going to affect coverage policy? Is it going to be used to help physicians make decisions, patients to make decisions, etc.?"

That's an implementation issue, and everything hangs on that. So I just want to mention, Cliff alluded to coverage with evidence development. This is something that's a set of terms that Centers For Medicare and Medicaid Services use, but there's something similar that's going on in the UK called Only In Research, OIR.

The idea is that you will pay to provide some form of care that, in some sense, is not quite experimental, but we don't know enough about how well it works in real-world populations, in other words the effectiveness question. So, you impose some kind of data collection requirement as a condition to pay for it, so that you actually learn going forward.

These are the criteria they use. They ask about the importance of the question, which diseases represent the greatest burden to Medicare beneficiaries. In other words, they don't start with the really rare stuff in general. It's stuff

that's really common, like heart disease, heart failure, diabetes, things like that.

Which diseases and their treatments are the costliest to the Medicare program? Very obvious one. What's the value of incremental information? So, this is the explanation for this criterion. Where do we have the greatest deficits of knowledge? So in other words, there are really some areas where we know very little.

And it's not simply a matter of refining an already decent knowledge base. It's a matter of really resolving whether a whole treatment approach works. So, there's an argument to be made, for example, that we may not need to learn a lot more about another statin, and in which subgroup it works to lower heart disease risk.

But there are whole treatment categories where we know very little about how they work. And I would contend a great deal of orthopedics fits into that category. So, then a question to ask is what kind of data, what kind of methods will be used, and will we really answer the questions by engaging in this?

So, Cliff has already mentioned these options. I just want to make a few points here. You can start by reviewing existing data, and this is what... There's a program, Cliff didn't name this, called The Evidence-Based Practice Center, Stanford and UCSF have one. There's one at UCLA, Ayn Rand, and the Evidence-Based Practice Centers engage in a very structured form of literature review to determine what works.

They do great work. It typically costs I think about \$200-250,000 for one of these reviews, which may sound like a lot of money to you, but once you've ever seen how these are constructed and how much time goes into it, you will see why that actually is a bargain considering what you get.

But clearly, if you're going to engage in new evidence collection, you're talking about orders of magnitude higher costs. The highest cost option is to do a randomized trial, which is of course the goal [unintelligible] randomized trial is not expensive because of the randomization, this is what's commonly known as a coin-flip trial where people are randomly assigned to get a treatment or a control.

And in the case of comparative effectiveness, it's actually going to typically be two active treatments. Something new perhaps, and standard treatment. In order to be able to determine whether the new treatment is better, when you're comparing it against something other than placebo or a sugar pill, you need many, many more people, and/or you need to follow them for a much longer time.

And of course, the implication is this is going to be considerably more expensive to conduct than a placebo-controlled trial. So this may not be done very often even though this is the study design that's most likely to provide you with definitive information. Then there are registries where you actually require specialized data collection, but it falls short of a trial.

So, Medicare has done this with implantable defibrillators. These are the devices that shock people's hearts if they are subject to something called sudden cardiac death due to an irregular heart rhythm. They get these devices implanted, and they really do work. But they're expensive, and the ongoing care of patients with these devices is expensive.

And there was some uncertainty about who really benefits, even though there's no question that there are some patients who get a substantial benefit. So, CMS said, "Okay, if we're going to pay for it, you actually, the doctor, have to fill out these forms and provide a lot of information," goes into a registry, and this will be used to determine things like complication rates.

Then there's routinely collected clinical data, which if you're with Jed and his colleagues at Kaiser Permanente and a few other very forward-looking organizations, you have very elaborate clinical data in electronic form that can be used for sophisticated studies. We are far behind where we'd hope to be in terms of the general health care system, and gathering data in this [more].

And then there are claims files and other administrative databases that allow clinical detail. The key question when you use anything other than a randomized trial, is when you see that a certain treatment approach is associated with better outcomes, is this just something that, does it really reflect the fact that the treatment is improving people's lives?

Or is it the case that the people who get the treatment will do better anyway? There's something about who's selected for treatment. And this was an example, you probably didn't get time to see it in Cliff's long list, but there was a belief several years ago, some of you may remember this, that high-dose chemotherapy or bone marrow transplantation for breast cancer was a life-saver for women with advanced breast cancer.

And in fact, many oncologists believed it was unethical to do a randomized trial. Nevertheless, after overcoming a huge amount of opposition and great obstacles, there were three randomized trials that finally were conducted, and they revealed that women who got this procedure were no better off, and by some measures were worse off.

Well, the women, when you looked at the non-randomized trials, women who got bone marrow transplantation or high-dose chemotherapy did do better, and there were very nice registry studies. And of course, the reason, if any of you are familiar with the procedure, is you would not give this procedure to a woman who is very sick in any other way other than her breast cancer, because it was so toxic and so

demanding that only relatively healthy patients could actually successfully go through the treatment.

So, we were misled, and that's why we always need to ask this question. These are some of the problems you have with observational data, and I'm harping on this because if you were running the agency that conducts this kind of research [unintelligible] you're going to have a budget. You may have a \$300 million budget.

No matter what your budget is, you will say it is too small, with considerable justification. And when you have that budget, you're going to decide, "Am I going to blow a third of my annual budget on one clinical trial, or am I going to do a lot of observational studies that are subject to this kind of scientific uncertainty?"

The selection effects are what I just described. And then I want to point out that there are some ways, novel ways to get out of this box of the choice between the very expensive, randomized trial and use of observational data. All three of us have been involved in various kinds of studies of this kind, where you look at a variation across the country and across states, across practices in how things are done, and

see if you can develop techniques to determine whether one treatment approach is better than another.

So we can talk later about these. These are methodological terms that I'm sure are not of great interest to you. But I just want to point out that this is an area of active research. There are some compromises in effect that we might adopt methodologically to get good results, that is results we can believe at relatively low cost.

Now, I don't want to take too much of your time, but I do want to get to what I see as a central issue. In the Bacchus legislation in Washington, which is leading legislation right now, I think it's viewed as having a pretty good chance of passage, would that be correct, Lou? There is a line there that basically says, "Let's not consider the costs."

An agency that conducts this research must not consider cost. In other words, the cost-effectiveness of research, cost-effectiveness of the interventions. So, here's a question: Should it ignore costs? I want to draw your attention back to Medicare because more than anything else, I believe that's what's really driving congressional interest in this subject.

This is from a study led by Dan Goldman, I was a collaborator, but Dana is the chief of health economics at Rand. It's from something called the Future Medicare Project, and he convened groups of experts to talk about what new technologies might be developed over the next 10-25 years.

Then they costed out what they thought these would cost, what the total cost would be, the per-patient cost, and then what would be the net effect on expenditures in the Medicare program. So, some of these interventions that could be developed actually will reduce costs down the line even though they cost something upfront.

So, getting back to the example of statins, the cholesterol-lowering drugs, in people with heart disease, actually they do lower costs, to treat somebody with a statin who has high cholesterol. So, one thing, this is actually a technology that exists today. Left ventricular assist device, or LVADs.

They're used in people with severe congestive heart failure, extremely severe. They cost a lot. The cost of the operation and the device combined is about \$80,000, but to keep someone alive like in the trials that were done of this device

originally, they're so sick, they spend a lot of time in the ICU, even if they get the device.

And they cost... We did a run of some commercial health insurance claims, and came up with about \$500,000 per year. So it's not that the device costs that much, it's all this other stuff. And so you end up with estimates here of an annual treatment cost of on the order of \$10 billion, and \$14 billion in 2030 for actually treatment of a small number of people, and this is a cost-effective assessment, not using qualities, but life years, about a half million.

Contrast that with anti-aging compound. How many people have heard of resveratrol? One person. If I remind you, I think a lot of the rest of you will... This is a compound that's found in red wine, and if you drink a couple of cases a night, you'll get as much as the lab rats did. So there are actually companies who are trying to develop analogs of resveratrol.

It's interesting, I'll just say very briefly, I don't know if you remember the details of this, but if you starve an animal, starve a lab animal to near starvation level, so they don't die, they actually live longer and they're less likely to get chronic

diseases. Sounds like the perfect health maintenance activity, doesn't it?

They're pretty cranky too. Anyway, if you feed the animal a regular diet so they're really happy, and give them the resveratrol, they live just as long as the starved animals. So, all of us want to line up and get some of this stuff, right? Well, the assumption here is if this is developed, everybody will get it when they hit age 65, they'll get this compound.

So you actually spend quite a bit. That's on the order of \$50 billion and \$73 billion. And it'll increase spending because everybody's getting it, but the cost per additional life year is really small. And part of the reason this works out is it extends life at a relatively low cost, but also it keeps people healthy.

So the idea is you get less of the arthritis, the diabetes, the other degenerative diseases that go with aging, which is what the experiments in lab animals suggest. What cost-effectiveness analysis does for you is it creates a set of incentives. That is, if you have cost-effectiveness analysis as a basis for determining which interventions get covered,

you're going to shift from things like LVADs to things like resveratrol by a cost-effectiveness criterion.

So I want to just go through an example to try to persuade you, and you'll see this is very similar to what Cliff showed you. But as Cliff noted, when you're evaluating an intervention, you look at the change in cost. Does it increase them or decrease them? And this includes not just the cost of the intervention, but downstream costs and downstream savings.

You prevent disease, there's going to be a negative cost in the future, and you look at the health benefit, ordinarily measured in quallies. So he was pointing out these quadrants here. You spend more and get less, you don't need high-priced consultants to figure out you don't want that. Spend less, get more, also you don't need the high-priced consultants.

You do want us for these two quadrants, and I'll concentrate on "spend more, get more," is it worth it? Here you're going to look at stuff that changes costs a lot, big increase in costs, not much health benefit, not cost-effective, you don't want to

do it. Down here, the opposite is true. These are the measures.

You look at the incremental cost-effectiveness ratio which Cliff mentioned, and I want to just give you a quick example to show you why so many of us view this as important even if you're a little uncomfortable with the idea of incorporating costs in our decision-making. This is an interesting study that was published in 2003.

Remember all the controversy about the Cox-2 inhibitors, that was Vioxx, Celebrex? And they came on the market, and they were to replace NSAIDS; non-steroidal anti-inflammatory drugs, like I use the example, Naproxen here, that would be Advil, Motrin, the anti-inflammatories that are in the same category as aspirin.

So they looked at the use of these to manage chronic arthritis, and in the base case, and incidentally when this was published there were some rumblings about the possible adverse heart effects of using these drugs, but they were only rumblings, it wasn't why we accepted. So they first excluded any adverse effects on the heart.

They looked at how much more expensive was the Naproxen, and the benefit, the benefit is you don't get GI bleeding, bleeding in the gut, which is the risk that these are designed to mitigate, and the incremental cost-effectiveness ratio is nearly \$300,000 quality-adjusted life year. And Cliff mentioned, the figures we would tend to use, I'm pretty comfortable, I think most people are up to \$70,000 per QALY.

But by almost any standard, this is not a good deal. Once you take into account the possibility that there's an adverse effect on the heart, today we would use different data and say it's much worse than what we see here. Then it looks even worse, of course, \$400,000 per quality-adjusted life year, because you're mitigating the benefit by having an adverse effect.

If, however, and I want you to remember this, if you remember anything today about cost-effectiveness analysis, it helps you to figure out how to target interventions. Not everybody's the same. High-risk patients is the group of people for whom this set of drugs was originally approved. People who, when they took an ENSED like Knapperson

before, had a GI bleed, especially one requiring hospitalization.

So they might drop their blood counts, have to be transfused. Some people can die from this. So if you focus on the people who've had one of these adverse effects before in an ENSED, and give them a Cox-2 inhibitor, the gain is much larger than for the average patient with arthritis, and now it becomes cost-effective at \$56,000 per quality-adjusted life year.

So this technique can help you to target it. So I want to just point out a couple other things about cost-effectiveness. You may think that it's easy to know if one treatment's better than another, but especially if its effects are complex, think back to Cliff's slide about evaluating a screening test or preventive intervention, but particularly a screening test.

There are many steps in the chain of logic that goes from test result to changing health outcomes. You cannot usually do this with a randomized trial. He mentioned the possibility of doing a study where you actually look at that entire chain all at once. You randomize people, you get a test or not.

I think the number of examples of that is on the order of five, maybe ten that have ever been done. It's really, really hard, really expensive to do that kind of a trial. So you almost always have to end up doing some kind of modeling, and that's a big part of what cost-effectiveness does. And plain comparative effectiveness, as envisioned in the Bacchus Bill, simply doesn't address cost, and it's not obvious that it's going to reduce costs.

So I just want to point out a few things that may be relevant to California. The agency that conducts the research may not be an arm of the payer, like Medicare or a private health plan, and probably should not be. We can talk about that during our discussion period. It can be used for a lot of purposes, but one thing that's relevant to California is when you use this kind of research, you can go back and say, "You know, you're not really delivering that much extra value; we're not going to pay that much."

It turns out Medicare can't really do it, they can't do that negotiation. They're constrained very much by a bunch of legal considerations having to do with legislation that created it. But in almost every other setting, including other

countries, it is used to negotiate prices. So, thank you. I hope that you'll think a little more positively about the role of cost, and maybe we can talk about that in our discussion.

**Cliff Goodman:** Alan is affiliated with the Veteran's Affairs of Palo Alto Health Care System. He's also with the Center For Primary Care and Outcomes Research in the Center For Health Policy at Stanford. Speaking next is Jed. Jed, do you have the...?

**Jed Weissberg:** Yes.

**Cliff Goodman:** Jed Weissberg is associate executive director for quality and performance improvement at the Permanente Federation. Jed, the floor is yours, sir.

**Jed Weissberg:** Thanks. You may not realize it right off the bat, but we just had this incredible survey of the field, and I met Adam last night, he's in the MPH program, you get academic credit just for being here today. [Laughs] You too. So I'm going to talk about Kaiser Permanente's approach to technology assessment, and you tell me afterwards how well we embody the approach and concepts and definition that Cliff described, and the broader implications that Alan was challenging us with.

So I'm going to talk a little about how we're structured, what we define as our scope, and then we're going to go into a current example to show what we do when we think the evidence isn't quite sufficient to derive a particular conclusion, and how we deal with that difficulty. Here's our scope: We talk about new technology as meaning something brand new, a new device that's never been used before in patients, or a new application of an existing device.

So that might be taking the kind of tools you use for a knee arthroscopy, say, and all of a sudden doing something new in the shoulder that's never been done with tools like that before. Now, that's the kind of thing that the FDA doesn't get involved with, and we'll talk about where the FDA falls short in guiding citizens and our health care system in what's safe and effective care.

We have other parts of our organization and other committees that look at drugs and biologics, because the concept of pharmaceutical formulary started years ago, way before technology assessment, and we have a lot of experience there. It's a very important area, though, because I think Cliff talked about, were you talking about Genentech

and Amgen, perhaps, as those two big companies in California?

The biologic molecules, all these monoclonal antibodies and the things that patients give themselves by injection or infusion, over the past ten years, for Kaiser Permanente's California members, spending on those agents has gone from \$70 million to \$700 million, and it's following the kind of exponential curve that, if it keeps going, just consumes all the money there is, and then we don't have any money for anything else.

So these kinds of pressures are all around us. And we very much try to take that definition of technology assessment that Cliff gave us, and apply a systematic approach and framework to looking at these new devices and approaches. Kaiser Permanente is largest in California, two-thirds of our members are here, but we're also in Hawaii and Oregon and Washington and Ohio and the mid-Atlantic states and Georgia, and we try to take people from the health plan, from the medical groups, bioethicists, legal, government relations people, media people, and get them together four times a year in a room to consider about eight topics a

meeting, and think about how we can garner the evidence that exists, and summarize it in such a way that it's helpful and useful both for our clinicians and the health plan that tries to manage benefits and costs.

We also have the additional benefit over a lot of private technology assessment companies that once we gather the relevant evidence, we can vet it and bounce it off a number of Permanente clinicians who have personal experience with the technologies in question. And we often get insights that you just wouldn't gather from reading the dry literature itself.

In addition to our committee, we actually support what we call our technology inquiry line. It's an 800 number within our organization, and anybody within Kaiser Permanente, any of our employees can call and say, "I just read about this new thing, what is that?" Or they can say, "You know, a patient is asking me about this technology, and I never even heard of it."

Happens a lot in the internet age. "What is this about, and do we have an evidence summary?" We field about 400 of those calls a year, and generally turn around a request if it's

something we've looked at previously, within a day or two, and inform our delivery system or our benefits people.

If it requires a new evaluation, we get it done within about a month or so. And in doing that, we utilize our own internal resources, we have all kinds of Master's-prepared people doing technology reviews and assessment, and we also have contracts with private technology assessment firms.

You may be familiar with a company called Eckry, the Eckry Institute or HAZE, these are the private companies. But in addition, just like Cliff said, you go to the internet, you find the Canadian agencies, the British agencies, the Australian or New Zealand agencies, the Cochrane Collaboration, there's huge amounts of material out there for the asking.

So this is our defined scope, and I want to point out that we actually don't do cost-effectiveness studies. The reason being is that we're working with a health plan, and the health plan is always subject to suspicion. Are they being objective in their assessment of a technology? Are they withholding beneficial technology to save money?

So what we've chosen to do is to summarize the evidence around medical appropriateness and effectiveness, and then hand that information over to the people who operate our delivery system and make the projections about what health policy premiums will be if we incorporate such technologies in the future.

That gives us a degree of insulation against potential accusations of bias that we think make our work more objective, more useful, and also allow us to collaborate with other tech assessment companies in the world, and other insurance companies for that matter, because we're not using this evidence in a direct linkage to a coverage decision.

Now, back in the 1980s we didn't really do so much of this. It was more what Cliff described, it used to be called "eminence-based medicine," as opposed to evidence-based medicine. But we had some experience. We had a court case in which we were assessed \$40 million in damages because for several years after in vitro fertilization became an accepted mainstream intervention, we in our own health plan in Kaiser were saying it was still experimental and denying that service.

So, we learned our lesson from that. And then we had a case in a region we used to have in Texas, not the most HMO-friendly part of the world, and this was one of the reasons why. You can see this, the parents of twins with a very, very serious congenital liver disease were asking for liver transplants for both of their children.

And this was at a point in the development of the medical science where adult liver transplantation was new. Kids were really just being investigated at the time, and there was really no experience with this particular disease. And as you see, we lost that appeal, paid for the transplants, and sadly the kids ended up dying.

However, the experience of those two cases prompted us to get much more organized around the development and training of evidence-based medicine for our clinicians and supporting them with all kinds of resources. We hired an eminent person, practitioner of evidence-based medicine, a guy named David Eddy to train us and help us develop this approach, and it's now spread throughout our program.

We also joined with Cliff and Alan as the co-scientific chair of this Blue Cross/Blue Shield Association Technology

Evaluation Center, which is I think the foremost proponent of a standardized, very, very rigorous approach to technology assessment in the country. I'll show you the language that we use when we finish an assessment now.

We developed it around 2004, and it's as good as you can get, and you tell me whether you think it suffices to answer the huge policy issues that Alan pointed out are facing us. I mentioned that the FDA is out there; people think the FDA is enough. We make the case that it does not nearly address all the issues of importance to the people trying to operate the delivery system and practice medicine.

We pointed out that we're looking in the FDA at substantial equivalents to a prior device, that lets a manufacturer slip a new thing through pretty easily, or in a drug trial perhaps, a comparison to a placebo, a medicine that doesn't have active effects. Whereas what we're really interested in in the delivery system is comparison to whatever we're currently using to treat that particular condition.

That's the increment we're really interested in. And in addition, we're interested in longer-term health outcomes. And I'm going to get to an example in a couple of minutes,

and we'll see just how long-term the studies actually report findings for a condition that afflicts people for the rest of their lives.

So we get topics from all over the place. Our clinicians are interacting with their patients, they're going to their own scientific meetings, everybody's reading the journals, we have people dedicated to the kind of horizon scanning that we heard Cliff mention, and we see what the other technology assessment companies are thinking about considering and decide how much interest we have in that topic.

And we put these together, and we're having another meeting in a couple of weeks. Let me see. I'm presenting the topic on a non-invasive assessment of how scarred your liver is. Because previously, what we do is we stick a needle in there and pull out some tissue, it's a technique with potential harm and pain, and if you can do that non-invasively and get an accurate answer, that would be a huge advantage.

So we're going to review the evidence for those things, and then decide whether we'd recommend the incorporation of those new approaches into our delivery system, thus

removing the liver biopsy from the delivery system. We use multiple sources, like I say. We don't do all of this ourselves.

We think that that increases our scope and efficiency, and it also increases our credibility because we're not the ones making all of our own insular, internal decisions on things. Speaking of David Eddy, he wrote in an article 10, 15 years ago, a very simple series of aphorisms and advice.

If you have evidence of benefit, by all means do it. If you have evidence of no benefit or harm, actively try to stop doing it, and perhaps counter the detailing or the marketing that might be happening on behalf of a drug or a device. And when there's insufficient evidence, simply be conservative. Use discretion and try to collect data as you investigate the new technology.

So if we look at our experience in assessing devices, technologies, procedures over the past couple of years we have, in some, found that there's sufficient evidence to say, "Don't do this." In some we found sufficient evidence to say, "This is an appropriate approach for this defined population. Do it in a highly competent way with good oversight and quality assurance and go out there."

The difficulty is that so many conditions and practices fall in this chasm. What do we call this? The chasm of insufficiency, where there's simply no evidence, and it's astounding, but true, that things are being marketed and are approved for use, and you can't find a smidgen of evidence about how they actually work or who they work for.

But more often there is some evidence, it's simply of insufficient quantity or quality, or it points you in different directions. Some studies say one thing, some studies say another. We're going to look at an example, I'm going to give you a little brief background in the technology, and then we'll see what the studies say.

So, if a condition is to treat vertebral fractures, your spinal column or your vertebrae, and mostly due to osteoporosis, lots of people, more women than men, as they get older, will suffer from osteoporosis and vertebral collapse. And you can see that it's the, this is the front of the patient over here, this is the back.

So the vertebra collapses there, and this is why you see older women in particular getting smaller and hunching over, it's due to this vertebral deformity from these fractures. And

these fractures can also be associated with quite a bit of pain. So they cause all kinds of difficulties in our patients.

One of the approaches is called kiphoplasty, which attempts to both remove pain and restore the height of the vertebra, so you don't get that collapse and forward kind of stoop. The other approach is just called a vertebroplasty, which I'll show you in a second, which is, only attempt is to relieve the pain and prevent further problems.

So here's an x-ray. This is an x-ray of the spine, and this is a big, nasty-looking needle, this is only done under anesthesia, so you don't have to wince too much right now, where you actually stick the needle right into the body of the vertebra, and inject a semi-fluid bone cement that hardens and gives off heat as it hardens, it's just like curing cement in your backyard.

And then you end up with a vertebral body that's strengthened by the application and installation of this cement. Now, nobody really knows the mechanism for why this works. It may be that it actually restores mechanical strength and prevents these little micro-fractures from moving against one another.

It may be that the heat generated through the curing process just simply destroys the nerves in the area, so it's sort of an existential Zen question, it hurts, but you can't feel it, so you're relieved of your pain. The difference with kiphoplasty, a variant, is that you try to restore the height of the vertebra as well.

So you put a balloon in there, you expand it, and then into that space you've created, then you inject the cement, and it hardens in a larger space. It seems more physiological. So, okay, we've got hundreds of thousands of women, fewer numbers of men, and younger people if they have trauma or a tree hits them on the head, suffering these vertebral fractures, suffering severe pain.

When do you perform a technique like this, and on whom, and which one do you choose? Let's take a look at what the evidence is, this is what it looks like after the kiphoplasty. This is a little kind of summary of reviews that we've devised in Kaiser Permanente. We call it the "at a glance report," and it gives you a sense of how much duplication and inefficiency there is in the realm of technology assessment.

So there's a Midwest organization called IXE that looked at these techniques, and they said they thought the evidence was pretty good for vertebroplasty, not so good for kiphoplasty. HAZE and ECRI are two companies we contract with, we pay for them. The license doesn't allow me to show you what they said about it, so I can't say, but it just shows you that both of these companies are considering this issue.

And then the group that Cliff and Alan and I have worked with, the tech looked at it, published the results in 2004, said, "Uh-uh, not enough good evidence to say this is an appropriate intervention," and just looked at it again recently this year, and we can perhaps talk about it, Alan, about why we thought the evidence was insufficient.

And we've looked at it on our Interregional New Technology Committee within Kaiser Permanente recently. So, here's the evidence. You have a patient, has a vertebral fracture. You're thinking, "Should I do a vertebroplasty or a kiphoplasty?" There is one trial in the literature comparing these two interventions.

It's a non-randomized trial. The patients were observed for eight weeks and treated conservatively with bed rest and pain medication primarily, and then they were enrolled in the trial. Pain scores in both groups improved. Disability ratings, and there are some standardized versions of that, were observed only for the kiphoplasty patients in the short term, but when they went back and looked at those patients two years later, there was no difference between the groups on this measure.

The study, which was in an orthopedics journal, didn't do a formal statistical comparison of the results. However, the authors concluded the kiphoplasty was superior and presented that information to their colleagues. Now, vertebroplasty has a little more evidence associated with it.

I'm not going to go through this in detail, but in the appendix that was handed out on the hard copy, you also see what we call an evidence table where we summarize the studies and all the relevant features of the studies, the number of patients, the length of follow-up, how the health effects were assayed and what they really were over a period of time.

But suffice it to say that there were two control trials, one with 34 patients and one with 79 patients. One trial looked at 24 hours, one day later, and at 12 weeks post-op. This is for a condition that basically afflicts people for the rest of their lives. And one trial looked at a two-week follow-up.

And then there were published case series. One highly expert group, as Cliff mentioned, at a Center of Excellence, was investigating a new technique, picked patients, did the procedure, and reported on the results in those patients. Kiphoplasty, the other of these kinds of interventions, has even a smaller group of patients.

60, 36 patients, and published case series. These were non-randomized trials. The interesting thing about this trial refers to something called selection bias. So the patients who were enrolled in this trial were first offered kiphoplasty. If they declined kiphoplasty, then they got medical management, and then those were the two groups that were compared at the end of the period of time, and perhaps Alan can tell us about the statistical issues in trial design that make this suspect.

So if we want to just summarize, we see that this is the evidence for vertebroplasty, this is the evidence for kiphoplasty, and this is the evidence for choosing between the two in an eligible patient. And again, we're making perhaps policy decisions for thousands and thousands of patients on the basis of experience with a few hundred.

It's true that in all the studies, there was a large decrease in pain by one of these visual analog scales for pain. However, there were multiple sources for study bias that are listed here, including the fact that the people who manufacture the bone cement or the instruments used were funding these trials, and it always raises a caution flag about how the data was generated and followed and analyzed.

What I didn't show you were the net harms. We were talking about the benefits here. And there are adverse effects of these things. That's a big, scary-looking needle, and you're jamming this stuff in a bone. The spinal cord is right there, lots of blood vessels are right there. Those blood vessels go right back to the heart, into the lungs, and there's potential harm that turns out to be actual harm, and infection and bleeding and having to remove extravasated

bone cement have occurred, you might need neurosurgery after this kind of thing.

And interestingly, you saw that the bone is reinforced. You can put a little cement in there, it's like filling in a cinder block, but what's not really known is whether that then makes the adjacent vertebrae more prone to fracture because all of a sudden you've stuck this really hard thing in the middle of them. And our patients in these trials haven't been followed long enough to conclusively answer that question.

And in addition, the FDA's been a little concerned about this bone cement because there have been a number of reports of severe patient harm, which I'll show you. So they sent out an alert. Cliff, you'll have to help me. The MAUDE database is maintained by the FDA, I think this is Manufacturers and Users Device Experience, a voluntary database that people report to if something really egregious happens.

With kiphoplasty, the one where you balloon up and then inject, there have been deaths, cement leakage, with and without long-term damage. The balloon can rupture, and sometimes you just have to leave it in there, sometimes you have to operate to remove it. Other cases where the blood

vessels to the lungs are blocked, people get these reactions to the cement, some sort of allergic thing when the blood pressure drops and you can have a heart attack.

You sort of wonder whether there's a thorough informed consent in all the patients going through these things. When you look for vertebroplasty in the MAUDE database, there were far fewer records. It seems on the basis of voluntarily reported incidents to the FDA that vertebroplasty is safer.

So, that's the evidence. Now I'll show you from a health plan point of view, IMR, if you're involved with health plan coverage issues, you know that IMR is independent medical review. So if Aetna, Cigna, United, Kaiser Health Plan have a member who wants something that their physicians don't think is appropriate, they can appeal, and the appeal can then, through a process, go out to an independent medical review.

And we looked at our own database of appeals, and we found a couple. We said, "Oh, one of the appeals supported the member," it said, "These two techniques have been shown to be very effective for immediate and lasting pain

relief. Ten years, excellent results.” That ended up getting paid for and done.

Another independent review, proven place in the management of compression vertebral fractures. Okay. The third one, the clinical efficacy of these two techniques has not been established, unclear whether this provides benefit. So what’s a poor health plan to do? They’re getting this conflicting information, and one of the issues is how evidence-based medicine and technology assessments are applied to both the policy and individual patient settings.

So, Michael challenged us to say, “You guys are the current and future policymakers for the state. What should the state be doing about this?” So I looked at the picture of the California bear for a couple of minutes, and I said that, “You should be supporting the practice of tech assessment,” and that applies to things, whether it’s embodied in legislation like a mandate, you should ask what does the evidence show about this mandate that might have been brought to your attention by an advocacy group?

We should be trying to apply the results of tech assessments to our publicly funded programs, and there’s great

precedent for this in Medicaid and MediCal with the use of drugs and formularies, so that is something that's starting to be done in the country. And mainly we want to avoid these kind of conflicting messages to people, to health plans, to clinicians saying, "One day you can get one decision, one day you can get another decision without any difference in the underlying data."

And lastly, and perhaps most importantly, we need to make the public aware that things are not as clear as they might seem when you read the glossy odd in the Time Magazine for a drug or a procedure, or the hip that Jack Nicklaus has when he got his hip replaced. And we need to foster a hunger for evidence, so that people are demanding that these evidence be produced.

That's going to make our jobs much, much easier when we try to make good policy decisions. So, going back to our information and advice from David Eddy says, "If there's insufficient evidence, be conservative," I'm going to show you that we have a interregional group of spine surgeons, highly fellowship-trained, specialized folks in the Permanente Medical Groups who came together, who all

had their own opinions about these things, from their training, from reading an article or two here and there.

And we've gotten them together, and we presented the data to them in a very organized fashion, and then we showed them that in Kaiser Permanente the cost of these procedures varies by about 35, 40% with kiphoplasty being more expensive. And we've asked them, "Based on this, what do you want to do?"

So we reviewed this stuff thoroughly, we agreed that there was very little evidence to choose one or the other. We also showed them the fact that both of these techniques were being utilized currently within Kaiser Permanente, and interestingly enough, another political and cultural dimension, spine surgeons do one of these procedures, and interventional radiologists do another procedure.

So there's one of these turf things going on too. And we wanted to see how they would influence the practice patterns of their own groups and their colleagues by examining all of this data together. They're in the midst of doing that now. I suspect we will see a decline in the number of kiphoblasties being done.

Ideally we would be collecting very detailed data on all of these patients. Sadly, however, that actually takes money, and we have to make choices about where to take member premium dollars and invest, and this is the kind of investment in learning and future refinement of practice that doesn't have enough money to support it, and why we'd like some public health [to that].

There's an astonishing pace of advancement. We use both the available public and private resources as well as the expertise of our own clinicians, and we try to use that evidence in the decisions about deployment, dissemination and use within our own delivery system. And I think that's all I'll say for now. Thank you. [Applause]

**Cliff Goodman:** Thank you very much, Jed. Although we started late, because our speakers were so concise, we're back on time, literally within, at the minute, which is I think unprecedented in any panel on which I have served, that's splendid. As shown in the agenda, we'll have a brief sort of panel discussion among the three of us here, and then maybe when we're done pontificating, we'll have some discussion Q&A from you.

If I might, I wanted to kind of pick up on what Jed was saying insofar as what can be done here in the State of California. Would you go so far as to advocate that here in the State of California, given the budget crisis and all these other pressures, that there be a push for really doing the kind of thing that Alan talked about, which is really specific cost-effectiveness analyses as part of tech assessment for the purpose of state decision-making? Any thoughts?

**Jed Weissberg:** I would. I think it's based on both our desire to do what's beneficial for patients, as well as the realization, whether it's implicit or explicit, that there is bounded budget dollars available for medical care within the state, and if we don't wish to see those exponential curves and see medical spending squeezing out everything else of benefit to society, we're going to have to bring those considerations to the fore, have the public discussions that make them credible and justified.

But I'd be curious to see whether Alan thinks that's going to happen.

**Cliff Goodman:** I'm very curious, Alan, and I'll kind of double down on the bet here, the state immediately to the north, if my geography

serves, had a very difficult experience with something like this in the early '90s with the Oregon Medicaid Program, and I think perhaps both you and Jed know Dr. Kitzaber and have spoken with him.

What do you think about the political wisdom of venturing down that road here in 2008, 2009 in this state?

**Alan Garber:**

The Oregon Medicaid experience is unique in a lot of ways. One thing, I'm sure most of you are familiar with it, but this was, the state was applying for a Medicaid waiver, and they wanted to extend coverage to a much broader category of citizens of Oregon than just people who are eligible for Medicaid, they need to get a Medicaid waiver to do that, and they decided to use cost-effectiveness criteria to decide what would be covered and what wouldn't.

So they were taking a fixed budget, and trying to spread it more broadly. I actually think that it really is useful to look at their experience, but the way they approached it is not how almost anybody would do it, and they wouldn't do it again that way, I think. And there are elements of it that were really good, like requiring or obtaining citizen input.

But it started out as something like what you described, Cliff, where you go down this list ranked by cost-effectiveness, they ended up with something very, very different, which is neither cost-effectiveness-based, nor based on any set of principles that you could easily describe. I think that everybody who advocates the use of cost-effectiveness and actually uses it, one example that I'm sure you're all aware of is the National Health Service in England.

They tend to allow quite a bit of wiggle room for various purposes. But you have to make sure that there's not too much wiggle room. And cancer treatment is one example where it's always been exceptional in the sense that we don't ask too many questions about cost. That is one of the reasons why Genentech is, right now they are the largest biotech company, in large part because of Avast and a cancer patient, which can cost on the order of \$100,000 a year for some patients.

But the implications for California I think are where this might play; the ways it might play out in California, I think are in two very obvious areas. One is MediCal. Are any of

you directly involved with MediCal, in legislation regarding MediCal or [SCHIP]? So, you undoubtedly are facing unbelievable pressures now.

California's not alone, but between the state budget, rising unemployment rates, and this just tremendous need, we are going to have to ask how do we stretch what may be a more limited budget, yet more people who need care? So that is sort of the motivation that Oregon had, except they wanted to go beyond their Medicaid population.

California also has state mandates for covering various services. I hope I won't offend anybody, but I did look up under Knox Keane what some of the mandates are, and it looks like, shall we say the mandates were not systematically chosen? And the suggestion might be that it had a little to do with who got to the legislature.

But California has a committee that actually evaluates potential new mandates put forward by the legislature. I don't know, Hattie, are you involved with that?

[unintelligible] Yeah, well, my understanding is before they didn't directly use cost-effectiveness, but they did have to consider cost implications, and so that's another very

obvious way to use, another obvious place to use this kind of information in California.

And of course, if California does this, presumably the way it would be approached is to draw upon work that's been done by others and commission work that needs to be done to answer California's specific questions, to Jed's point.

**Cliff Goodman:** Let me just push it a little further down this road here. All of us talked about different resources nationally, and even internationally, technology assessment. I'd also call your attention to the very nice white paper done by Lucian Wilson and Madame Doherty that document some of these technology assessment organizations around the country that could be used as resources.

Can either of you imagine some entity or institute or policy advisory body here in California that could pull together information like this and actually do their own cost-effectiveness analyses for MediCal, or would you want it to draw on other sources around the country? How would you envision this being implemented, if you do favor it in the first place?

**Jed Weissberg:** One of our speakers mentioned that there's a very known and noted technology assessment group in California, and it's the California Technology Assessment Foundation, which was created by the Blue Shield Foundation when it spun off. They commission work from UCSF and General people to do the technology assessments.

It would not be impossible to further capitalize on the incredible expertise we have in our UC system, in our private universities, and Alan's group, to add onto those technology assessments that are giving us that initial estimate of the health effect and the benefit, with some of the cost information. Now, that could still be advisory, but it would be the next step in bringing such information to consciousness, which can then be used in decision-making.

**Cliff Goodman:** Alan, anything else on that?

**Alan Garber:** Yeah. I just want to ask, are any of you aware or been involved with the Blue Shield Foundation's report about what should go into a minimum package, minimum coverage, basically? I heard about the reaction to this, this is something that's worth looking at for any of you who are

going to be involved, especially you with the MediCal involvement.

This is a national issue too. The question really comes down to, "Which types of care does a health plan need to cover to be a decent health plan, one that we would all agree is sufficient?" Either in the context of Medicaid program like MediCal, or a commercial insurance. So the Blue Shield Foundation got a group of experts, physicians together to discuss what should go into this package of this minimum coverage package.

It's what the Clinton Administration tried to do when they were developing their health plan, and they ended up with something like 2,000 pages. It's really instructive because that's ultimately where this kind of stuff goes. That's the number one use, probably, even though all three of us have mentioned other uses, and we think those are important.

But this is really, from a policy point of view, the first place it goes. I just want to point out that that effort is probably what needs to be done with MediCal, but you have to look at the experience to figure out how to not make the same mistakes. I was at a meeting in Berkeley, and somebody who

had been a legislative staff at the time recalled when this package of minimum benefits was presented to the legislature, and she said everybody had trouble suppressing their laughter when they heard this.

You know why? This minimum package of benefits looked a lot like a conventional Blue Cross Blue Shield plan. A Blue Cross Blue Shield plan is expensive. A typical Blue Cross Blue Shield plan. We are not in a world where we can afford to say that the minimum benefits should equal the average of what people get today or something that's better than average.

It seems to me for California, we really do, we would benefit from having a group, and I like the ideas that Jed was talking about and Cliff was talking about in terms of how to do it. But this is something the state probably needs to do. There's a lot of talent around the state, Southern California, Northern California, a lot of expertise not only in academia.

There are organized physician groups. Jed and his colleagues, tremendous resources. And here again I think California can take the leadership in the country, because

Oregon was viewed as an aberration. California will be viewed as a bellwether.

**Jed Weissberg:** I'll add an element to that insofar as what's special about this state. It is not just the individuals and organizations about which you've heard, but California, as much as any other state in the country, has a vital interest economically in fostering innovation. Witness Amgen, witness Genentech, witness all the high-tech outfits here in the state.

And so in an environment in which all those parties are subject to a downward economy, and all the things that go with it, the state and even the manufacturers have an interest in saying, "How do we, in a well thought out, rational way, maintain and improve the health of the citizens and foster innovation, which is so important to, let's face it, the tax base of the state and employment rates?"

So what you could do in this state is, if you're going to be involved in HT or comparative effectiveness effort or whatever you might call it, is actually engage the innovators. Get them on your side; have them be part of it. Now, under well-monitored circumstances, but part of doing HTA is not ignoring or quashing innovation.

And there may be a way to do that in a cooperative way. I did see a hand about that. Did you want to comment on that?

**Female Voice:** [Unintelligible] medicine, genomics-based, DNA-based, information-based companies. What we're interested in doing at BT&H, again, in collaboration with HHS, is to do more research around the cost implications of personalized medicine on our health care system in California, in particular on our MediCal system and on our CalPIRGs population.

So what we're trying to do is actually identify funding streams, foundation funding streams to do that. Also, what we'd like to do is examine what the regulations are around personalized medicine, and finally examine what the information technology issues around data sharing... The reason I tell you this is that I'm interested to know how some of your work could potentially apply to this initiative what we want to get off the ground.

**Alan Garber:** I hope you don't have to leave soon, or give us a chance to talk about this again. Since you're looking at me, I'll lead off with an answer to that. We at the Lewin Group have been

looking at that issue. If you want to remember one thing about this, it's the following: Health technology assessment and evidence-based medicine for the most part, comparative effectiveness research, are largely population-oriented inquiries.

We want big studies. We want large RCTs, we want big observational databases. They are inherently population-oriented. Now, in parallel, there's this extraordinary interest in personalized medicine. That's the rub. Where is the convergence or maybe even the contradiction between population-oriented studies and the information needed to make individual-level patient decisions?

Now, methodologically, we've got some ways of getting at that through sub-group analyses, through modeling and other approaches. But that's where it's headed. Anytime you start talking about smaller groups of people or individuals, your evidence, the significance statistically and clinically, your evidence may look to diminish, and that's what we've got to overcome.

It's interesting you should say this, and I don't know, Jed, if you even know this yet, I think the next Medcac[?] meeting

in February is going to be on genomics, as it turns out. I just saw an announcement. So, Medicare, this is the panel on which the three of us have served, the Medicare Evidence Development Coverage Advisory Committee is going to be looking at genomics on February 25<sup>th</sup>, actually.

**Alan Garber:** They're open meetings.

**Female Voice:** The reason I have to go is we have actually a grant that's due to AHRQ tomorrow morning around personalized medicine and how the state could actually get into the conversation around personalized medicine. Otherwise, we don't have any money coming in to do research.

**Cliff Goodman:** You better get going.

**Alan Garber:** Jed and I, for the Blue Cross Blue Shield Technology Evaluation Center, this is very much a topic, and we actually, we're doing a lot of work on this too at Stanford, because a lot of the companies are right around us in the Bay Area. And there are now tests, genomic-based tests that actually can be shown to reduce costs under some circumstances.

So there's a lot of potential, and they're difficult to understand. That's why you need to study it. It's not really easy stuff.

**Female Voice:** [Unintelligible] some other foundation grant is to scope out a study, and we want to set up three workgroups, and the first workgroup is actually a macroeconomic group, where they would scope out a study, not do the economic analysis themselves, but actually scope the study, because based on just the initial research that we've done, you need experts to even put together the study to determine which diseases you're going to take a look at, what diagnostics, what therapeutics, and how are you going to apply that?

Make sure that it's independent, that kind of thing. So what we're hoping to do is get people like you, frankly, to be able to participate on a semi-gratis basis, to inform the state about these issues and this emergence of personalized medicine, which personally I see it as something that's going to completely revolutionize our health care system.

We're talking about disease-specific things in your presentation that in ten years, with personalized medicine,

our insurance-based system may be vastly different as a result of personalized medicine.

**Male Voice:** Here are two sources for you. We supported some work, the Secretary for Health and Human Services US has under him the secretary's advisory committee on genetics and health in society.

**Female Voice:** The PCAS report on personalized medicine?

**Male Voice:** No, that's a presidential one. The SAC GHS, which is a national blue ribbon panel, produced two reports. One was on pharmaco-genomics challenges and opportunities. The other one was on genetic testing, overcoming regulatory pain and other barriers. So, those are very good sources as well as the PCAS study, of course.

**Female Voice:** And I know Deloitte is also coming out in January with their national macroeconomic analysis of personalized medicine. But what we're interested in at BT&H, again, in collaboration with HHS, is something that's California-specific. But from BT&H's point of view, we're interested in fostering innovation as well because these are a lot of jobs in

California, and we think a lot of intellectual capital that obviously we want to keep in California.

**Male Voice:**

Along the lines of fostering innovation in some of those companies, large and small, like in the state, they're looking at this too. We don't have a lot of time to go into detail, but basically, as we heard today, we all like these large studies, because they provide, as I said, statistically and clinically significant findings and so forth.

But when you're looking at cost-effectiveness, as both Alan and Jed suggested, interventions may be more cost-effective, dollars per quality, dollars per life year, for narrower and narrower populations, right? So if you're a health care products company, you know that you're being scrutinized along the lines of cost-effectiveness.

You're thinking about your markets, and you've got a bit of a conundrum. You want your drug or biotechnology to be bought and used for a lot of people, but you also know that it's more effective in smaller groups, subgroups. So they're in a position where they've got to balance highly focused, targeted products that provide a lot of bang for the buck for

those people, but that becomes such a small market, is it worth it?

And it gets back into how they make pricing decisions, in fact. So, industry too is caring about the tradeoffs between population-based value and individual-based value, and it is incumbent upon them, and they are looking very much at that issue, seeing the emerging importance of personalized medicine.

**Female Voice:** The last thing I would say is I just hope you do another presentation, and I hope that there's a larger audience, and more people can hear what you have to say.

**Female Voice:** Get them to shut down across the street and we'll bring the people in.

**Male Voice:** Thank you very much. Yes, sir?

**Michael Weinberg:** My name is Michael Weinberg, and I do healthy policy work for the New America Foundation. I'm interested in, Alan, I saw you present that terrifying chart at your conference at Stanford a couple months ago. We've already been doing some health technology assessment comparative effectiveness for a while.

So I suppose that health costs could've been growing like this instead of like this. But clearly, we haven't gotten health costs under control as of yet. And not necessarily for lack of extremely good work. But if we're thinking about institutional design going forward, what are the best organizations that are doing this kind of work, and how could we perhaps, as policymaker, either fund those organizations more or give them greater authority or link them with the payment mechanisms better?

Rather than, of course the standard approach is going to be, "Let's create an Institute For Comparative Effectiveness for California and place it in some agency that doesn't make sense for it to be there, and create a whole new bureaucracy." But I guess I'm interested in what's the very best work that's being done right now, and how could we elevate that work?

**Alan Garber:** I think you really hit the nail on the head in raising the issue. You recall I said implementation is critical. I would actually contend we don't have that much the right kind of study done in the US, actually. There hasn't been enough demand for it. And by that, I mean cost-effectiveness studies that can

be translated directly into either coverage decisions or practice.

There's nothing inherent about that, that it can't be done. It's simply a matter that we haven't had a need to use it yet, or a mechanism by which to use it. I see that as what your question is really about. How do we actually put this into practice? And I'm sure that Cliff and Jed will agree with me that just about every meeting that we attend where health reform is discussed, there's agreement, at least among the health policy committee, that the number one issue is getting payment right.

And getting payment right is really what this is about too. So you could think of coverage as being a very crude form of payment reform, which say, if it's not covered, you're not going to get paid anything for it. So that's kind of an extreme. But there's the intermediate stuff. And it will take a lot of work to figure out how to redesign payment, and Kaiser has a model where this kind of work can be directly implemented.

They haven't been to, judging from Jed's description, which corresponds to what I've heard, they haven't been very

aggressive about saying, for example, and “You can’t use kiphoplasty.” But to get directly to your question, a huge amount of the effort has to be on the implementation end.

Redesigning payment, whether it’s for MediCal or for commercial health insurance or for Medicare, is incredibly complex. You can have all kinds of unintended consequences. Like you pay a little too much for one procedure or say to manage a diabetic, if you want to go up to the disease management level, which I do think is generally the right way to go.

You pay the wrong amount, and you get huge over-utilization of something or other. So I can’t tell you exactly how to do it, nor are there many private organizations that have really tackled this yet. This really has to be a government activity, I believe. It’s not going to be done by a private organization because any private organization that would have the money to do this in a detailed way would be an interested party, and their findings would likely be rejected.

**Michael Weinberg:** But there are interested parties in government as well. Your pointing out that the things that are mandated as minimum

coverage may have some connection to the people who are responsible for putting those mandates in place. That's something we know about how government functions. Even within government, how would we design an institution to attempt to isolate it from these political pressures, but not then just have it more open to sort of capture by industry?

**Alan Garber:**

Right. There's an impulse in Daschle's book, for example, my colleagues and I have had a health plan proposal that has similar features to create a government agency that would maintain accountability, be transparent with things like open meetings, have a fairly assured funding stream, and would be isolated enough, have just enough protection that people would widely view it as procedurally fair and not subject to capture by interest groups representing narrow interests.

So a year ago I think most people would've said that that's a good model. But maybe that's not the right... The model is still probably good, but talking about the Fed may not be so good. But you understand that the characteristics of the Fed are it has a guaranteed income stream because their income

comes from open market operations, not from congressional appropriations.

They have long-term staggered terms for Fed members, and they do have to report to Congress, and they have a fair bit of accountability, but not enough that a Congressman can call up Ben Bernanke and say, "I want you to cut interest rates," and expect to have any kind of favorable reaction.

So that's the sort of thing that people are thinking about.

**Cliff Goodman:** Jed, anything on that question?

**Jed Weissberg:** So, not like the SEC. [Laughter]

**Male Voice:** Social Security Administration.

**Cliff Goodman:** Let me just add something to that. What California would probably not need to do, at least very often at all, is generate new primary data. So, you don't have to set up a great big clinical trials operation here. California probably would not have to do a lot of its own systematic reviews of existing data, because there are ECRI, Blue Cross Blue Shield, Tech, Hayes and IXY and all these other groups.

AHRQ as well. DERP does it, comparative effectiveness reviews of drug classes. But what does it have to do? As I sit at the outset, technology assessment isn't the same as policymaking; it informs policymaking. So what policies pervade the state of California that are unique or special to it?

Well, whether it's MediCal or any other payer here, you have a specific set of benefits, and laws that go with that, that you say, "We will cover certain services." And that set may be unique to California, and the population of beneficiaries has certainly a unique profile. Every state is different.

And you see, "What are our populations, age groups, racial, ethnic, socioeconomic, cultural makeup, and how do we serve those needs?" And so, much of the information from things like randomized, controlled trials and systematic reviews could be readily applicable, but what you might need is a policymaking body or analytical support that translates all that evidence out there at primary and secondary levels into answering your state's policy questions for your benefits package and your beneficiaries.

Yes, your name and affiliation please?

**Trisha Wen:** Trisha Wen with the State Treasurer's Office. I also sit on the CalPERS health committee. One question we talk about a lot at CalPERS is improving quality of life, which actually drives costs down with some of the chronic diseases. So if you take a disease like diabetes, if you improve the patient's quality of life, you probably are also driving the cost down to the payer, the health plan, and ultimately the state.

So my question is with your QALY, the definition I wrote down is, "How much will you pay for a better quality of life?" Does that also take into account your cost savings by improving that quality of life?

**Alan Garber:** Yes. There's a very big literature on QALYs and the basis for assessing quality-adjusted life, and so forth. I couldn't get into that. Yes, if there are savings, of course. Remember, it's the difference in cost divided by the difference in some outcome, in that case, quality-adjusted life years. So, of course net savings are quite welcome.

Typically, the answer is positive dollars per incremental QALY. There may be instances where it's actually negative

dollars per additional QALY. Rare, but there are some circumstances of those too. Absolutely.

**Jed Weissberg:** There is a perspective, though, to keep in mind, which is that if you treat your diabetics very proactively, and according to guidelines and keep their blood sugars and lipid, blood fat levels and blood pressure according to where they should be, that's great, and you'll avert many of the diabetes-associated complications like heart attack, kidney disease, stroke, blindness, amputations, those sorts of things.

At the same time you're keeping those people well enough that they then become eligible for things like hip replacement and knee replacement and all the other things that accompany the process of aging. So, it depends how you define cost savings and what your unit of analysis is.

**Ruth Holton-Hudson:** Ruth Holton-Hudson, and I'm with the state controller's office, but I also sit on the health benefits committee for PERS, for the controller. You mentioned Oregon had tried to apply some of this. Are there any other states that are sort of actively looking at sort of using assessments, and in particular you talked about large populations.

So the PERS equivalents in other states, is there any sort of active movement, do you know, of how you use this information to really start looking at outcomes? And since we are a huge payer, what sort of role do you think PERS could play in moving this kind of agenda?

**Jed Weissberg:** PERS I think could play a huge role. I don't know about all the states, but I am a little bit familiar with the history of Massachusetts, because the academic who designed a lot of it is a friend of mine, and he was looking around for, "How can we deal with this problem?" And they didn't want to use Oregon as a model, of course.

And they couldn't get any answers fast enough. As you'll recall, when Massachusetts put in their health reform plan, their key challenge was how to get what they call minimal creditable coverage, and how to define what had to be in that plan, and to get somebody to offer a plan as low-cost as what they claimed they could offer it to.

And that required a lot of arm-twisting. And they did what almost any actuary would do as their first move, which is instead of defining the benefits package narrowly, which is what Oregon was sort of trying to do, they ended up going

with catastrophic plans, which is probably not the best thing to do for low-income people.

And they got one plan that had such a limited network, that actually did have a low-cost plan, but their network was so limited that many, many people could not get an appointment with a doctor who are covered by that plan. So it's tough. In other words, I don't think that you will find a blueprint out there about how to do that.

But we are the most populous state, and this is the place to start to do something like that.

**Female Voice:** [Unintelligible]

**Cliff Goodman:** Alan's looking at me, so I'll start. You've heard the term "NICE." It's the National Institute For Health and Clinical Excellence. NICE is the policy advisory body to the UK's National Health Service. Here are some good things about the NICE model: First, it provides analyses on what we've been calling technology assessments, as well as clinical guidelines.

And it is evidence-based, quite detailed, it is evidence-based quite specifically. It also does cost-effectiveness analyses. It

does use cost-effectiveness ratios, particularly in the form of cost per QALY. It does have a flexible and explicit threshold from between 20,000 British pounds to 30,000.

It doesn't mean that 20,000 and below is considered in the plan and will be accepted by the NHS. It doesn't mean that things 30,000 and above will not be paid for by the NHS. But if you look over time at the yes's and the no's, they line up pretty well with 30+, less likely to be brought in the service, less than 20, more likely to be brought in the service.

What NICE has done is learn over time, NICE has become, as Alan suggested, transparent, and they have, despite very rocky beginnings, developed a rather productive and largely cordial relationship with industry, not just in the UK, but around the world. And so they provide for industry to provide evidence upfront and have input to the process.

The process is largely transparent, and there's even an appeals process there. So there's some very good aspects to that model, and many countries and perhaps even states around the world are looking to NICE as not the model, but a model.

**Alan Garber:** I think that's really accurate. I was just going to add a couple of things. There are other models too, like the Australian pharmacy benefits advisory committee that are worth looking at. NICE is in a very, very different political environment than ours. There were a couple of times in its history when its life could've been ended. It had extremely strong support from the government.

It's a parliamentary form of democracy, not at all like either California or the federal government here. And I don't think if we had an entity like NICE; it would've survived under that kind of pressure. Nevertheless, there are many, many things we can learn from them. And there's a lot of effort now to coordinate what NICE does with what goes in the US.

There have been a lot of connections made. And certainly, California can learn, and again, MediCal is one place where you might directly transpose what they did to figure out how this would fit with commercial health insurance. It's a little less clear, but there's a lot that can be learned.

But again, I want to underscore, and especially for the record, if we create an entity like NICE, it will be different in

the sense that NICE actually does make what amounts to coverage decisions, even though technically it's another part of the NHS that does. It's part of the NHS, and it essentially makes coverage decisions.

If we had something like that here, not necessarily if it was part of the state government, but a broader entity, it would probably not be making coverage decisions. The uproar would just kill it at the very beginning if it were directly making coverage decisions.

**Jed Weissberg:** And many, including Alan, have written that the national identity and culture in England is different than in the United States. There's a more pervasive sense of egalitarianism than we seemingly have here, and that supports that kind of work.

**Cliff Goodman:** Furthermore, last point to this is that the UK National Health Service really is subject to a budget, fixed amount of money. Now, we at the national level and California, what have you, we try to act like there's some cap to it, but it's not a fixed budget. And if you are operating in a true fixed budget, iron clad, fixed budget, you are forced to make these choices in a more explicit fashion.

So that's one element of the different environment to which Alan referred.

**Paul Velevan:** Paul Velevan with the Employment Development Department. My question is regarding how technology assessment is used with regard to drug formularies. You said it's being used to help lower costs. Well, the model seems to be a very generic model that you're using. It suggests, then, that the data being used that's being pumped into the model is variable because you've got HMOs having different drug formularies, the VA having a different formulary, etc., etc. down the road.

How do they take into, measure cost-effectiveness? Are they basing it on their own contracts and such when the different HMOs are doing this?

**Alan Garber:** As Cliff mentioned, I work at a VA, and the VA uses federal supply schedule. I think that fundamentally, what the VA does is you take a class of drugs, and they treat everything within the class as large equivalent. They don't do that so much for psychiatric drugs, but in other drug categories they do.

And you basically figure out where you get the lowest price. It's a crude form of comparative effectiveness research. They assume all the ase inhibitors are roughly the same, all statins are roughly the same, and so on, or at least high-potency statins are roughly the same, and they take it from there.

But elsewhere in the world they really do base the formulas on cost-effectiveness considerations. And for most of us here, we get our drugs through a PBM, and PBMs tend to use a mix of some, depending on the particular PBM, a mix of cost-effectiveness considerations and acquisition costs in designing formularies.

**Jed Weissberg:** You're quite right. Your basic point is that the output of the analyses depends very keenly on the prices of the inputs, and you can model some of that with sensitivity analyses, but as Alan was sharing with us on a planning call for this, there are many other issues that go into the inputs and process of care rather than little sensitivity analysis around a particular price. You want to mention that?

Well, you were saying that, for example, for kiphoplasty, it's not only the cost of the needle and the cement and the professional time and the room charge for radiology. How

many visits do you do before hand? How long are you in the hospital? How many visits are there afterwards? And there's enough variation in those components of the treatment episode that it really makes the range of cost-effectiveness very, very broad.

So you have to say, "At this price, at this model of care, this is the cost-effectiveness."

**Paul Velevan:** Then I'm wondering in terms of drug formularies, are they taking that type of information into consideration? Let's take a class of seizure drugs, where if you're on one seizure medication, and you need to, because the formulary doesn't cover it, you need to withdraw from it slowly, and then start taking the other medication over a series of maybe 90 days.

And then in the period of time, take a lot of blood tests. Whereas if you're taking a pure class of drug case, where is the cost-effectiveness being...?

**Jed Weissberg:** Right. In Kaiser Permanente we definitely do exactly those analyses, and furthermore look to see in our actual experience, for example, in the statin drugs it's been claimed by certain companies that their drugs cause less muscle pain

than others. So if you start with that, you won't have to get more doctor visits and switch.

So we looked at our experience when patients were started on one or the other drugs, how often did they have to switch to another, and it turned out it was exactly the same no matter which drug you started on. So, informed by that experience, we could inform our clinicians, and then just go for the lowest price of that class.