



Setting Food Safety Priorities: Toward a Risk-Based System

RFF Conference Center
May 23-24, 2001
Transcript of Proceedings

Thursday, May 24, 2001

Day 2: *Improving Methods to Inform Risk-Based Priority Setting*
P R O C E E D I N G S

WELCOME AND OVERVIEW OF DAY TWO

DR. HOFFMAN: Good morning. Welcome to the second day of the RFF conference on Setting Food Priorities: Toward a Risk-Based system. I think we had a really good set of sessions yesterday. The primary focus of yesterday's sessions was looking at substantive issues on what do we currently know about the distribution of risk and the distribution of our regulatory efforts and industry efforts in relationship to those risks.

Today we are going to move ahead to looking at some methodological issues. The underlying question that we are trying to deal with in the conference is what can we do to work on improving coordination across our food-safety system. Today we will be hearing from a series of speakers who will be addressing that issue from methodological perspectives.

Yesterday we heard from both Bernard Schwetz and Tom Billy at FDA and FSIS that part of our goal in trying to improve food safety is to move toward a more risk-based system. Kathy Woteki posed the question of that is something that the National Academy of Sciences has recommended and there is a lot of consensus that if we are going to improve the effectiveness of our food-safety system, we need to move toward a more risk-based system.

While there is a lot of consensus about the need to do that, what there is not is a lot of understanding about is just what that means. Yesterday we were exploring some substantive questions about what that may mean in terms of allocation of resources, but I think, as both Tom Billy and Bernard Schwetz mentioned yesterday, this is really also about process and about methodologies for informing decision making, and that is going to be the focus of our discussion today.

We are going to start out with talking about risk assessment. I think the speakers yesterday all emphasized that the basis for all of this, for any kind of priority-setting, has to be good information about the biological sciences, the hazards that are posed by microbial risks, and the hazards that are posed by chemical risks. Today we are going to hear from Bob Buchanan from the Food and Drug Administration, and Harold Zenick from the U.S. EPA, about where we are currently with the state of our understanding of how to conduct risk assessments of biological and chemical hazards.

We will also, then, hear from Joe Rodricks and Kaye Wachsmuth, who are both very experienced with chemical and biological risk assessments, their views on where we are moving with risk assessment and how that can inform priority setting in and of itself and as a basis for other kinds of analyses that are important as we move ahead into management questions.

Understanding the biological basis is only a beginning, and it is only a beginning when we move into the question of how do we manage the whole food safety system, how do we manage government activities that are trying to help promote food safety, how do we manage risk communication and public education about food safety risks.

Paul Fischbeck, from Carnegie Mellon, will be talking to us about a method that EPA has been using extensively to think about priority setting across a wide range of environmental risk-reduction programs, in

thinking about where one puts regulatory efforts. He and his colleagues at Carnegie Mellon have been doing a great deal of work on comparing risks, soliciting preferences from the public about their view of risk rankings, and Paul will be talking about that.

We are also very fortunate to have Don Barnes from the EPA Science Advisory Board, who has worked for many, many years with EPA's efforts at doing comparative risk ranking as part of their priority-setting efforts, as well as Stu Sessions, who has also worked with those efforts. We hope that we can learn something, as we are looking at food safety -- there are undoubtedly lessons that we can learn from other agencies' efforts to kind of use risk ranking to look at priority setting across a wide range of hazards in different types of programs.

Then, in the afternoon, we will move on to talking about different ways of looking at how the public values risk reduction. As an economist, I have to argue that it is not enough simply to know how many illnesses we are preventing or how many deaths we are preventing, even though, of course, that is incredibly important information and we need to understand that.

But as was pointed out yesterday, not all illnesses are equal and we need to understand the values people attach to reduction of risks of different types of illnesses. There are many kinds of methodological questions about how we go about understanding how people value reduction of those different types of risks.

The second part of the day is really a continuation of a discussion that has been going on, as Elise Golan pointed out in her talk yesterday. It is a reexamination of questions about how we value reductions in risk to human health and human life that has been going on in the food safety area since at least the early eighties and, more recently, this is a continuation of discussions that began last fall in thinking about how can valuation methods help inform priority setting across a wide range of food safety hazards, especially when we are talking about end points as different as cancer and diarrheal diseases, hazards as different as microbial hazards and chemical hazards that may have very latent kinds of impacts.

With that, I would like to reintroduce Mike Taylor, who works with me here at Resources for the Future in the food security and food safety program. He will talk a little bit more about priority setting, and then I would like to introduce our speakers for the day.

THE IMPORTANCE OF RISK-BASED APPROACHES TO PRIORITY SETTING

MR. TAYLOR: Thank you, Sandy, and good morning, everybody. Welcome back. We appreciate your continued interest and look forward to a really interesting day today.

Sandy and I thought it might be useful to have someone who has labored in the vineyards of risk management make some sort of prefatory comments today about the importance of the issues we are talking about today, which have to do with the tools we have to set risk-based priorities for food safety. We had exhausted ourselves so fully in finding the great array of speakers that we have, and we realized that I have spent a little time laboring in that vineyard and I was easily available and would be here anyway, so I am the designated former risk manager to talk a little bit from the perspective of people trying to manage food safety programs, why what we are talking about today is so important.

Of course, I think yesterday's discussion gets at one very important level at which risk-based priority setting is important. If we are going to design a government food safety program that is using its finite resources optimally to reduce the risk of illness, obviously we ought to be understanding as much as we can about where the risks are and prioritizing those risks in terms of food safety benefit to people. That is what the purpose of the program is in the first place.

We ought to use the best tools and the best information we have to prioritize those risks and then, in an ideal system, allocate our efforts accordingly. So that is what we talked about yesterday and that is an obvious context for the discussion today -- it is sort of what this conference is about: How do we improve the way in which we do risk-based priority setting for purposes of system design and resource allocation across the system.

There is another level, though, that is perhaps more immediate and more day to day for people currently involved in trying to manage food safety programs, certainly a more immediate concern. In reality, our food safety agencies are battered constantly by petitions from various interest groups and stakeholders, by the prices of the day, to do certain things that, from the standpoint of that stakeholder or given the issue raised by the latest crisis, seem like a thing that the agency ought to be working on.

It is a constant challenge, I think, particularly for the Food and Drug Administration, which, as Dick Merrill led out with so well yesterday, lacks statutes, any mandate, as to what it is to do with its statutory tools and

resources. It is left to figure it out every day, essentially, as to how it is going to use those resources to improve food safety. It can become very vulnerable to all the pressures that come in every day, and there is a constant struggle among people leading that program to keep the eye on the ball and to be working on those things that are the most important from a food safety standpoint. The agency works hard every day to do that.

Obviously, to the extent that FDA -- and the same applies to FSIS, again, to a lesser extent, given the degree to which the statute drives so much of the use of FSIS resources -- but to the extent that risk managers in the agencies like that have gone through analysis that can be defended publicly, that sets priorities from a food safety standpoint, from a risk-reduction standpoint, for the use of their resources, that provides at least some insulation from the day-to-day pressures to be diverted in this direction or that. This is never going to be an ideal regime in which agencies get to do the ideal every day with their resources, that is not the way life is, but to the extent that agencies can be building up their own agenda for risk-reduction activities, supported, again, by good solid analysis generated by good and defensible methods of analysis, that gives the agencies the ability to stay on target, stay on track, and resist some of the pressures that come along day to day. Again, it is a very hard job that the folks in the agencies have to get through the day; any tools they can have, I think, to establish and defend a set of priorities that keep the agency focused on reducing the risk of food-borne illness is to the good. Independent of any conversations of a broad, long-term redesign of the system, that is a day-to-day need that regulators have and, I think, can be met by the kind of work that we will be talking about today.

The question that I will pose at the end of the day -- I thought about it some more overnight, the centrality of coming to grips with (again, this is another perspective of the risk manager) how good is the information that we have to rely upon when it comes to risk and to inform priority setting, how good are the evaluation methods that we have? What is good enough in terms of being able to rely on the outputs of that sort of work? Where are these methods going? What do risk managers have to look forward to in terms of an improved ability to make and defend risk-based priorities for food safety?

I am going to be sitting here as a former risk manager, absorbing this and trying to learn something and I look forward to it very much.

Thank you.

DR. HOFFMAN: For our first session today, we are going to be hearing from Bob Buchanan, who is senior science adviser and director of science for the Center for Food Safety and Nutrition at the FDA. He currently leads FDA's microbiological risk-assessment efforts. He has worked on food safety. He is trained in food safety and has worked in this professionally throughout his career, and we are very glad to have him here.

Then we will have Hal Zenick, who is director of EPA's National Environmental Effects Research Lab in Triangle Park, North Carolina, speak on the role of public health impact in environmental decision making. With him will be Brent Suhre, from EPA's OPP Health Effects Division.

THE STATE OF RISK-ASSESSMENT METHODOLOGY MICROBIAL RISK ASSESSMENT

DR. BUCHANAN: I would like to thank you all for inviting me to make this presentation this morning. Mike asked me, as I was getting ready to do this, how about covering all the methodologies of microbiology and, by the way, microbial risk assessment and, by the way, you have 20 minutes.

(Laughter)

So you are going to get the really condensed version of this and I would like to focus on not a lot of really detailed methodology, but give you a general feel about how we approach these problems and then a little bit about the usefulness of the information that is generated.

I might note this has been a rapidly advancing field. There are a number of people who are real experts in this and there are details of this scattered around the audience from both USDA and FDA, so strike up a conversation. I would say you have a probability of about 0.15 of talking to somebody who has actually been doing microbial risk assessment if you just pick somebody at random. So let's talk a little bit about this and what is our challenge.

[Slide]

Our challenge is really basically twofold; one, to come up with some kind of measure of the frequency and extent to which the American consumer is exposed to pathogenic microorganisms and then actually convert that into a value that has some measure of public health impact. This is something that came out of, believe it or not -- necessity is the mother of invention -- this came out when they published the SPS agreement and somewhere buried in the SPS agreement and the accompanying TBT agreement it says you are supposed to

be able to do risk assessments when you have a trade question.

A bunch of us were sitting around, saying, gee, we have not the slightest idea how to do that but it sounds like an interesting project. That was really the impetus for getting this started.

[Slide]

I also want to note that this is a rapidly changing field. This is from a WHO-FAO consultation on the application of risk analysis to food standards issues. Our conclusion, after going on for quite a long time about chemical risk assessment and its usefulness, they basically said this sounds like a really great idea, but do not expect us to be able to do it in the future.

[Slide]

Since the issuing of that report and, luckily, the people who wrote the report, the people out in the field who did not read the report, because in those short intervening years -- it has been six years since that report came out -- it has progressed to the point where we are now saying, for example, the U.S. food safety regulatory agencies are actively using risk assessments and developing them for decision making.

These are the major ones that have either come out or are due to come out from the regulatory agencies in the next couple of years, Salmonella enteritidis in eggs, which I think Kaye will talk about later on, Listeria monocytogenes in ready-to-eat foods, vibrio haemolyticus in oysters, the E. coli 0157 that Tom Billy mentioned yesterday, and then a fluoroquinolone resistance in Campylobacter.

[Slide]

This has also progressed to where it has become an international activity. During this past year -- actually, year-and-a-half -- there have been five risk assessments either started or almost completed that have been under way at the FAO-WHO joint activities group.

[Slide]

I do want to compare and contrast this a little bit with chemical risk assessments. One of those things in perceiving risk, I went up to Dr. Schwetz after his talk yesterday and said, "At the risk of getting into trouble, I'm going to disagree with you when I make my presentation, because I don't think there is that much difference between chemical risk assessment and microbiological risk assessment. The basic framework is the same."

It is really the same framework that we have been using for 30 years, the National Academy of Sciences approached us, we refer to it as the Codex Alimentarius framework, but it is really the NAS framework of a pattern identification exposure assessment, hazard characterization, and risk characterization. That is the general framework. We follow basically the same kind of thought processes that take place.

What has been significantly different is, one, we are primarily focused on post-market activities, not pre-market and, two, to date, because they are very new and, three, because of a conscious decision, we have not included a lot of risk-management decisions in the assessments. There is a history of assumptions that have been built up about chemical risk assessments that have been included in the assessments and they are so ingrained now that they are no longer transparent. We have made a conscious effort to try to make our risk assessments as transparent as possible.

But there are some real differences and I do want to highlight those just briefly. One, the hazard-identification phase of a microbiological risk assessment is usually minimal, because the hazards are real. Typically, it is a well-characterized syndrome, there is often a substantial body of both clinical and epidemiological information that we can draw upon. There are case studies that we can pull out, well-defined clinical studies that looked at a variety of factors. So we are not talking, as Vern said yesterday, about theoretical risk, we are talking about real risk, because we have real adverse reactions that we are finding.

[Slide]

One of the other factors that make this a little bit unique is the levels of pathogens that we are dealing with, and on the basis of the exposure assessment, can change drastically in a short amount of time. It was pointed out yesterday by Rob Tauxe that if you leave Vibrio out for about six hours, one can become billions. Likewise, just cooking it for a couple of seconds, the billions could become one again, so it can happen very fast and the extremes that we follow that Rob talked about are often highly complex and include a variety of different steps.

[Slide]

We also have to deal with the fact that we are dealing with a balance in a three-way system. We are dealing with a balance between a pathogen, the host, and a food. These are all biological systems, they are all incredibly variable, so you are dealing with three major sources of variability and things can get very confusing very quickly because of that.

[Slide]

The other thing that is different about this in relation to chemical assessments is disease in microbiology is almost always associated with a single exposure; that is, you deal with individual eating occasions and you usually get sick or you do not get sick. We typically do not have chronic problems in terms of building up a dose within the body. In fact, it is just the opposite. If you build up a dose within the body, you probably become immune and your risk actually goes down. So typically we are dealing with an immunologically naïve population.

Even if you have chronic syndromes coming from it, things like reactive arthritis or hemolytic uremic syndrome, they are still the result of a single exposure.

[Slide]

We also have the fact that a single viable cell has a definable probability of producing an infection. This is a concept referred to as independence, and while it is controversial sometimes, it is one of those default assumptions; you would make this assumption one way or another -- that is, a single organism has the potential for causing disease.

Now, if you get a whole bunch of them together, the probability of causing disease is much greater but, typically, we use non-threshold models for looking at infectious agents that are either linearly or log linearly extrapolatable at low doses.

[Slide]

That immediately restricts the number of models that we can actually use for dose-response models down to three major ones that we are now using, beta Poisson models, exponential models, or Weibull-Gamma models. I might note that these vary in terms of their complexity. The exponential model is a single parameter model. It is pretty good at looking at fairly severe consequences of infections. It is fairly steep in its shape.

The beta Poisson is pretty good at infection, it is a two-parameter model and, finally, the Weibull-Gamma is the most flexible model, a three-parameter model, but when you get something for flexibility you have to pay for it in uncertainty, so while it is more flexible, it is much more uncertain.

[Slide]

A point I want to make, and I make it consistently, is there is nothing magic about doing risk assessments. It is just doing what we have always done but we do it at a much more sophisticated level. In fact, we have addressed concerns of microbiological risk assessment in FDA ever since FDA was founded.

[Slide]

What we have done, though, is we have gone from relying solely on expert judgment up through safety assessments to now being able to use quantitative microbial risk assessment. In the process of doing this, we increased both the level of sophistication we bring to a problem, the level of formality we bring to a problem, hopefully, the level of accuracy that we bring to a problem and then, finally, the level of transparency.

[Slide]

There are basically three types of risk assessments that we do in microbiology, or we could do, and I have sort of lumped them together into these three. One is a risk ranking. This is a priority-setting activity, where you are looking at multiple foods for a single agent or multiple agents for a single food. I will give you an example, as we go through, of *Listeria monocytogenes* risk assessment, which is a priority-setting risk assessment.

The second one is, we will call it, a risk-management or risk-reduction type of activity, or it is sometimes referred to as a found table [phonetic] risk assessment, where you take a single product/pathogen pair and you follow the pathway of that product, seeing where the organism enters, where it leaves, where it goes up, where it goes down. This is typically used to identify interventions or risk-control strategies.

Finally, we could do a risk/risk analysis. To my knowledge, there has not been a microbiological risk assessment that has been a risk/risk analysis. I believe it has been done in conjunction with drinking water and chlorination, but that would be a risk where reducing one risk raises the risk of something else. It is a risk-optimization process that you are going through.

[Slide]

Let's look at the two examples that I am going to bring forward a little bit and how we approach these. The first is the draft assessment of the relative risk to public health from food-borne *Listeria monocytogenes* on selected categories of ready-to-eat foods, a title that only a regulatory agency could love.

(Laughter)

And we do.

[Slide]

Just to remind you what we are talking about, Listeriosis is a severe life-threatening set of infections. It is a relatively rare disease compared to others. There are somewhere between four and eight cases per million U.S. population. Now, however, of those, 20 percent of people who contract that severe Listeriosis will die. That has been fairly constant over a bunch of years. Twenty percent of the cases that wind up being hospitalized will lead to fatalities.

There is also another disease syndrome, *Listeria gastroenteritis*, which is a mild, flu-like disease. We did not model that in the risk assessment, there is not as much known about that.

[Slide]

The problem in dealing with this was *Listeria monocytogenes* had been isolated from a wide range of ready-to-eat foods. If you go through the literature, you can find literally hundreds of foods that at least periodically you can find this organism in. Outbreaks and sporadic cases have been attributed to a number of different foods. We have had a great deal of positive reduction of risks associated with this organism in a period from 1988 to 1995, but since that time our progress has leveled off and we have not further reduced the incidence of the disease.

The question that was facing us is we have this broad array of foods we are concerned about. Which are the ones that we should really focus our attention on and really start to look more actively at? So this priority-setting risk assessment; we are trying to figure out where our focus should be.

[Slide]

To do this we look at 20 groups of ready-to-eat foods. We made an underlying assumption here, for the sake of transparency, that Listeriosis in the United States is associated with these foods. We know that there are some limitations to that assumption. Then we set out to determine the consumers' exposure, develop a dose-response model based on annual disease statistics, and then calculate the relative risk both per serving and per annum for each of those 20 groups. What we are doing is assuming that these 20 groups are responsible for food-borne Listeriosis and then comparing their relative importance among that group.

[Slide]

We did follow the Codex framework for doing the risk assessment in all attributes. We did start off with an exposure assessment. These are the key attributes of that, where we determined the frequency of contamination for these different foods, the number of *Listeria* when the food was contaminated. We were using primarily retail data, but we did take a look at the potential for growth between purchase and consumption. We looked at the frequency of consumption and we looked at the amount of food being consumed. The end result of that was estimating the number of *Listeria monocytogenes* ingested for each of the food groups for the U.S. population.

[Slide]

This is what it looks like. The bars in the back represent basically the food servings, the servings per year, where there was less one *Listeria monocytogenes* per serving, and then it moves up in groups of one to a thousand, a thousand to a million, a million to a billion, et cetera. You can see that most of the food has less than one organism per serving and then, depending on the food groups (those are the 20 different bars there), you can see different patterns of contamination.

[Slide]

We then moved into a hazard characterization, where we started off with data on dose-response relations in mice to help us set the shape of the dose-response curve. We then took into account variation in the virulence of the different L.M. strains using mouse data. We did, then, a conversion factor for relating mice to men, because if you just use mice, you will come up with really strange results. Then we grounded or anchored those data actually in disease statistics, so that when you went through this process and you calculated exposure, et cetera, you had to come out roughly with the right number of cases per year. It was actually internally looking for reflecting the real world.

[Slide]

We came up with a series of dose-response curves. There were actually three, because we looked at three populations. We looked at the prenatal population, children either not born yet or up to 30 days of age, and we had the elderly, which is defined by CDC as age 60 and above. That turns out to be the most controversial thing in the risk assessment, is how we defined elderly.

(Laughter)

You learn a lot of things about human behavior and perception and that has actually been very controversial, and we now have learned to say CDC defined elderly and we just use their data.

(Laughter)

The middle group is if you are older than 30 but under 60 you were put into an intermediate age group.

[Slide]

We did generate dose-response groups. This is the neonatal group. You will note that the center line is the dose-response curve and then there are uncertainty limits around it. You will note that the uncertainty increases down substantially as you get to low-dose extrapolations.

[Slide]

This brings up to the final stage of the risk-assessment, which was the exposure assessment, combining the exposure assessment and the dose-response models, to come up with our estimates of relative risk assessments among the groups.

[Slide]

We actually presented in three different formats. We presented in predicted frequency of Listeriosis per 100 million servings (and I will admit we arbitrarily picked 100 million -- it worked), predicted per-serving relative risk ranking among the different groups -- remember, the primary thing was we were looking at comparing the different groups for their importance, and then the predicted per-annum relative risk ranking.

[Slide]

This is the predicted relative risk per serving for the 20 groups. It starts on this graph. The P stands for pate and it goes down all the way through the 20 groups to ice cream on the far end. This is the kind of relationship that we saw when we looked at the actual calculated risk.

I might note, I do not want to give the impression that ice cream had no risk at all; all foods had risk -- that is one of the things that we confirmed in our analysis. It is just the five orders of magnitude between ice cream and pate in terms of the relation risk between them.

[Slide]

We did do a ranking exercise. We examined all of these. We would run the model and find out where they came out. This is the median risk ranking for each of these. I do want to point out the one ready-to-eat food that was a little different from others: frankfurters. This was the one we assumed people actually might reheat before they ate, so we had actually had three values, or three rankings for them, one when you combined all of the data, one for when you considered only reheated frankfurters, and one where you considered only nonheated frankfurters. You can see that that reheating has a tremendous impact on the relative risk.

[Slide]

We also examined uncertainty here. What we did is, we generated the model, got a ranking, and then we repeated that process 4000 times. We then looked at the risk ranking, using this as a measure of its uncertainty.

This, for example, is vegetables. It came out pretty consistently low. You can see a fairly tight pattern about No. 17 ranking. Again, this was a way of measuring our uncertainty.

[Slide]

We did the same process for risk ranking per annum. Let me talk a little bit about risk ranking per annum and per serving and what they mean and what they include. The risk ranking per serving is the risk that a consumer would face if he walked into a store, picked up a serving of the food and ate it.

The per-annum value is correcting that for the total number of servings consumed within the United States per year. If the per-serving risk ranking is what the consumer faces, the individual, the per-annum value is what the country faces with regard to the hazard.

[Slide]

A little bit about what this is useful for. Obviously, this is a tool for focusing regulatory activity, for effort. By indicating to us that we have foods that have substantially increased risk, it certainly would behoove us to try to focus our efforts toward those, so we obviously would be looking in our action plans to focus on those foods that are on the left-hand side of the chart.

[Slide]

Finally, and I think this is one that economists would be interested in, because it lays out, again, what would happen in a standard. If you could actually segregate, based on microbiological numbers, this would be two graphs here, the upper one, or solid line, is the number of disease cases that would be prevented versus the amount of the oyster harvest that would have to be no longer available to the public, if you set the standards at the bottom, ranging from 100,000 per gram down to 100 per gram.

Again, it is a way of arraying. Here is your decision-making process and here is what is going to be at least our predicted consequences of that.

[Slide]

I might note that we have published these risk assessments in draft. We have done some things that I am really proud of, the team that put these together. One, that these are written by risk assessors for risk managers. We have taken a great deal of effort not to write risk assessments by risk assessors to risk assessors but actually write it for the people who have to use it, so we are really quite proud of that. We have also tried to be as transparent as possible. Any decision we made in the risk assessment is there in the risk assessment. We have also put it out for public comment for a total of 180 days, seeking comments, advice, new data, et cetera, that we will put in before we finalize it. For any of you would like it, it is available either by writing us or you can download it from our Web site. We have sat down with risk-assessor groups and actually walked them through to explain how the model operates and to show it in detail.

I have tried to provide an overview. It is a very exciting time for microbiological risk assessors. I think we are doing some really good work and, people out there in the audience, I applaud you, those of you who have been involved in it. It has been a field that from nothing to, in five years, we have gone to a high degree of sophistication.

Thank you.

DR. HOFFMAN: We will now have Hal Zenick and Bart Suhre from the U.S. EPA to talk about current efforts in chemical and nonmicrobial risk assessments. Then we will have two discussants and then we will have time for questions and answers and discussions about these presentations.

CHEMICAL AND BIOLOGICAL RISK ASSESSMENT

DR. SUHRE: Actually, I am Bart, not Hal; Hal will follow me. I am a long-time employee of the Office of Pesticide Programs. I work for Margaret Stedokowski. She is the director of the Health Effects Division in OPP. Margaret has a large staff, she has about 150 people working for her. About 10 percent, or 15 of those, are involved with the development of new risk-assessment tools. We have been very busy since the passage of FTPA in 1996.

I have 10 minutes, I am splitting 20 minutes with Hal, so it is going to be a really quick run-through of what OPP does to conduct risk assessments for pesticides. The real information will probably come in the question session afterward.

[Slide]

I am going to start by very, very briefly introducing the statutes that authorize the regulation of pesticides. I will then move on to a description of the components of the dietary risk assessment and the process by which the assessment is refined, depending on the regulatory needs, and then I will close the talk by directing you to some recently developed science policy papers for addressing the changes that were mandated by the FTPA amendment.

[Slide]

The first slide here identifies the statutes that OPP works with. They are the Federal Insecticide, Fungicide, and Rodenticide Act, FIFRA, the Federal Food, Drug, and Cosmetic Act, FFDC, and the amendment to these acts, the Food Quality Protection Act. These are the ones that we generally work with.

[Slide]

FIFRA gives us the authority to register pesticides, to license the use of pesticides. If you are not aware of this type of thing, the next time you are in a hardware store or a garden supply store, pick up a pesticide label and look for that EPA registration number. That is the license to sell that particular pesticide.

In order for EPA to license these pesticides, there is a rather significant data requirement that is listed in 40 CFR 158, and this is levied on the registrants, the individuals who produce and market a pesticide. FIFRA also requires that the registration of a pesticide not cause unreasonable adverse effects on people or the environment, so contrary to FDA, we have a very clear mandate to do something about the risk assessment associated with the licensing of a pesticide. This is why we are so big into conducting risk assessments.

[Slide]

The other statute, FFDC, provides EPA with the authority to establish tolerances. Tolerance, of course, is the amount of a pesticide that is legally allowed on a food. The idea is, if a pesticide is used on a food item, it requires a tolerance.

In addition, there is the authority for both the Food and Drug Administration and for USDA to monitor raw agricultural commodities for pesticide residues. Data generated by these two programs, PDP, and FDA surveillance, are used routinely by EPA in conducting risk assessments.

On August 3, 1996, everything changed with the passage of the FQPA, which amended both of these statutes. Prior to FQPA being enacted, OPP tended to do its risk assessments -- or its risk assessment paradigm was to conduct these assessments on single pathways, dietary, occupational worker, residential

environment, individually. Generally, the decision was made on whichever of those pathways was the dominant pathway.

FQPA changed that. It mandated that these pathways be aggregated, be brought together, for the same chemical. It mandated that this be done in order for tolerances to be set, so all of a sudden OPP needed to consider what was happening in a residential environment in order to set a tolerance on a food. This is quite a different paradigm from what we were working on before.

[Slide]

The Act did not stop there. It went on to also have us consider multiple pesticides via multiple pathways; in other words, cumulative risk assessment, where all of the O.P.'s or a group of pesticides having the same toxicological mode of activity needed to be considered. So we are starting from one chemical, one pathway, and changing that to one chemical, multiple pathways, and then going a step further to multiple chemicals, multiple pathways. This was all mandated on August 3, 1996. We are five years down the road and we are just now beginning to be able to do this.

[Slide]

This slide is a little busy, but what it is trying to do is introduce you to the risk-assessment process. This particular slide focuses on the dietary risk assessment. At the top of the slide we have just the basic algorithm that we are dealing with, as you saw from the previous speaker, Bob. The risk assessment is really a function of both the exposure and the toxicity.

Exposure is then broken down further, depending on whether or not you are dealing with dietary or whether you are dealing with residential. This particular slide is focused on dietary. It shows you the components and, actually, the way in which a risk assessment is refined.

What we do at OPP is we do our assessment initially on a screening level basis and then we refine it as the regulatory situation dictates. This means it gets very confusing to interpret what you are looking at, because there is not a direct comparison. If a particular risk assessment is highly refined in one chemical and, for another chemical, is just a screening level, then the direct comparison of these things is very difficult and relatively meaningless. One really needs to be focused as to what extent that risk assessment has been refined.

On the left of this slide is an attempt to describe some of the refinements that occur as one goes through a risk-assessment process. For a dietary risk, if one wanted to do it very quickly and use a piece of paper at his desk, it would start with a tolerance-level residue and assume that the entire crop, that particular crop, was treated.

An initial refinement of that process would be, instead of assuming that 100 percent of that particular crop is treated, to incorporate a percentage of the treatment of that crop, coming from USDA's data. Another refinement would be switching from tolerance-level data that reflect that maximum use down to some of those monitoring data collected by USDA and FDA.

Another step downward is starting to use probabilistic techniques as opposed to deterministic techniques. At this stage we are trying to deal with not only what is the risk but what is the probability of the risk occurring to individuals within a population.

Then we go on down and on particular occasions we will ask registrants (or they will volunteer) to collect very specific data in order to define the risk assessments. Again, we are dealing with a progressive series of refinements. This series continues to be applicable, whether or not you are doing an FQPA risk or a pre-FQPA risk. What changes is the array of chemicals that is used in the risk assessment.

Just for completeness, I have added the toxicity data, the type of toxicity data that are usually available to OPP in conducting one of these risk assessments.

[Slide]

This is my last slide, and what I want to do with this slide is to alert you to a number of things. Again, there has been a significant impact on the program with the passage of FQPA, which amended FIFRA and FFDCA. There has been tremendous involvement by the stakeholders. In the five years since that amendment, there have been many, many meetings chaired by under secretaries of agriculture and deputy administrators of EPA. There has just been a lot of input, a lot of involvement with what does this mean and how do we go about implementing this regulation.

The next thing I want to mention is that out of these efforts a number of papers have been developed and they start to appear on our Web page. If none of you has had the experience of looking at government Web pages recently, they are becoming a wealth of information. They are the first choice for getting information now. I would direct anybody really interested in this topic and understanding what is going on to go to the EPA Web page, click on pesticides, and start looking at that.

If you are interested in the developing issues, what is happening, how things are changing, I would send you to some of the advisory committee reports, much like this organization here. There are a number of those. There is the FIFRA Scientific Advisory Panel, which sits and listens to EPA present any significant changes to the way in which we do business, so a lot of the models that are being developed to support these particular changes, the aggregate and cumulative risk assessments are going before the SAP, and they are commenting. Those are public meetings.

[Slide]

The guidelines that are listed here all went before the SAP and those will all be on the OPP Web page. The way it really works is the agency develops the guidance documents. It develops the guidelines and these all go on the Web site. External to that, those guidelines are used by consultants, by industry, to develop the tools that are then going to be used in order to conduct these risk assessments.

These are the guidelines that you would find, I think, applicable to this particular discussion: how to do an aggregate risk, how to combine pathways to general principles for that, the cumulative (which is a little more complicated, where you have relative potency factors for different chemicals and you are combining), and then the tivec [phonetic] safety factor, which is really the code for dealing with the concern with children's and infants' safety that were mandated by FTPA.

I think I have used up my 10 minutes and Hal will now talk about some of the future trends for risk assessment.

DR. ZENICK: Thanks, Bart. I am going to take about five or six minutes of what is left here to run through sort of how we look at this issue from a research perspective in the Office of Research and Development, and I will start with the fact that we believe our way to best inform the agency is to really try to tackle many of the issues related to the entire what we call environmental health paradigm in order to best inform the decisions.

[Slide]

Obviously, efforts to gather this type of data within each of these cells allows us to understand something about how we might want to appreciate what the risk is and to whom the risk is greatest in doing our assessments, and also gather data that allow us to work back to identify the actual pathways and sources of that risk that we can best direct risk management and pollution-prevention efforts. Each of these, in and of itself, has a substantial research program that is undertaken in best support of the agency.

[Slide]

What I was asked to do was to talk a little bit about where this science is headed and what we are actually trying to begin to accommodate. What I have listed here are a number of the issues that are beginning to impact our research thinking. As Bart pointed out, we certainly are moving away from the concept of a single pollutant to the challenges of understanding what is the aggregate of exposures, cumulative risk being the aggregates of those aggregates.

Obviously, this is a tremendous challenge. It is fairly daunting to try to understand something about how a single chemical actually impacts mechanisms to produce adverse health effects. That becomes equally challenging when you begin to try to understand how the exposures through multiple pathways accumulate to produce a dose that produces those effects.

Then you get to the even more daunting task of understanding what some of those aggregates mean.

Cumulative risk for us might apply to not only a wide number of pesticides being experienced through multiple pathways, it may also deal with other constituents that may constitute that particular pesticide product. It may reflect other contaminants that may be present in that dietary source, metals, for example, PCBs that may be present, and so forth. This becomes extremely challenging.

I think OPP is to be commended in taking the first steps in developing guidance that can not only guide decisions at this stage but also allow us to set a research agenda.

[Slide]

We are moving, obviously, to issues that allow us to target who is at the greatest risk. We are trying to establish efforts to allow us to understand something about what the distribution of exposures is within the population and who is at the high end of those exposures. There are issues that are related to age and we, of course, talked about kids having a major focus for the agency in the last couple of years, but there is the reality of another very growing population that we need to recognize, and that is the CDC's "old people," and the realization that not only is that the fastest growing population in this country, we are all going to live longer and not necessarily be in great health. We have very little understanding of what it means to become older and understand risks that we might have been better able to tolerate at a young age when we had greater capability and capacity.

[Slide]

I will talk momentarily about genetic differences and, obviously, life style and other factors that the individual is in control of may impact, also, placing them at risk.

We are beginning to recognize the need to move away from some of our classic understandings of cancer and non-cancer, where cancer was viewed as a non-threshold phenomenon and non-cancer was viewed as a threshold phenomenon, to really trying to better understand something about the actual mechanisms and beginning to move toward using more common approaches to assess risks that appreciate what the contributions of these particular phenomena may be; marked pressure in moving away from high-dose descriptive studies to an ability to better understand what actually operates at exposures that may be more realistic of what the population is exposed to, and begin to try to develop models that allow us to extrapolate those risks and predict what they are, trying to develop an understanding within the agency with regard its being able to obtain mechanistic data and apply them within these models.

I think one of the challenges in the paradigm I presented to you in the preceding slide is that although there are a number of components in that paradigm, historically we have been very stovepiped in examining the determinants within each of those components, whether it be exposure, dose effects, or source.

It turns out the more effective way for us to carry out risk assessments is to actually understand what the predictive linkages are between those elements, the ability to understand how we take what we understand to be the exposures the individual may be experiencing from measurements in different pathways to actually predicting that to dose and then being able to link that dose into actual effects. It is the ability to do that actually allows, I think, to have the better interpretation of the data that is available.

We have spent most of our research efforts in trying to provide science that will better inform the risk-assessment decision. I think a question that is becoming increasingly apparent on the public side, as the cost for incremental improvement in reducing risk is increased, is how do we know that those decisions were correct. We are beginning to see efforts to orient our research programs to also assess the impact of those public health decisions, not simply how well have we informed those decisions.

This has implications, obviously, beyond EPA, because this not only, perhaps, argues for developing markers that allow us to do that assessment, but it has impact for monitoring and surveillance systems in this country, how good are they and how adequate are they to allow us to assess impact.

The other reality is that we do not operate in isolation and many of the activities we engage in that have ecological impact have, in turn, human health impact. Obviously, agricultural practices that may impact, water quality that impacts exposure to humans are of concern, and there is an effort now to actually develop a much more integrated and holistic approach to understanding interactive risk to these systems.

[Slide]

New tools always bring promise for new answers but they also bring new challenges. One of the themes that is apparent is the tremendous development of molecular and mechanistic techniques that allow us to measure a wider variety of events that we have had available in the last several decades. For example, there is a recent report that has been released by CDC that has presented data on 27 pollutants present in the blood and/or urine of the U.S. population as part of the NHANE survey.

[Slide]

This is just a subset of those data -- I have not bothered to fill all the information in. This is a subset of pesticide metabolites that have been reported in that particular survey, showing what the population distributions are of these. This is just a small subset of what we will begin to see accruing over the next several years as the efforts by CDC give it some status and trends information on what is present in humans.

This presents, I think, a rather interesting set of challenges. For example, it would be interesting to know who is it who is at this higher percentile number, but I think the greater challenge that we have is understanding how you actually interpret this information. We have not had this information available before.

It was interesting, in looking at the materials that Pennea-Crist presented yesterday, the last conclusion on her slide was that if the accurate national profile of exposures remains elusive, this is going to begin giving us the type of data -- the challenge is going to be the interpretation of this information. The challenge is going to be understanding how food and other pathways may contribute to these body burdens in terms of effective pollution- and risk-management strategies.

[Slide]

As I indicated, we are developing sophistication to measure events at smaller and smaller refinements. The problem we have is not interpreting the ability to measure to be an estimate of risk but, instead, actually

being able to interpret what these events actually mean. One of the biggest concerns and challenges of our risk assessors is that you begin to present mechanistic data, you have failed to establish this relationship and, as a result, it is very problematic that you can actually base a risk assessment on such events.

We are now into the "omics" era. This is terrific, it brings great promise with it. The challenge we have, to some extent, is what the hell do you do with it? This has some very major promise. The problem for us, of course, is being able to understand the genetics that determine these and understanding how environment contributes to that.

That has some interesting implications, because what that argues is very unique individual factors that may determine whether an individual is at risk, even in the presence of exposure, which suggests that perhaps risk assessment down the road may not be population-based but may be more individual-based, as we understand this, and how does an agency actually go about protecting the individual as opposed to the population.

I think that this is going to create a tremendous opportunity, but it is very intimidating as to how we are going to make use of the information. It is literally intimidating in terms of the onslaught of the data that are going to be provided. This well exceeds anything that we have anticipated and, unfortunately, knowing that this was coming, we spent very little time actually dealing with the technology that we are going to need to advance this particular science.

I think we have other issues, obviously, related to having access to this information and, at the same time, trying to appreciate confidentiality issues that surround these. So we have tools that are certainly going to greatly increase our ability to understand what the contributions of environment are, what the impacts are on the individuals and, I think, give us insights into better health intervention, pollution prevention, and risk-reduction strategies, but there is a tremendous challenge for the science to be able to be in place to take these advances and make sure they are used appropriately in the best protection of the public health.

Thank you.

DR. HOFFMAN: We will now hear from two discussants. Joe Rodricks is currently a principal with Environ Corporation, which he founded. It is a consulting firm that does technical consulting on chemical risk assessments. He was one of the early developers of risk assessment in FDA, where he served from 1965 through 1980. He served on multiple National Academy of Sciences committees on risk analysis, and we are very glad to have him with us today.

DISCUSSANTS

DR. RODRICKS: Good morning. I am long an advocate of the use of risk assessment in decision making but I am also an advocate of keeping limits in mind. My discussion is aimed at some of the limits of assessment. You have heard a lot about exciting new developments; I want to talk a little bit about limits and I am going to take you back 30 years to begin that discussion to a time when I was at FDA and responsible, in part, for a rather potent chemical carcinogen called aflatoxin, which is still around.

We did research on aflatoxin for some time. The question arose about how to set limits on aflatoxin in food products, because it certainly was fairly widely dispersed in foods at quite low levels. I looked around and I talked to some of the toxicology people at the agency at the time, and they said, "Well, we have the methodology for dealing with threshold agents. It is this ADI methodology, it is kind of a bright-line methodology. We don't do that for carcinogens, that would be wrong."

If you look in the literature of the time, it goes back into the 1940s, 1950s, and 1960s, when carcinogens were of great concern, that was the prevailing view, that carcinogens were somehow different from other toxic agents, that the threshold idea was not appropriate for carcinogens, there were no safe levels, as we used to say at the time.

The idea of using the ADI, this threshold kind of approach, to set some limits for aflatoxin was out the window. What to do? If you, again, look at what regulators had been doing there in the early to mid-1970s with carcinogens, there were sort of three approaches. One, you banned it, if it was easy to do so. That was, of course, a very limited activity, because it was limited to things you deliberately added to the food supply, so it did not have many applications. Pollutants, contaminants, and so forth, you could not easily ban. The approach with pollutants was either to set arbitrary limits, which was done for a few (not just food, but also applied to water and air), or to ignore them completely, for lack of methodology, and, in fact, most of them were ignored.

I looked around and worked with some other folks in the agency and we decided to try to look at risks from carcinogens. Looking through the literature, there was quite a lot developed at that time in the scientific literature about how dose-risk relationships might behave at very, very low doses, the low-dose problem, as it is called, and different ideas about how that relationship might look, different ideas in the literature about

one might move from animal data to people. Most of the information we had back then was based on animal experiments, there are a lot of them. It is a very important carcinogen, but the levels needed to produce cancers in animals were still five to 10,000 times greater than the levels that we would commonly consume (at least in the Western World) in the food supply. It is a typical problem in cancer risk assessment.

So what to do? We had an arbitrary limit for aflatoxin at the time. We decided we were going to move toward a risk-based limit, make the hard call about risk, and some risk goal, some risk target, for aflatoxin, but that required that we undertake a risk assessment. So I and some others went into some back room and we worked for a few years on a risk assessment for aflatoxin, looking at all the different approaches you use, all the assumptions and models you had impose, to get to some kind of estimate of risk.

When you did that, what you found is a huge difference in the projected risks, the predictive risks, at low dose, depending on the choice of assumptions and models.

It was quite apparent there was nothing much you could do about that. Science did not give you a clear answer as to the best choices to make. There was some support for each of these, both in theory and in empirical information, but not a clear answer.

So what do you do with risk estimates that vary by 10,000-fold? We did. we made what is clearly a policy choice. We opted to base regulations on the high-end risk, the risk that we could not discount but the risk that looked as if it projected the greatest risk for the human population, a conservative public-health-oriented view based on that scientific uncertainty. We began there.

I tell that story and that story is still applicable today to most risk assessments, because it points out that you cannot get, in at least most chemical risk assessments, whether it is for carcinogens or others, to some kind of picture of risk without imposing assumptions and models that do not have complete scientific verification -- and, in some cases, very little scientific verification.

We are working with risks that are, in most cases, well beyond the current power of epidemiological science. Dr. Zenick talked about moving in some direction that gets us closer to an understanding and direct measurement but we are still quite a long way from that point.

There is a 1984 study from the National Academy of Sciences, quite a prominent study, a seminal study, on the whole field of risk assessment that is primarily a study of what it called science policy choices in risk assessment, the kind of policy choice I just described. Given the limitations, we either remain silent on risks, say nothing, or we impose some assumptions, some models, that do not have complete verification.

The Academy said you should do this, you should impose models, using the best available science, recognizing the limits and uncertainties. You should be very careful in how you pick those assumptions and models. These things have come to be called -- I think Bob used the term -- "defaults" in sort of the common vernacular, and there are quite a lot of them in a typical risk assessment.

The Academy said do it, write it down, agencies. It said to the agencies, "Put this in the form of guidelines so everyone knows what you're doing." That is something I would recommend the microbial people begin to think about. You need some experience with actual assessments before you can begin to formulate guidelines, but you ought to be saying what you are doing.

The other thing about defaults is that you ought to allow for exceptions, for the kind of new information Hal Zenick talked about, to be brought in, in specific cases, if real data shows the default to be incorrect. We have not done a terrific job in that. That is something that has been very, very slow-moving.

The other point here that all this comes down to is that when you have a fairly consistent and generic approach to assessment, and I think you can probably say that for EPA and FDA now with respect to carcinogenic assessment, there is a pretty consistent approach. That is an oversimplification but it is not bad. That allows risk comparisons to be made. That is one of the beauties of the approach. There are some limits in that but you can now compare dioxins and PCBs and aflatoxins and acrylamides and vinyl chlorides, all these carcinogens that have quite different risks.

The problem, of course, is that because we are using assumptions and models of unproven validity you cannot say anything about the actual size of the risk, the predictive risks. You can make some broad generalizations about it, but it is difficult to say anything about the actual risk, so are we preventing one case per year, 10 cases per year, 100 cases per year when we control aflatoxin? That is much less certain than saying we ought to pay more attention to aflatoxin than to -- well, I will not say dioxins, okay.

Anyway, a couple more points, but my main point there is that we have got quite a good system now, I think, slow-moving toward change but still a good system that is now encoded in some guidelines. EPA has been quite good about guidelines, FDA has not done that, but their assessments are very much like EPA's. We are seeing the same thing developing in Europe right now as they move toward a risk-based approach

to carcinogens.

A couple more points, then I will be quiet. The approach for dealing with chemical agents is thought to have thresholds, we treat every other toxic effect as if it has a threshold, without complete proof that that is the case, but that is the procedure now that exists. EPA uses what is called a reference dose. The ABI approach is a similar concept. It is a bright-line approach. It says if you are below, everything is fine. If you are above it, something will go wrong, but we cannot say what.

It was developed for premarket clearance purposes, originally, for pesticides and food additives. It is not very informative, as in, for example, the recent case of methyl mercury in fish, where the National Academy of Sciences talks about fairly large numbers of people at risk, because their exposure seems to exceed this threshold, this reference dose that is proposed for methyl mercury, but a very unclear picture of just what that risk is and how it is distributed. There is a lot of work that needs to be done there.

One more point. Data. I have heard people many times accuse risk assessment for problems that are really problems of the underlying data, and it is not fair to risk assessors to do that. Risk assessors have to use what is available. They may make recommendations for getting data of the type Hal Zenick talked about to improve the assessment, but you cannot make up in risk assessment for lousy data, deficient data, or just meager data.

Many substances, many additives, for example, pesticides tested in recent years, have very extensive data bases, but some of the older additives, some of the older pesticides, some of the contaminants have varying degrees of quality and sufficiency in the data base, so that makes a problem for risk assessment.

I am now working on a committee at the National Institute of Medicine on upper limits for nutrient chemicals, because, of course, they can cause problems. That is a data base that is quite poor, in fact, because people have not studied nutrients for their toxic effects very thoroughly. A huge difference in the data base, again, if you ignore that, it makes risk comparisons quite problematic. Risk assessment has a way to sort of compensate for some of those deficiencies, but not very satisfactorily, and it is not a good solution.

Those are just some of the limits I wanted to point out. I am still sort of an advocate in decision making. It may not be informative for all kinds of decision making but informative for many kinds. My attitude toward risk assessment, I like to paraphrase Sir Winston, who was talking about democracy as a form of government. My paraphrase is risk-based decision making is the worst possible basis for decision making except for all the others.

(Laughter)

Thank you.

DR. HOFFMAN: Kay Wachsmuth is joining us to provide us with some additional comments on microbial risk assessment. Kaye is deputy administrator for the Food Safety Inspection Service for public health and science. She comes from a long and distinguished research career at CDC and has spent a short time at FDA, and so understands a bit about that side of food safety, too.

DR. WACHSMUTH: Thank you.

[Slide]

I would like to start by saying this first slide is not entirely accurate. I sit where many of you may sit, which is some place between the risk assessors, the highly technical risk assessors, and those people who are writing the policies and even implementing the policies, and that seat gets warm sometimes, but that is where I think we need a lot of work.

I am going to talk to you about the risk of Salmonella enteritidis in eggs and egg products. Bob described it already, micro-risk assessments that are already under way right now, but the portfolio is pretty slim. This is not AN example, I think it is maybe THE example, of how we have really applied formal quantitative risk assessment to our regulatory efforts.

[Slide]

The problem came to our attention in the late 1980s with a multi-state outbreak of Salmonellosis, which was eventually tracked to Salmonella enteritidis in a raw egg product in a frozen food. Since that one outbreak, multiple outbreaks occurred, and the number of outbreaks seemed to be increasing with time, as did the sporadic illnesses. That is what this graph depicts.

These are pre-FoodNet data, old CDC passive surveillance data, with a soft denominator, and there are some other limitations, but if you look at the bright yellow line at the bottom, within all Salmonella, enteritidis was growing in the 1980s and up to 1996, where it became the most commonly documented and reported cause of Salmonellosis.

Where we were situated at CDC, when we looked at Washington, D.C., it looked as if things were broken

here. I think at that point in time, to those of us in public health, it seemed that the egg problem might be the poster child for a single food agency. I think we have made a lot of progress since then; we certainly are not at the solution yet.

[Slide]

In 1996 we commissioned a risk assessment. What we asked those risk assessors to do was to develop a model so that we could go from farm to table to the consumption of eggs that would be internally contaminated with S.E., to identify through sensitivity analyses the variables that would most likely contribute to the risk of illness, and to evaluate the mitigation strategies, possible mitigations that would affect the public health risk that we could control and, also, very importantly, to identify research needs.

[Slide]

We structured the risk assessment in modules. There were some advantages to doing that. It broke the risk assessment into manageable bites, if you will, because this was an assessment that we considered very large in scope. Even though it is a single pathogen and a single product, it is farm to table.

The risk assessors were asked to do this and to run their assessment without any consideration of the authorities of the regulatory agencies, by the way. I think later we viewed that as potentially some information that might have been missing to them but it also gave them some freedom to look at the whole problem without trying to look at particular risk-management concerns.

Another reason the modules gave us an advantage was that each module could then be analyzed for how changes in input data would affect the output and could affect the risk.

[Slide]

Within each module we know that some variables are more important than others with respect to how they affect the risk. The sensitivity analysis for each module was completed by correlating modules' input with the output. The outputs that were most affected, or the inputs that most affected the outputs were then viewed as good candidates for greatly influencing the eventual model output on risk, so we identified what we thought were sensible variables.

These had to be things that we could affect. They had to be something that, for instance, we could affect the prevalence of S.E. in flocks, so this was a sensitive variable, unlike affecting the growth of S.E. within an egg. So within a module mitigation variations were changed and the entire model was run again to compare that mitigation to the base line.

[Slide]

This is an example of two of the strategies that were first analyzed. The first example is that of an egg that is laid into ambient 45 degree air temperature and then it remains at 45. In that case, compared to current industry practices, where this does not exactly happen for a number of reasons, if that were the case, we would get approximately an 8-percent reduction. Bob explained earlier that this is a mean value for a distribution, and it is a fairly wide distribution, because of the uncertainties that we deal with.

The second example would be if the hen laid the egg and it was immediately brought down to 45 degrees internal to the egg and then held at 45 degrees until the time it was consumed, we could predict a 12-percent reduction in the illness.

We found these to be rather exciting scenarios that we could anticipate and tried to find other ways to express this.

[Slide]

I guess these are economic terms: elasticities. This was a way of looking at other variables where selected input variables were reduced by 25 percent. That would be all of the potential mitigations that are listed. If the decrease in human illness were less than 25 percent, the mitigation elasticity is less than one. If the decrease was more, then it was more than one.

I think what is notable here is that in order to get an elasticity of greater than one, we had to combine things. We felt that this validated what a lot of public health officials and industry officials had said in terms of needing a multi-hurdle approach to this problem.

[Slide]

Then the egg safety plan: In 1998 and 1999, all of those federal agencies that have responsibility for egg safety came together. We came together in doing the risk assessment and then those same agencies came together in terms of how to deal with it within their authorities.

The egg action plan I am certainly not going to go into in any detail. This is not a two-page guidance document. This is about a 40-page, very detailed document with specific objectives and then actions under those objectives, including time lines and agencies responsible. This is from FDA on the farm through even CDC's part in monitoring food-borne disease to assess how well we are doing.

[Slide]

Again, the risk assessment did guide the development of that action plan. In fact, working groups were created to develop that action plan. In fact, the working groups that were created to develop the action plan corresponded very much to the modules of the risk assessment.

[Slide]

We feel we have worked very hard and we do have a useful risk assessment but we also feel it is the beginning and it needs to be improved, for all sorts of reasons.

[Slide]

The challenges ahead -- these are, actually, some people might view them as limitations, and I was trying to spin it a little more positively -- are those data gaps that have been mentioned that really hamper us have given us a rich research agenda. For this particular risk assessment we know that in order to narrow that distribution of probabilities we are going to have to know much more about the epidemiology on the farm, the bacteriology of S.E. in the egg, the human behavior in handling and preparing the food, and the human response. We have to know more about all of it to be more accurate in our predictions.

"Right sizing" -- this means making these tools more timely. It takes us quite a while to produce this. This is more than a two-year effort, really, and now, adding the action plan, we are pretty far out. It needs to be a little more user-friendly and we are undertaking that with some contracts with risk assessors.

It is simplistic in a lot of ways, even though it is complicated; the level of complexity needs to be evaluated again. We do need to be more responsive now to specific management needs of those specific regulators and their authority limitations. I feel that this third bullet is perhaps the most critical challenge for us now. We have an assessment, we are trying to put it into action, to create the policies, to write the regulations, and that is requiring going back and forth, so even though we functionally need to keep assessment and management separate, we have to have, we know now, a tremendous interaction between those two groups, those two activities.

I think the best side of it, the up side, is all the organization that this brings to bear on any particular problem and the focus that it puts on human health outcome. I think for food regulators, for food safety officials, it is key, it is critical, it keeps from counting bacteria in food.

The down side is, now we know we can do these quantitative risk assessments, I would guard against an approach that says we have to have this before we can take any regulatory action to protect public health, said, probably, more with my CDC hat.

Thank you.

DR. HOFFMAN: We would now like to take about 10 to 15 minutes for questions and discussion.

GENERAL DISCUSSION

DR. ZAKARIADZE: Ira Zakariadze, University of Arkansas. I would like to address my question to Dr. Buchanan. First of all, thank you very much for your presentation, I really enjoyed it. I have two questions, actually.

Working in food science and the poultry department, we are working on pre-harvest poultry production risk assessment. This is particularly at hatching and grow-out stages or, as we call it, nodes of the poultry production. This is, as you described, product pathogen pathway; in other words, farm-to-table risk assessment.

My first question is, in this case, could we use the same methodology as we are using for slaughter plants? In other words, could we use dose-response relationship in regard to pre-harvest animal production risk assessment? What particular methodological approaches could we have in this case?

My second question is, what are your thoughts about microbial risk assessment from the poultry company standpoint? To simplify my second question, could a production plant manager, for example, put his particular parameters, for example, temperature, or microbial Salmonella level, into the computer and in that way improve or enhance his HCCAP plan and determine what risk-management decisions should be made in his particular plant?

DR. BUCHANAN: Let me address them in the order that you presented them. Kaye mentioned -- basically what you are looking for here is are you interested in eggs or are you interested in broilers?

DR. ZAKARIADZE: Broilers.

DR. BUCHANAN: Currently there is risk assessment that addresses broilers. However, there is a framework of one being developed by FAO and WHO in consultation with an international team of risk assessors. Hopefully, what we will be seeing here is a relatively simple model that is built up in modules that, then, as more data become available, we will be able to modify them relatively easily in order to take in the variable in question.

I would recommend that you take a look at the WHO-FAO Web site. The draft risk assessment is available there. It still needs some work, so I am a little hesitant to say that is ready for you just to put in your individual data sets, but it is getting close. They also have established a second model for *Campylobacter* in broilers, and that is just being started. There are models available in the latter case from a combined risk assessment that was done by Health Canada and U.S. FDA that is available. You should contact Anna Lamberding, who is with Health Canada.

In regard to your second question, that is an interesting one. It is one that we have had a lot of debate on, on who can use risk assessments and who cannot. Generally, most people look at the daunting amount of work that has gone into microbiological risk assessments and they say how could anyone other than a government ever think about doing a risk assessment.

Luckily, I am not a member of that school of thought. I certainly feel that there are some attributes about risk assessments that ought to be done by the government. Probably the dose-response end is that kind of data, but there is no reason that you cannot take the exposure assessment portion, get hold of the model, modify it to your own personal uses, and then use it to look at your own unique situation.

In fact, I think that one of the futures for risk assessment is going back and looking at HCCAP. That will be the next degree of sophistication, particularly as you get into foods where the process itself is less clear-cut, where we are looking at multiple steps that have to be controlled. I think it has a very interesting feature, but I agree that there are some components of that risk assessment that will probably always have to be done by government.

DR. BYRD: My name is Daniel Byrd. I work in a consulting firm called CTRAP. I will put this on the skids for the entire panel. The effort this morning, beside being very informative, seems largely to try to put microbial risk assessment as it is currently practiced, and chemical risk assessment as it is currently practiced, into the same basket, because the steps they go through have similar labels attached to them, and I want to disagree quite strongly with that.

My comment after the Salmonella risk in eggs came out -- I sort of walked through that step by step with my USDA colleagues -- was that it was a landmark, a real benchmark, a change in the weather, so to speak. They thought it was because, A, it was an application to a food product and, B, it was an infectious disease risk assessment.

My answer was no, it is the risk-assessment model. You have explicitly identified the uncertainty step by step and you have integrated that uncertainty over the assessment. Chemical risk assessors do neither of those things, so the result is, I am constantly in a position in chemical risk assessment where I have no basis to disagree, unlike what Dr. Buchanan described as the epidemiologic grounding of the risk assessment.

I will give you a brief example and ask the panel to comment on this. I think we are talking about two completely different things, and there is a real danger in trying to compare these risk assessments. For benzene, if I took all of the high-benzene-exposure data, which is where we have linkage to AML, and threw them out and looked at only very low-level exposure, which is very well-validated, because the petroleum industry monitors benzene levels all throughout the United States, and I took EPA's benzene estimate for potency for acute myelogenous leukemia, I predict three times more AML cases than the National Cancer Institute says occurs.

When I put this on the table for the EPA officials, they said this is good, it shows we are not underestimating the risk. If you take that kind of attitude, there is no ability to deconstruct the risk assessment, there is no capability to do what Dr. Rodricks has identified, that for historical reasons we do not explicitly identify the assumptions or attempt to assess uncertainty of them.

DR. RODRICKS: I think uncertainties ought to be assessed and described and dealt with when new data appear, if they can somehow reduce uncertainty. I do not know about your benzene case and how well-documented that excess that you describe is, but one of the things the National Academy committee said in 1984, in its first report on risk assessment in the federal government, and in 1994, an update on that report, is flexibility is key. Describe uncertainties, let the manager know what they are, and also be ready to substitute data, when you have them, for that.

I do not think destroys the framework for risk assessment and I do not think that says that microbial assessments are a different framework from chemical risk assessments. It simply says we have good risk assessments and not-so-good risk assessments and we need to find a way to make them as good as possible, given our limitations, but I do not see that that says we necessarily have a different framework for the assessment of risk.

Let me just say, the point you described, I mentioned defaults and the selection of defaults as a very important part of the process. Selecting defaults, because they have a huge influence on the outcome of an

assessment, is a very difficult process. The early procedures were simply to select defaults at every step in the assessment that gave the highest possible assessment of risk; that is, better safe than sorry at every step in the process.

That was done out of fear more than anything else, you did not want to underestimate risk. There are better ways, there are better criteria, for selecting defaults now. We have not moved strongly in directions of picking defaults, so we avoid what I call the scenario of hundreds-exposed, thousands-will-die kind of assessment, which is maybe what you are talking about. At the same time, I recognize the difficulty of selecting the sort of science policy in risk assessment.

DR. SUHRE: I would just like to add a little bit, since I was speaking on the chemical side of things, that, indeed, we are developing new risk-assessment tools. Those tools have been taken in front of various panels. The issue of uncertainty is never escaped in those panel discussions.

I do want to mention, there are certainly great benefits in dealing with uncertainty, but there are tremendous challenges. The people who have really done this have indicated that when you go down that uncertainty road you become very vulnerable, and you become very vulnerable to your judgment, your assumptions, et cetera. It is a double-edged sword. We realize it is there, we realize we have to approach it. We tend to deal more with the idea of sensitivity analysis and contribution analysis as our way of looking at these.

The other thing I might add is that as we move from deterministic to probabilistic assessments there is the idea that that probabilistic assessment applies to every input in the algorithm. We are getting pretty good at dealing with the exposure side of the algorithm and probabilistic inputs. We are not so good on the toxicological side in the probabilistic treatment of those inputs on the tox side.

MR. TAYLOR: Mike Taylor with Resources for the Future. For the risk manager who is looking to use risk information to set priorities and allocate resources in accordance with risk, the question is what is the actual public health impact of a particular situation, obviously, and then how do those compare with one another? As we are beginning to develop quantitative risk estimations for microbial pathogens, which can be laid down next to quantitative estimates for carcinogens and chemicals, we are going to find ourselves in a situation in which we are looking at projected -- let's just make it up -- one in 100,000 risk for this array of microbial hazards and one in 100,000 risk for chemical hazards.

I am just wondering what risk assessment is able to tell us today about how one would compare those quantitative estimates of risk if one, again, wanted to use that information for priority setting and resource allocation?

DR. BUCHANAN: Let me take a shot at that and answer it in two ways. One is that what has not been done well in a lot of risk assessment has been to really take a look at the severity analysis that is associated with it, particularly when you are dealing with an adverse -- or a hazard that has multiple end points. One of the things I would recommend as you do risk assessments is that you look at multiple end points to be able to measure and have some kind of estimate of the total impact of the agent.

The other one is a question of perception. I do not think that we are ready to have risk assessors start making assessments of people's perceptions of risk. That is another whole field that is separate. I want to be careful about trying to do things across the board, because the ways people perceive of chemical risk, carcinogens, microbial risk, et cetera, are not the same. It is like comparing apples and oranges.

Until you can find that Rosetta Stone that lets you translate perception into actual risk-management decisions, we are making it up as we go along at that rate. I do not know if we will ever be able to say four microbiological cases is equal to one case of a chemical problem, or vice versa. You are dealing with human perception here.

DR. HOFFMAN: That will be an excellent segue into our next session.

Any other response to that?

DR. SMITH DE WAAL: Caroline Smith de Waal with the Center for Science in the Public Interest. This is probably more of a comment than a question but I want to interject a little politics in here.

Excuse my faulty memory if some of the details are not quite right, but in the early 1990s there was a big effort on Capitol Hill to get a regulatory reform bill passed. This had mandates for cost-benefit analysis and risk assessment to back up every rule, particularly aimed at public health and safety rules.

My recollection is the bills never passed because OMB essentially adopted these requirements anyway, and at USDA we have two levels of cost-benefit and risk-assessment review going on. There is one that happens at the agency and then they all go over to OMB and do it, again, so we have two levels of economists, essentially, reviewing these rules before they get out.

Again, my recollection is that consumer and environmental groups were very concerned about these bills, because they saw risk assessment as vehicles to delay needed health and safety regulations and they also

saw that risk-assessment and cost-benefit analysis might be a way to say, well, it is okay, the costs over here justify the lives lost or the illnesses over there, that somehow we were going to be balancing economic interests against life and health.

So that is the introduction to what I think is the concern that I am having right now. Here we are, nearly 10 years later, we have an S.E. risk assessment, we have a Listeria risk assessment. CSPI has petitioned the agencies to get on-farm controls for S.E. before we ever had the risk assessment done. We now have it. The risk assessment essentially confirms what CSPI petitioned for. We still do not have on-farm controls for S.E. in eggs.

Listeria monocytogenes, we have a huge outbreak, 100 people, 20 people dead. We have to wait for a risk assessment. The risk managers say, "We can't act, because we don't have a risk assessment." You have dead people but you still need the risk assessment. So now we have waited and we do now have a proposed rule out to deal with Listeria monocytogenes in ready-to-eat meat products, not in the other products.

My issue here, and I am laying it out to this audience, but I think it is a serious problem we have. I think the pendulum has swung too far. The question is, is risk assessment the servant or the master? Right now I feel as if risk assessment is the master, because the risk managers say, "Can't do it, we don't have risk assessment." Risk assessment should be the servant. We should be able to move to risk-management policies and then await risk assessment to help refine those policies and make them better.

DR. BUCHANAN: I will try answering at least part of it to Caroline. Caroline throws out the bait and I cannot help but at least sometimes go after it.

I am going to reflect on it in two manners. I am going to pick on Listeria in the first instance, and it is an interesting one, because the requirements in place now are actually very, very conservative and, in fact, are default to zero tolerance in a ready-to-eat food, so the risk assessment is not going to be something that is more stringent than zero. In reality, we have taken the conservative approach to risk by establishing that zero default value.

Let's also go back and look at your definition of a risk assessment. It is interesting here, because at a recent FDLI meeting I had to make a presentation about the integration of this in the regulatory process. I would put forward to you as a thesis that the preamble of any proposed regulation is a risk assessment. If you cannot establish through some evaluation of the risk that there is a problem, you are not going to have a regulation.

If you are looking for everything has to be a very detailed quantitative risk analysis, as you see here, no, I do not think any of this panel would agree with that, but when the science gets complex, when there is a great diversity among the stakeholders who are looking at data and interpreting them, when the data can be interpreted in more than one direction or they are incomplete, then it is a tool to help us through this process. It is not the be all or end all, but it is a tool to allow us to look with more accuracy in very complex situations. There is no restriction to taking action while you are waiting for a risk assessment and, in fact, typically we take action and then follow it up with risk assessments.

DR. HOFFMAN: I am afraid we are going to have to cut off questions at this point. What I would ask is if you could hold those thoughts. What I am aiming for is to have a good chunk of time at the end of the day to come back and integrate these issues. Also, panelists will be available during the break. I ask that we come back at 11:15 and we will try to get back on track here.

[A brief recess was taken.]

DR. HOFFMAN: We have an emergency in terms of a family emergency. Paul Fischbeck had written me he would be coming in late last night because he had a family emergency yesterday. He has not shown up yet. I am hoping he will be here after lunch, so what I have decided to do is to go ahead and do the panel on public health approaches (is the way I think of this) to evaluation of health outcomes.

I had set up the day to kind of progress through methods that began with addressing the point that Bob Buchanan raised, that as we start to think about risk ranking and as we start to think about issues of not just the incidence of disease but also the severity of disease, we quickly move into also thinking about what are the kinds of risks that people care about most.

As we are setting priorities, it is not simply how many diseases are we preventing, but it is also a matter of the severity of disease, the concern that people have for different kinds of health outcomes. So we are going to do comparative risk assessment first and then move on into the more dollar-enumerated methods of trying to assess the value that people attach to reducing risks of different types of disease or bad health outcomes.

Instead, we are going to start with talking about quality-of-life year indexes and cost-effectiveness analysis. Milton Weinstein, from Harvard Public Health School, will be doing a presentation on this. Milton, as

many of you know, is probably one of the leading experts in the U.S., if not the world, on cost effectiveness and QALYs, and we are very happy to have him.

Clark Nardinelli from FDA is sitting in for Richard Williams, who was called out of the country and will not be able to attend. Clark has been very helpful in helping plan this conference and was also instrumental in planning a conference last fall on valuation of health for food safety purposes, at which we began a discussion of the kinds of research developments that are needed to adopt methods that we are using in either clinical settings or in other environmental health settings for food safety purposes.

I should not say beginning, because, actually, that discussion about food safety has been going on for many years, if not decades. As we are thinking about trying to have a risk-based integrated system, that raises new questions for the applicability of the methods we have for comparing risks. That part of the discussion, I think, is something we were beginning to have last fall and this is a continuation of that.

Anne Haddix is a professor at the Rollins School of Public Health at Emory University. She has had a long history of working in food safety, although most recently she is working on health policy in developing countries. She has worked extensively with CDC on food safety issues and we are glad to have all three of the panelists to join us.

USING QUALITY-ADJUSTED LIFE YEARS IN THE EVALUATION OF FOOD SAFETY PRIORITIES

DR. WEINSTEIN: Good morning. Since I have been pressed into service early, I have an excuse if I mess up. I am also a novice at Powerpoint presentations, although this is my third one this week and so far, so good, so I am optimistic.

I am a relative fish out of water -- pardon the food metaphor -- in this group. Usually what I do is talk about things like how to make decisions to give drugs to lower cholesterol and how often to screen for colon cancer and with what kinds of tests. When I give those kinds of talks and speak to audiences in managed care or in the pharmaceutical industry or NIH, I usually talk about QALYs as a way of comparing apples and oranges -- again, I am a little timid about using those kinds of metaphors in this crowd, but basically that is what this is about.

It is one of many methods, and you are going to hear about three methods, I guess, during the course of the rest of the day, for making these kinds of comparisons, basically, the valuation end of risk analysis. We heard about the risk-assessment task, which is a daunting task, as you have heard, whether you are talking about microorganisms or chemicals.

The valuation task actually is also very daunting, but there has been a lot of progress made in this area over the past decade or so, some motivated from the medical decision-making fields, some motivated from the environmental decision-making field, and I think there is a role to play for these methods.

I think the purpose of the next three presentations is to discuss three related but different approaches to try to reflect the values of the public into decision making involving relative risk analysis.

[Slide]

I am going to be talking about a measure called quality-adjusted life years. I really do not know how many of you are familiar with this measure, but I am going to assume that you are not very familiar with it (those of you who are, please bear with me, I will try to be brief). The idea here is if you are concerned about a health consequence, whether it be a case of cancer or a case of food-borne infection, you ought to be able to compare the health impact of those consequences as valued by the affected population.

The measure of quality-adjusted life years is essentially a measure of years of life weighted to reflect the value of health status. It is designed to reflect both effects on mortality, loss of life or reduction in length of life, and a change in quality of life as might result from either chronic disease or a short-lived acute episode of illness.

The weights that are attached to the various health states range from one, to reflect perfect health, to dead, represented by zero, with intermediate health states, such as the time spent living with a disease as between zero and one. The idea, and this is really the point, is that the weights are designed, if they are elicited properly, and can reflect the preferences of the affected individuals regarding those health states.

To satisfy their purpose as a way of weighting different health states, these weights must be based on preferences, so we rely a lot on surveys of individuals who might be affected by these conditions to elicit those weights. They have to have an interval property in psychometrics such that the ratios of differences are meaningful, so a difference of 0.1 on that scale can be compared to a difference of 0.1 somewhere else on the scale to be called equal and, moreover, that a difference of 0.2 can be interpreted as twice as bad or twice as good, depending on which direction you are going, as a difference of 0.1.

Finally, in order to allow this comparison between mortality and morbidity effects, the scale has to be

measured or transformed on to the particular interval scale where the reference point death has a score of zero and the reference point optimal health has a score of one.

[Slide]

QALYs can be used in priority setting and have been used in priority setting in a number of contexts. One such type of study is a burden-of-illness study, where the idea is to estimate the total number of QALYs lost from a disease or a health risk. This could be used, for example, in the setting of NIH trying to decide how to allocate its resources to do research on various diseases, or in cost-effectiveness studies, where the purpose is to estimate the efficiency with which resources can be used to gain quality. Here we are talking about interventions, actions that can be taken with the intent of improving health or reducing health risk. QALYs have been used a lot in decision making in medical care, in helping to develop guidelines for medical practice. You see a lot of studies using QALYs and cost per QALY in the medical literature. It has really been quite a revolution in the last decade or so. I do not know how many of you ever look at the pages of journals like *The Journal of the AMA* or *The Annals of Internal Medicine*, but you will find cost-effectiveness studies using QALYs quite frequently in those journals.

They are becoming, I will not say they are overwhelmingly influential in medical care, but they are becoming more influential. QALYs are used to some extent in setting priorities for preventive health services, for example, at CDC -- I do not want to steal too much of Anne's thunder; she may say something about that later.

I just came from a conference on pharmacoeconomics, where we learned about various HMOs that are now using cost-effectiveness analyses, some using QALYs, to guide decisions about reimbursement of pharmaceuticals and other medical technologies.

In the government, the QALY idea has been used to set, or at least recommend, priorities for the development of drugs and vaccines, and I will comment on a study done at the Institute of Medicine that I was involved in recently that did that.

[Slide]

The technology of cost-effectiveness analysis has gotten to the stage, in health and medicine, at least, where it was felt important to have some rules and regulations, some guidelines (this came up earlier with regard to risk assessment, whether there ought to be some sort of standards for how to do risk assessment). The same is true of cost-effectiveness analysis using QALYs.

To that end a panel was convened, the Panel on Cost Effectiveness in Health and Medicine, by the U.S. Public Health Service in the mid-1990s, which resulted in a set of recommendations presented to the U.S. Public Health Service for cost-effectiveness analysis as it applies to health and medical interventions.

One of the recommendations from that panel, which I co-chaired, along with Louise Russell from Rutgers, was that for a reference case analysis -- "reference case" is a buzz term for a set of standards (not that you have to adhere to these standards but it would be good if you at least paid attention to the standards and maybe included a set of analyses that followed these standards) -- the incorporation of morbidity and mortality consequences into a single measure should be accomplished and this should be done using quality-adjusted life years.

[Slide]

The question here is can QALYs help guide food safety priorities? I am going to briefly review the methodology itself, describe some applications of QALYs in health care, including the Institute of Medicine vaccine priorities report, and some of the challenges that remain in estimating QALYs in the health care context, then describe some of the theoretical limitations, what QALYs do and do not reflect in principle in terms of values regarding risks, some of the advantages and disadvantages relative to other methods, which you will be hearing about later, and some of the research needs for the methodology.

[Slide]

One of the advantages of the methodology is that it keeps separate the issue of estimating probabilities from estimating the values. It is consistent with a view of risk analysis that says there is risk assessment and then there is risk valuation, and this is about risk valuation.

[Slide]

The data requirements to get to an assessment of the number of QALYs or the probability distribution of the number of QALYs lost or gained are fairly substantial. Say you have an estimate of the risk of some disease, say, a food-acquired infection, Listeriosis, for example, or a chemical-induced cancer, you need some estimate of the disease incidence that you are talking about, or the increase or decrease in the disease incidence, but you also need some estimate of how this affects health outcome, so you need some estimate of the disease progression. What does this really mean for people if they get Listeriosis or if they get

cancer, say AML?

The first step is to estimate the lengths of time spent in different health states as a result of this incident case. The data sources potentially available for accomplishing this task are experimental studies, clinical trials, for example, natural history studies based on observation of patients with the disease observed over time. Expert judgment always come in, although I see it higher on the list in the environmental context than usually in the medical context; medicine is influenced a lot by the evidence-based philosophy that if there is no direct evidence, you had better be careful. But sometimes you cannot avoid expert judgment. Then, of course, models are used a lot in medical risk assessments.

[Slide]

So there is a risk-assessment component to using QALYs. The risk assessment, then, would not stop with an estimate of the number of cases or the incidence rate for Listeriosis or cancer. You would have to play out, then, okay, what does this mean for people in terms of their health status, in terms of their mortality and their longevity.

Having then estimated the consequences in terms of length of time spent in different health states, we come to the valuation task, which requires health-state-specific utilities that reflect the preferences of citizens, tradeoffs between time and risk of health states, but scaled so that a value of one corresponds to perfect health and zero corresponds to dead.

[Slide]

There are a number of ways to elicit those preferences for health states. A health state might be a day or a month spent with some health problem, severe diarrhea, for example. How would you get at that? A couple of methods, for those of you who are not familiar with them, are the time-tradeoff method and the standard gamble method.

[Slide]

The time-tradeoff question goes something like this. If you would do this in a survey of some citizens, you would ask them, "Suppose you faced some number of years in some health state," which you would describe or perhaps which they are experiencing at the time, and you would ask them, "How many years in perfect health," some number Y (presumably less than N), "would you be willing to accept if you could have perfect health instead of being in this health state S?"

Then the estimate from the time-tradeoff method of the utility association with health state S would be the ratio of Y to N. There is a whole theoretical literature defending that, going back to the fundamentals of decision theory and utility theory, which I definitely will not go into this morning.

[Slide]

Then there is another method, called the standard gamble, where you get the risk aspects of the health state, or the valuation of the health state. Here the starting point is imagine that you are going to spend the rest of your life in some health state -- or some finite time, it does not have to be the rest of your life -- and the question is, "What chance of death would you be willing to accept in order to improve your health to perfect health?" In these surveys people are willing to take chances of death, depending on how bad the health state is.

The utility here to find is one minus the probability of death they would accept, so if they willing to accept a very small probability of death, that means the health state really was not so bad, so one minus P would be very close to one. If the health state is really bad, they would be willing to accept a very high probability of death, maybe close to one, and, in that case, one minus P would be small, so the utility would be very low.

[Slide]

The cost-effectiveness panel from the U.S. Public Health Service recommended that the preference weights that should be used for public policy and broad resource allocation decisions and priority setting should be based on the preferences of the general population in a general context. The recommendation is that these weights should be linked to a health state classification scheme that reflects the domains or attributes of health that are important for the particular analysis that you are considering.

Ideally, the health state classification scheme should be generic. It should include a broad range of health states so you could really get all the apples and the oranges and the pate and the ice cream into the same scheme. This is a very strong recommendation. So where does that leave us?

[Slide]

Fortunately, there are out there a number of generic health-status classification schemes that have been weighted according to the preferences of some representative, or allegedly representative, sample of some community. One of these is the Health Utilities Index, developed in Canada but used a lot in the United

States. It has eight domains of health in it, and the weights were elicited from their community samples using primarily the standard gamble.

There is a scale developed in Europe, called the EuroQOL. It has only five domains, using the time-tradeoff method to get at the weights. The Quality of Well-Being Scale is the oldest one, developed in California back in the early 1970s, that has four domains. One of the limitations is that the weights are based on a rating scale, an ordinal scale, rather than a cardinal measure of intensity of preference. It has been criticized for that and some of the weights do not necessarily have the kind of face validity that you would want for these weights, so it tends not to be used as much as it used to be.

More recently, some of you may be familiar with the SF-36 as a measure of quality of life. It was developed at the Rand Corporation for use in a study called the Medical Outcomes Study. The SF-36 was not preference-weighted, so an economist in England, named John Brazier, decided it was time to make it preference-weighted, so he took six items out of the SF-36 and then went out into the field and got the citizens to estimate their preferences for these six domains of health, using the standard gamble method.

[Slide]

Just as an example of what one of these schemes looks like -- this is the Health Utilities Index, the most recent version of it, which includes dimensions for vision, hearing, and speech, ambulation and motion, dexterity, cognition, and pain and discomfort. Which of these might be affected by a food-borne illness depends on the illness. If you are talking about arthritis, it might affect ambulation and dexterity, possibly emotion, almost certainly pain and discomfort.

If you are talking about cancer, it could affect many of these same things. If you are talking about dementia, neurological effects, you might get into some cognitive deficit, which would be captured within the HUI.

[Slide]

I do not expect this to be absorbed, but this is just to show you the mathematics of calculating the utility, so this is based on multi-attribute utility theory. You have to take my word for it that it is firmly grounded in theory; they did not make these numbers up. These are based on the elicitations from their sample population and, together with a few assumptions about the way these preferences hang together across domains, then you can come up with a scoring system that allows you to assign a weight or a utility to all possible combinations of these attributes.

For example, somebody who is at vision level No. 1, and hearing, speech, and ambulation -- say the only thing wrong is that he is in pain level 4, so he has a health state where everything is okay, except he is in pretty bad pain. You plug in b1 through b7 equals one for all the attributes except pain. Pain, b8, would have a value of 0.77. You plug it into that formula, which I still have not corrected. I presented this slide in Jim Hammitt's class and he pointed out that b8 is missing, so just put a b8 in there. So you plug in 0.77 for b8 and out pops the utility for that health state as elicited from this population of citizens who responded.

[Slide]

There are other preference-weighted health classification schemes that are disease-specific -- I should mention them. For example, if you interested in only comparing, say, interventions that affect the risk of trauma, then you can use something called the Functional Capacity Index, which was developed at Johns Hopkins. If you are interested only in cancer, you can use something the Quality of Life Index, which was developed in Boston at the Dana Farber Cancer Institute.

[Slide]

I should mention DALYs. You have probably heard more about DALYs than QALYs. Depending on how you feel about DALYs, you might say, well, DALYs are basically the same thing as QALYs, with a few little differences, or you could say they are quite different and you could say they are quite different and much different, or quite different and much worse.

I am sort of in the school of thought that it is really just a different flavor of QALYs. It is the same thing, except the scale is upside down and you are measuring losses rather than gains, so zero is good and 100 is bad. There is only a single attribute there; it is called disability. It is a six-level attribute. The weights are not obtained from the community, they are obtained from experts.

In addition to the health status weights, there is an additional superimposition of weights per age, so that people at older ages, people older than age 35, as you get increasingly older than 35, the value attached to one year of life goes down -- consistent with Bob Buchanan's comment.

(Laughter)

I am with Bob on that one, I do not like that.

Similarly, at the other extreme, infants, neonatal years of life would be valued a lot less. This is going on

inside the DALYs, it is not going on inside the QALYs.

Also, there is a little glitch in there. It assumes that people whose deaths are prevented would have survived to a normal life expectancy in perfect health, so it sort of overvalues life saving relative to quality-of-life improvement, because, in fact, the lives you save are not going to have lived their lives at a normal life expectancy in perfect health.

[Slide]

These are some examples of utility weights based on a number of surveys, preference surveys. There is one called the Beaver Dam Health Outcome Study that went directly into the community of Beaver Dam, Wisconsin. They did not use a generic health-state classification scheme, they went directly to the people who were experiencing these conditions and people who were not and, using the time-tradeoff method, got their utility.

For example, they found that, on average, people who had arthritis, who said they had arthritis, had a time-tradeoff utility of 0.816, whereas people without arthritis, on average, had a utility of 0.9. You can see that these various conditions subtract anywhere about between 0.1 and 0.2 units on the zero-to-one scale per year spent in those states.

The National Health Interview Survey was used together with a generic system called the HALex to assign weights to a wide range of conditions. This was done, by the way, by the government in the NCHS. You can see here these weights come out a little bit less. These were based not on the time tradeoff but on a set of expert weights assigned to various attributes of these conditions.

Notice that constipation may be a food-related consequence of some interest. It has, to me, an extraordinarily low weight, meaning people value it as pretty serious relative to some other conditions that you might not think would be as serious. This is per unit of time, though, so if you are chronically constipated, it would be almost as bad as arthritis or, actually, worse than arthritis, on average, as perceived by the people who responded to the National Health Interview Survey.

[Slide]

Just to show you that you can do this by age, this shows you that the weights for these conditions, from the Beaver Dam study, at least, were not very sensitive to age, so the value people place on arthritis, on having a health state of arthritis, whether you use the Quality of Well-Being Scale or the time-tradeoff direct assessment in the affected population do not vary very much by age -- there is a little bit of a drop-off with age, but not that much.

[Slide]

The Institute of Medicine study is a study that was done using an incidence-based burden-of-illness methodology but where the outcome was not dollars but QALYs. Public Health data were used to estimate disease incidence and mortality. Years of life lost per death were estimated from life tables and weighted by average age-specific health-related utilities for those diseases.

[Slide]

The consequences for health status were estimated from the literature and by experts, how many years you spent with chronic pain or with disability of various kinds. The health states were then mapped by clinical experts who were consultants to or members of this IOM committee into the health utilities index classification. The weights of the sort that I showed you for the HUI were then used to calculate the expected loss of quality-adjusted life years due to each disease, and then the potential QALY benefit estimated by incorporating the expert judgment regarding the efficacy and utilization of the possible new vaccines, so the vaccines here are analogous to a regulation to reduce the tolerance for a pathogen, for example, so the experts were used for that purpose. The rest of it was done using fundamentally Public Health data and survey data on the preferences.

[Slide]

Let me conclude, then, with some of the challenges. What does this have to do with food safety priority? First of all, to use QALYs you have to have a detailed assessment or some model of the consequences of each disease for health status over a lifetime for the affected person; otherwise, you cannot assess the number of life years and changes in health status necessary to calculate QALYs.

It requires, secondly, mapping from disease-specific health states to a set of preference-weightable health states, like the generic scale in the Health Utilities Index, where you have health status described in terms of pain, cognition, emotion, dexterity, ambulation, and so on. Now, this mapping may be available from the literature. There are some studies, like the Ontario Health Survey, where they actually administered not only the usual kinds of questions you would find in the National Health Interview Survey, but also the Health Utilities Index, so you can actually see how people with arthritis view their health-related quality of

life as measured in the Health Utilities Index, so you can get a utility for people in Ontario who have arthritis.

Or the mapping can be done by clinical experts or it may be worth, depending on how important the food safety problem is, the condition that would result from it, doing a population survey of persons with the condition of interest. For example, people with Listeriosis, you might want to find out how bad that is by administering a Health Utilities Index to them to find out how their health status is, and then use community preferences, or you could use direct preference assessment in the patient population.

[Slide]

A few theoretical limitations that imply risk neutrality; that is to say, a 50-percent chance of losing 10 QALYs is equivalent to losing five QALYs for sure. That is true of a lot of valuation techniques, it is true of QALYs. There is a way of adjusting for that, which, in theory, can be done -- nobody ever does it, maybe it is time to start doing it -- but it has not really been worked out in detail.

Here is an important limitation and this is where I will stop, because I think this serves as a transition into the other methods. Unlike the other methods, willingness to pay, various other methods of eliciting risk attitudes and perceptions, QALYs ignore the psychological attributes of risk and risk reduction. It focuses on the health consequences. It does not really focus on the process of how you get to those consequences. To the extent that there is loss of peace of mind along the way, every time you eat egg salad you worry about you might get Salmonellosis, that is not reflected here. It is only the consequence once you get Salmonellosis, how bad it is and what it does to you. Whether the risk is voluntary or involuntary is not reflected in this method. Dread. Those kinds of things are not reflected.

It need not, although you could look at QALYs in different populations or subgroups, it does not force you to consider distributional equity. On the other hand, QALYs are basically democratic. A QALY is a QALY (there is a paper in the literature that has that title). It does not, for example, weight health consequences by ability to pay, so willingness to pay can be criticized -- I do not want to get too much into that now -- because it is known to be related to wealth as a basis for the response. QALYs are not related to those things.

[Slide]

What are the advantages? The method incorporates scientific evidence and expert judgment about the health consequences of exposures -- that is the risk assessment part. It is not flawed by perception. If you went directly to a consumer and asked him, "How do you feel about the risk of Salmonella in eggs," he does not really understand, necessarily, what the risk is, what the probability is that he is going to get a case of Salmonellosis, let alone the consequences of Salmonellosis, what are the chances he is going to die of it as opposed to getting diarrhea for a period time, so you use the experts and the data to do the risk assessment.

On the other hand, the method incorporates measured community preferences regarding the relative disutilities of the health effects, so you go to the people who are affected, you get their perceptions of the values, you use the data and the experts to get the estimates of the probabilities.

I am sorry I ran over, but I did not have time to prepare, so that is my excuse.

(Laughter)

DR. NARDINELLI: I am standing in for Richard Williams. Richard, every Monday, prints out his outlook calendar and sees his week, and this week he noticed that he was scheduled to be here this afternoon and also scheduled to be in Portugal. One of us had to go to Portugal, one of us had to be here.

(Laughter)

I am going to give his presentation as closely as I can to what he would say, though I will not promise that I will not deviate a little bit from Richard's views.

First of all, let me say I enjoyed the presentation and the paper, which contains a lot more information than Dr. Weinstein had time to give, and that our agency, particularly in the Center for Food Safety and Applied Nutrition, we use QALYs quite a bit. We use them as part of our method of estimating the benefits of food safety regulations, so we are very familiar with them and we are always glad to know that more and more new information is coming out on QALYs and there are QALYs for more and more different kinds of conditions and situations.

But using QALYs for putting a value or benefit on food safety regulations is much more complex than using them for medical interventions or to estimate the cost effectiveness of medical interventions and, again, this is another key point, cost effectiveness. We do not use cost effectiveness but benefit-cost analysis. For food safety regulations this is really, we think, a much better way to go about it.

What I am going to do now is talk a little bit about how we go about doing food safety regulation,

prioritizing public health expenditures in the small, if you will, and show where QALYs fit in and how we use them.

[Slide]

Let's begin at the beginning: public health base-line risk. We start with the risk estimate, what we might think of as a very crude risk assessment, not the full-blown risk assessments you have been hearing about earlier, but some unbiased risk estimate: What is the problem? Again, this is a risk estimate, not a regulatory risk-management decision. As Richard would say, that means no conservative assumptions to begin with, just what is the problem?

[Slide]

We then look at ways we can deal with the problem, with the food safety problem. We have six options. We look at all of those and then we notice that we call our lawyers and they say that we have to look at what we can do. As has been mentioned earlier in this conference, the Food and Drug Administration has certain things it can do and things it cannot, there are restrictions on what our regulatory powers are. So we do that.

What we find is that we have to eliminate options two, three, and five, so we are down to options one, four, and six. Those are the ones, then, that we try to estimate the costs and the benefits for.

[Slide]

In order to do that, in order to estimate the cost, to be able to get an idea of the cost effectiveness of these options, it is much more difficult in the case of medical intervention, because the costs depend, for one thing, on the reaction of consumers and the change in the risk will also depend on that. We can put out a food safety regulation but how will people react to that?

Of course, how will producers react to that? For example, if you institute a particular form of control in processors, the processors could do one of three things -- let's suppose it is a performance standard -- they could adopt the performance standard. An alternative, however, is that they could go further and have a stricter standard. Another alternative is they could decide we do not want to be in this business any more, they can shut down.

These are different ways that the costs have to be estimated. We have to wonder what is actually going to happen. So the cost, then, is how the value of resource allocation changes in response to the option we are estimating. It is not just a take it down off the shelf and what is the cost of this particular program.

Suppose we get the decision, option four, it involves certain reactions by consumers, reactions by producers, we estimate the costs of those. Then we estimate the QALYs saved and we would use some of the methods that Dr. Weinstein has depicted. Then we put a value on that. Typically, we would use willingness to pay to save a QALY, and there are many different values of that (that is actually the subject of some of the talks this afternoon, so I will not get into that).

The important thing is to add a dollar valuation. To put it simply, we multiply the number of QALYs times the dollars per QALY, and that is our estimate of benefits, and that is our benefit-cost analysis. If the decision is option four, that typically will be the one that has the greatest difference between benefits and costs -- remember, we are not talking about a ratio, we are talking about the difference.

[Slide]

This is one of Richard's favorite slides, so I have to throw this in. The costs and benefits go into the decision but they are not the decision. All of these other things matter: law and precedent; the uncertainty involved -- the risk managers in the Food and Drug Administration and higher up may often take into account the uncertainty; and, of course, we look at all of the sciences that we draw on, the nature of the risk, not just the numerical value but other factors, some of the things that Dr. Weinstein indicated are not included in QALYs, like dread, for example, or public perceptions. All of these things could fit into the decision.

[Slide]

Let me just follow up by saying, then, that we use benefit-cost analysis to bring QALYs into our decision-making process because we think of the problems in cost effectiveness. One is that cost effectiveness does not allow all costs to be considered -- this has been mentioned. Cost effectiveness cannot allow us to compare different food safety policies.

If I may, for just a second, speak a little bit of economics jargon, economics jargon you have all heard, cost effectiveness, let's suppose dollars per QALY, we can use that to trace out something like a marginal cost or supply curve of efficient allocations. What is the equilibrium allocation? Then you need the demand side. That is where valuation, perhaps through willingness to pay, comes in.

Finally, part of our job is risk communication, talking about our decisions, defending them. Not everybody

knows what a QALY is, but dollars, net benefits, or benefits minus costs in dollars, is something people do understand. So this is why we use QALYs to generate dollar benefit estimates.

Let me conclude (and I will actually beat the timekeeper here) by saying we do think this is an extraordinarily useful tool, we are glad to see more and more being done on it. We follow Dr. Weinstein's and others' publications on this regularly and with great interest.

Richard has come up with a symbol for a QALY. We urge you all to take QALYs seriously and find ways to use them.

Thank you.

DR. HADDIX: Now I have to think of something to say, because it figures that Richard and I were on track on this and, Clark, you just said every single thing I was going to say. This is a great opportunity for me, because I get to put back on my CDC hat and, hopefully, Art Liang in the audience, will not shoot me for doing that, even though I am at Emory University now.

I was one of the first full-time economists on staff at CDC, so I have to confess right off that I have been totally brainwashed by the epidemiologists, although, when things got really bad there -- they were not very comfortable with economists in the beginning -- I would tell them epidemiology was just economics without the dollar sign.

One thing I did want to mention today, because I was actually going to talk about just exactly what Richard thought was important to talk about here, and that is the use of QALYs in cost-effectiveness analysis and then the use of cost-effectiveness analysis in evaluating food safety interventions, which was something that we at CDC attempted to do all the time.

Something else that Milt pointed out was that you can use QALYs for two different things. In addition to cost effectiveness, you can also use them for burden of disease. I think that that is a very important point that could be explored further, so I just want to mention an exercise that we at CDC did and are just now publishing -- in a very timely fashion we started it six years ago.

In this particular case, because we were creating new methods for prioritizing health problems, these methods and this study went through lots and lots of review and lots of lots of corrections. CDC is a mess, because you write up a paper and you address a problem that occurs in every single center at CDC, you have to send it to everybody to clear and get everybody to agree on it. It can take a while.

One of the things that we did with QALYs there, which could be used in food safety, is we set about the task of prioritizing the clinical preventive services recommended by the Task Force on Clinical Preventive Services. All I want to say is just a note about the methods. We came up with a scoring system and we used two criteria. Both of those criteria included QALYs. The first criterion was the burden of disease, and it was the preventable burden, so, first, we estimated in QALYs the burden that could be prevented if the preventive service were delivered at optimal levels.

Second, we used the cost effectiveness of that particular clinical preventive service. We assigned scores depending on the value that each one of these criteria gave and ranked these and were able to come up with the top -- we would like to say the top 10 but, as it turned out, because many have the same score -- 14 clinical preventive services. As it turned out, eight of those services are currently being delivered to fewer than 50 percent of the recommended population. It was a very, very innovative method for prioritizing interventions. I think the use of these methods could be explored in food safety.

Before we can do that, I want to reiterate some of the things that Clark mentioned about once you get QALYs, then you can put them into cost-effectiveness analyses, but cost-effectiveness analyses in food safety present some particular problems.

[Slide]

First of all, as I think Clark mentioned, is the challenge of measuring the net cost of the intervention that you have in mind. One thing nice about CDC, the limit it does not have that FDA does have is we never had to consider whether it was legal or not.

(Laughter)

I mean, seriously, think about it. We did cost effectiveness of needle-exchange programs to prevent HIV. That was not a legal intervention at the time, but because we are not a regulatory agency, we are not as bound to looking at only options that could be solved through regulation. So we have a little bit more flexibility but, still, when you are looking at the cost of interventions to prevent food-borne illness, it is much more of a challenge than doing cost effectiveness of health interventions that are delivered in a clinical setting and, compared to food safety problems, are fairly easy to measure -- think about when it was being proposed to ban the importation of unpasteurized cheeses, particularly from France, measuring the net cost of such a strategy.

Then there are food-borne illnesses such as cyclospora, where the interventions at this current time are fairly unclear and may involve increased surveillance, early outbreak detection, and a temporary importation ban. These become somewhat of a nightmare to measure and they certainly did at CDC when we were attempting to do it.

Then, as Salmonella has been mentioned in the past, it has been very, very challenging to cost out changes from farm to table, a strategy that involves changes at many points along the way, including changes in animal production practices, changes that may involve substitution effects on the parts of consumers.

Measuring the costs of strategies is a particular challenge.

[Slide]

Another challenge is describing the base-line comparator. When you do cost-effectiveness analysis, you are always comparing it against the current base-line strategy, the current standard of care, and it is not that we do not do anything at the current point in time.

At CDC, certainly, one of the things that we do is we have surveillance systems in place to detect outbreaks and then we have outbreak investigations. Conceivably, if you have a program that is going to reduce the incidence of food-borne pathogens, you are also, hopefully, going to reduce the number of outbreaks, so you actually have to be able to have a good description of the current operating procedures and their costs. I would like to point out that in this particular case it is not only the cost of illness of individuals who get sick that we are trying to avoid but, oftentimes, we are trying to avoid significant public health costs to respond to outbreaks. One, for example, was the cost of an hepatitis A outbreak in Colorado, where about 95 percent of the costs associated with that outbreak were public health response costs, not costs of infections of hepatitis A.

[Slide]

Then there is the whole issue of capturing the benefits and do QALYs capture all of the benefits associated with a food intervention? I think both the previous speakers have pointed out that, no, they may not, that in addition to the quality of life that is saved by preventing illness you also decrease outbreak-related anxiety, which can be tremendous, and when outbreaks are multi-state outbreaks and are fueled by the press, outbreak-related anxiety spreads far beyond where the outbreaks are associated.

For example, with cyclospora, even though the outbreaks were limited to certain cities, there was outbreak-related anxiety on the part of anybody eating a raspberry, which leads to the final one, which is what are the benefits of consumer confidence in the food supply and does cost effectiveness capture that? The answer is probably not. When we are thinking of cost-effectiveness analysis we have to take into consideration is our denominator here really capturing all of the benefits associated with an intervention?

[Slide]

Finally, the last thing I wanted to say is how does cost effectiveness fit into decision making for food safety strategies? Cost effectiveness, I think, as Clark mentioned, tends to be a supply-driven measure. It does not usually factor in the demand side, where cost-benefit, particularly when using willingness-to-pay methods, may look at consumer preferences for particular interventions.

Finally, cost effectiveness is currently not an acceptable method for the regulatory process, unless you do the conversion that Clark mentioned, where you actually convert those QALYs to dollar value so that you can report cost-benefit analysis.

That is it, and I think I came in on time. I think I have to say that at CDC one of the things they train you in very early is -- I mean, they get the hook out if you go over. Thank you.

DR. HOFFMAN: We will take about 15 minutes for questions and then I will let you go for lunch.

GENERAL DISCUSSION

DR. ROBERTS: Tanya Roberts, Economic Research Service. I guess I would like to ask Milton to talk a little bit about what do you do when you have a fetal injury, how do you value that using QALYs? That is one of the things you really run into -- or if you have a spontaneous abortion or you have a fetus that is born severely mentally retarded, what do you do with QALYs in those cases?

DR. WEINSTEIN: That is an easy one.

(Laughter)

No, that is a thorny one. It depends on what the counterfactual is, of course. If you are talking about a child who is born with a deformity or disability that he or she would not otherwise have had, then it is relatively - - I will not say it is easy but it fits well within the framework, because you can look at the difference between the quality-adjusted life years without the insult and with, and that is the loss, it is the difference in that quality-adjusted life expectancy.

If you have talking about spontaneous abortions, then it is not well handled, you are sort of outside the

construct, because it is not clear whose lives you are talking about.

DR. ROBERTS: But this was a child who would have been born normal and is now aborted because of the infection.

DR. WEINSTEIN: Yes, it is a thorny issue. I do not want to stick my foot too far in my mouth. You can talk about replacement children, and things like that. Jim and I were talking the other day, if you save the life of somebody either during her fertile years or before, you could think about the lives and quality-adjusted life years you are saving and future generations would not have been born. I think you are sort of getting beyond what the methodology is really good at doing.

While I have the microphone, I would like to just comment on the notion of putting a dollar value on a QALY at the end and then doing a cost-benefit analysis. You are really doing the same thing that you are doing in a cost-effectiveness analysis, except you are making explicit what the cutoff is, so you are getting a cost-effectiveness ratio, you compare it to some number, whatever number you are going to use to put a weight on the QALY, and then you are going to decide whether the ratio is above or below that, unless you use a different number, depending on the context. You might want to value QALYs at a different value for one exposure than another, and in that sense you are doing some different.

DR. HADDIX: But then you can use it to go along with the regulation and that is why we do it.

DR. RODRICKS: A quick question on the community preference approach to establishing QALYs and getting community preferences. I am surprised, and maybe you covered this and I did not really follow it, that there are no very large differences among individuals in how they value different disease states, according to age, according to gender, perhaps, socioeconomic status, and how do you handle that?

DR. WEINSTEIN: Until you got to the last phrase of your question, Joe, the data support you. There are huge variations across individuals in preferences for the health states, for example, the health states in the HALex score scale or the Health Utilities Index. However, most of the data suggest that there are no systematic variations with respect to demographic or social characteristics; the preferences are remarkably stable by age -- I showed you some data on that: the same for men and women; the same for rich people and poor people; the same for people of different racial and ethnic groups.

They are very similar across countries, at least in the countries where these kinds of preference assessments have been done. This is currently some work that, actually, Chris Murray, who is identified with the DALY approach, is using, person-tradeoff and time-tradeoff methods in developing countries to see whether they get similar responses.

The data suggest that there really are not systematic differences; at least we have not found any.

MS. LILYGREN: I am Sara Lilygren from the American Meat Institute. I realize we are talking about values associated with costs and benefits, but I wonder if any of you could comment on whose job it is, if any, to go back and look, after a regulation has been promulgated, supported by the work that you all and others have done to establish the cost and the benefit of the regulation or the intervention, whose role it is, if any, to go back and evaluate whether your theories or your hypotheses actually played out?

I am thinking specifically now about the cost-benefit analyses that came out when mandatory nutrition labeling was promulgated in the early 1990s. Today, 11 years later -- and I remember looking at it and how many lives would be saved by nutrition labeling, et cetera, and how many lives would be extended, and now we have a worse obesity problem in this country than we had before -- whether it is anyone's job to go back and look at whether the hypotheses and the evaluations that were made about lives saved actually came to pass.

DR. NARDINELLI: That was a Food and Drug Administration regulation. I guess I am stuck with the job of answering it. To the first question, is it anybody's job to go back to explicitly look at it? The answer is no, but people have, both within our agency and many people outside of our agency, and we have found out a lot.

One of the things you mentioned -- in fact, there has been some published material on this fairly recently -- is this continued increased in obesity, which we had not expected. On the cost side, if I can refer back to my presentation, on the producer-response side, one of the things we also did not anticipate was the large amount of product reformulation in response to having to put the nutritional values on the label, and we are starting to look at that.

So although there is no office of reviewing old cost-benefit analyses of major regulations, it does go on. As I said, we have a great deal of help from our friends in the think tanks and universities who are doing some of these analyses.

DR. WEINSTEIN: This is actually done quite often with medical interventions. One example is there are some cost-effectiveness analyses of cholesterol-lowering drugs that were done before there were large-

scale clinical trials and before the epidemiologic data, obviously, reflected the increase in the use of these drugs. People have gone back and looked at the public health data on coronary-heart disease mortality, and so on, and found that, actually, the public health data track pretty well with the predictions based on the changes in the utilization of these drugs, so there is some evidence that maybe the predictions -- not that they had to be right. You know, when you are making these risk assessments, you cannot be held to the standard that you have to be right; you are doing the best that you can do at the time. It is useful to go back, so that you know that the next time that you do it, you revise your approach.

Also, the National Cancer Institute has a program to look back at the projections of the effects of various screening programs for colon cancer, breast cancer, to see whether the public health data have actually generated evidence that increases in the utilization of screening mammography or occult blood testing and sigmoidoscopy for colon cancer have resulted in reductions in cancer mortality. That is ongoing now, funded by the National Cancer Institute.

DR. HADDIX: I think the question raised is a very, very important question and it is one that we do need to consider, to try to make the time to go back and do some of those. At CDC we did frequently go back and do those and I think one of the things that it drove us to do at CDC was to develop models that anyone could use and that would be available and transparent.

For example, when the rotovirus vaccine came out, we had done a cost-effectiveness model of that and had it set up in such a way that when major problems occurred with that vaccine that we did not anticipate, we were able to go back and plug in new numbers and, when we got new disease burden numbers, we were able to go back and do that, and do that fairly quickly.

The facts of life in an agency like CDC is that we have always moved on to the next problem and we always have to have the answer yesterday, and it really is very helpful if we can make it as easy as possible to have models that we can use to easily reassess things. That is a challenge to all of us who are developing those models.

DR. HAMMITT: Jim Hammitt from Harvard. I wanted to make a comment to this topic. I think it is a very important point that all of us doing these kinds of models try to work to go back ex post to see how accurate our models were. Some of us have done that in some contexts. People in this building, Winston Harrington and Richard Morganstern, have looked at that for the costs of some environmental regulations. I just to highlight that it is not necessarily easy or even possible to do, because if you implement a program, you do not know what would have happened had you not implemented the program. We are always trying to estimate incremental costs and benefits and you observe only one state of the world, you do not observe the counterfactual with which we want to compare.

DR. BARNES: Donald Barnes from EPA's Science Advisory Board. I wonder if, just briefly, the panelists could make some comment on what philosophical objections there might be to the use of QALYs and what are the responses to those objections?

DR. WEINSTEIN: I listed some of them at the end. There may be some valued consequences of a regulation, or the lack thereof, that are not reflected in the way people feel about health status. QALYs are linked to health status and there is more to it than health status. The worry about eating something because you are not sure if you are going to be the one who is going to get that, you are going to be the one in 100,000 who is going to get something. That in itself is not reflected in the QALYs.

The fact that every QALY is considered equal, regardless of who gets it, that a 50-percent chance of losing 10 QALYs is the same as losing five QALYs for certain -- the risk neutrality assumption. Those are all limitations.

The responses to the objections, at least the one that I always give and the one that is sort of the official position of the Panel on Cost Effectiveness in Health and Medicine, is that you should never use QALYs or cost per QALY as the sole criterion for any decision. It is one of many inputs. The Richard Williams slide, I think, is apt, it is just one piece of information that decision makers, I think, should have, but certainly should not dictate the decision.

DR. HADDIX: Just to follow up on that, having faced this, I think these issues that Tanya raised, too, on how do you treat spontaneous abortions and fetal deaths, is a big one. I think we faced this when I did one of the examples for the Panel on Cost Effectiveness in Health and Medicine, which is a project I worked with Richard Williams on, on folic acid fortification acid to prevent neural tube defects, and we were dealing with the issue -- one of the consequences of neural tube defects is spontaneous abortions as well as fetal deaths. The question came up, who has standing in economic evaluation and cost-effectiveness analysis, whose QALYs do you include?

Clearly, with many diseases, many more people than the patients are affected, and their QALYs. How

about the QALYs associated with caregivers, a spouse caregiver of someone with Alzheimer's? Or the QALY of a fetal death? We really struggled with those and, again, the way we present cost-effective analyses is that in addition to a cost-effectiveness ratio to look way out to quantify as many of those outcomes as possible. It is very challenging to do it all then come up with interventions that are comparable with other studies. It is a challenge that I think we are still facing.

DR. NARDINELLI: Most of the philosophical problems I think have already been covered, so I will just say what we do for some of these other things is we try to add them to the value of the QALY. If there is something that a QALY does not cover and we feel preventing it is an important benefit from our food safety regulation, we add it. It is a little ad hoc, but it gets it in there.

The one, I think, philosophical problem that I do not really have a solution for -- I think it varies from analyst to analyst -- is this question of dread. One thing that always seems to show up is that dread often does not really correlate very well with an objective QALY measure of a particular condition. It can, but it often does not. There is no real solution to that other than to go strictly to willingness to pay without regard to QALY, a solution which we have been hesitant to use.

DR. WEINSTEIN: Or education. There may be some reason it is immutable or it may be that people, if they were informed about the evidence, might change their dread.

DR. NARDINELLI: Sure. The idea is not to add dread into the measure but let's adopt some measure that --

DR. WEINSTEIN: It could go either way.

DR. HOFFMAN: Thank you very much. Let's break for lunch.

[Thereupon, at 12:35 p.m., the meeting was recessed, to reconvene at 1:20 p.m., the same day.]

AFTERNOON SESSION

COMPARATIVE RISK ASSESSMENT

DR. HOFFMAN: I still have not heard from Paul Fischbeck. I got an e-mail last evening from him saying he could not join us for dinner because he had a family emergency, and I have gotten no other communication, so I am assuming that it is a real, serious family emergency.

But we have with us Don Barnes and Stu Sessions, both of whom have had extensive experience with comparative risk assessment. We talked about it and decided it probably was worth going ahead and having a short session on comparative risk analysis and the role that it might play in helping to inform priority-setting for food safety across this wide range of different types of health outcomes and different types of hazards that we have to deal with in a food safety regulatory setting.

Don Barnes is staff director of the EPA Science Advisory Board and has been involved extensively from early efforts at EPA to use comparative risk ranking as part of its effort to assess priorities across different types of environmental programs and environmental end points, and that has been an ongoing activity since the mid-1980s. Don has been actively involved in that through the Science Advisory Board.

DR. BARNES: Thank you so much. No Powerpoint slides. No overheads, even. Can you imagine that?

This is going to be strictly verbal.

First of all, I want to congratulate the organizers of this conference on bringing together a diverse mix of people. I was here yesterday for portions of it and I was very impressed to hear the interaction between the economists, the scientists, and the risk assessors.

I might note that this integration, I might call it, between disciplines is now finally coming to fruition. Obviously, there have been people who have been talking about this for years but, even as we speak, over on Connecticut Avenue the Science Advisory Board and EPA are sponsoring a workshop that is bringing together economists, social psychologists, cognitive psychologists, and people of the social sciences realm to talk about how it is that environmental decisions are actually being made in the real world.

The case study we are looking at there is nitrogen deposition in Tampa Bay. For a number of years the estuary program down there has been working to clean up Tampa Bay and the idea is to try to figure out how is it that they made those decisions and how is it that the social sciences can play a more direct role in rationalizing how it is we make these decisions.

These decision, as you all know, have been made all along. Some people say those decisions have been made politically, on a political basis, and they say it somewhat pejoratively or derisively. But politics, in many ways, is the art of the possible and there are many things that bear relationship there that bear relationship with economics, where you are trying to distribute scarce resources to do various kinds of jobs. There is a lot of commonality here and I think this is now coming to the fore and people are beginning to realize that. What I heard yesterday here, I think, was just another example of that.

When people are making these decisions about what to do, they are really making comparisons and they are

making choices. The new administrator, Governor Whitman, has come to town and very quickly she found that she had to make certain choices about global climate. She found out she had to make certain choices about arsenic. She found out that has to make certain choices about dioxin. You go down that list and none of these choices is easy. She is finding that -- I think she is finding -- that she can be well served by some more concrete analysis of talking about what these risks are and what their comparisons might be.

One of the points that I want to have us focus on here is the paper that Paul Fischbeck has written, and is available electronically. I would commend it to your reading. What the people at his university have done is to have done, I think, a very good job of taking a very academic approach to looking at the whole concept of how people make choices when they are faced with different alternatives.

I think it is the most exhaustive analysis that I have seen of looking at how people make choices and comparing that to what their individual feelings are. They have looked to see what their feelings are about certain issues and then they look to see how that reflects itself in the choices that they actually make, their attitudes and then their behaviors. They cross-check these things back and forth. I think this is very nice.

There is one technical question that I did want to ask him, however. They went through and they surveyed a lot of these people and they indicated that they did not give them any compensation. However, it does say in his paper that at the end of the interview they offered them a candy bar. The question I have is what was that candy bar? What I have seen on television about the Klondike Bar, they have this question, "What would you do for Klondike Bar?" I just think that is something we need to track down.

(Laughter)

Another question I think we need to ask ourselves as all of us here are developing analyses and providing information and recommendations and advice to someone who is going to have to do something with our decision -- again, referring back to the Richard Williams slide, with all of the different factors to go into making that decision -- the analyses we are talking about are only one part of that.

I think the question we need to be thinking about for our leaders, the decision makers, is do we want those decision makers to be implementers, taking all these, the legal analysis, the economic analysis, the science analysis, and implement the implications of those, or do we want our leaders to be leaders?

For example, a number of the things we are talking about now, and in Paul's paper there is this implicit idea that we are asking the public, the people, what their inputs are and how they feel about issues, the question comes down to do we want the leaders to implement what they think they want or do we want the leaders to do what the leaders think is right? It gets back to John Kennedy's Profiles in Courage, and so on.

It occurred to me -- those of you who follow NBA basketball probably have not heard that here in Washington we actually have a National Basketball Association team --

(Laughter)

They used to be called the Bullets -- you may have heard of them. The Wizards you probably have not heard of at all. But we did manage, for reasons that are quite unclear, to convince Michael Jordan that he ought to come and be president of the basketball operation. He now is having to figure out whether or not he is going to come out of retirement and play sort of the seniors tour and whether he can take on Koby Bryant at this age or not.

Also, another decision he has to make is, the Wizards, because of their abysmal record, have managed to get the number one draft choice for the upcoming draft. Who does he choose? If you look at the prevailing opinion of the people in the Washington area right now, people say Mr. Battier from North Carolina would be a wonderful choice to make, so there is this drum beat of that is who he ought to select.

Well, let's suppose that he does make that selection and suppose Mr. Battier does not live up to those expectations. Does that mean that Michael Jordan was simply being an implementer of what the public felt the right choice was or should he have been a leader and made some other choice? But he, as the leader, is going to have bear the consequence of his choice.

The public, we know, is very fickle about things. There is an op-ed piece in The Post this morning by George Will that suggests that people's opinions that they form in their minds are, in large measure, a reflection of the information they receive. Of course, he has never been one to really be all that high on the media, and he makes this as another case to show why people are being already conditioned.

He says if you ask the average person about some issue, he will simply reflect what he has read in the media. Again, do we want our leaders to follow that or do we want leaders to form their opinions?

I will give you an example of somebody I thought was a really good leader at EPA, and that was Bill Reilly. Back in 1988, when he came to town, thanks to some people here who were and others who now are at RFF, decided that EPA's agenda was out of control, that is, the agency just did not have control of its own agenda, that it was the Congress telling them what to do, it was whatever showed up in the newspaper

in the morning, and so on, so they said, let us try to see if we can do this more rationally.

They teamed up with the suggestion of getting 90 career people within EPA to pull aside for about 180 days and see if they could decide amongst themselves -- people who have been doing this professionally -- what the big risks were compared to what some of the smaller risks were. That is not to say that risks are irrelevant, that any of the risks are irrelevant, but can you distinguish bigger risks from smaller risks?

In the late 1980s EPA's professional group came up with a paper that is called "Unfinished Business."

When Reilly came to town, he saw that sitting on the middle of his desk and he was simply asking the question, what does this mean and is this a tool that I can use?

At that time he came to the Science Advisory Board and asked the Board to basically take a look at what was in "Unfinished Business" to see whether or not it made sense and whether or not the Science Advisory Board could then take it one step further, add value, and come up with what it thought the largest risks were compared to some of the smaller risks.

The Science Advisory Board worked for about two years and put out a report called "Reducing Risk in 1990." That report did, in fact, identify, in the ecological area, issues, saying, look, we do not have enough data, there are never going to be enough data, but we are willing to bet that if data were available, we would see these as the larger kinds of risks.

The attributes would be that they were global in scale or widespread in scale and that they were, in fact, affecting a lot of people, giving rise to irreversible problems or problems that would take a long time to reverse. Those principles dealt with ecological risk.

In addition, they looked at the issue of health risks and they said, well, even there we would like to have some information, but we think the following attributes are useful guides: that there are direct effects, direct exposures to people to toxic materials (and they went through a list of those).

Based on that, Bill Reilly, again, being the great leader that he was, said, "This is exactly what I need, because it allows me to then look at my budget, and if I can take 5 percent of my budget over the next four years while I am administrator and refocus from what is acknowledged as lower risks, refocus some of that funding to higher risks, at the end of five years I will have made a significant change in the direction of this oil tanker that is called EPA."

That was his intent. He worked out a situation with Senator Moynihan to have hearings on this -- Moynihan asked him what he could do to help -- and so on. I thought he orchestrated that very well. He knew exactly what he wanted to do with the report when he asked for it and he went ahead and did that.

In 1995, the Science Advisory Board was asked to take another look at that. "Look, it has been five years since you guys ranked the risks, maybe the risks have changed, maybe you ought to try to update it." At that time, people in the Senate told us what they had in mind. They wanted to see an array of risks so that they could say, okay, we will fund down to this level and anything below this level we will not fund, but we will cut in some way. That was the worst of all possible worlds. That was not the intent of the exercise and it was not the intent of that next exercise.

What happened in 1995 was the Science Advisory Board was asked to take a look at comparative risk, again, but look at it on a broader scale, that is, see whether or not in the decision-making process, of which risk-ranking is only one part, things could be improved in that aspect. The answer to that was, in fact, yes. The Science Advisory Board came back and said, yes, you should be trying to identify larger risks versus smaller risks but, at the same time, you ought to be thinking about the broader way of making decisions. The kinds of things that they talked about are the kinds of things we are seeing happening even today, even in this conference, and that is, bringing together a wider array of people, particularly when you start an exercise, during the problem-formulation stage, the risk assessors, risk managers, representatives of the public, and so on, so that everyone understands what this exercise is going to be about.

One of the things that the Board was asked to do and decided that it could not do was can you, in fact, merge the rankings of health risks and ecological risks and quality-of-life risks? The Board felt at that time that it was probably not the sort of exercise that the Science Advisory Board ought to do; that might be able to be done in larger context, but it would not be able to do it itself.

As a result of that exercise, in 2000 we finally finished up and put out a report called "Toward Environmental Decision Making." It is on our Web site, and it lays out the Science Advisory Board's suggestion as to how environmental decision making should be done in the 21st century. It bears a lot of relationship to other reports that have come out along the same lines.

The Risk Assessment Risk Management Commission of 1998, I believe it was, chaired by Gil Omenn and others, talked about -- they had a little benzene ring, and so on -- at each step on the line they said you need to check in with the interested and affected parties.

If you take a look at Understanding Risk, published by the National Academy of Sciences, the so-called "Orange Book" -- some people say it is a pale comparison of the "Red Book" -- it talks about an iterative process of risk assessment and risk management that goes between an analytic phase and a deliberation phase.

In the Science Advisory Board's report you can hear echoes of those earlier reports as well, but it also goes on to talk about the discussion that was made earlier this morning, that is, evaluating the impact of what you have, in fact, done. Just to do the exercise is not adequate; one needs to go back and look at outcomes. You hear echoes of GPRA in that, the Government Performance and Results Act, that the Congress has suggested that we do.

This is a quick overview of where I think comparative risk has gone in terms of the Science Advisory Board. I will just say that EPA funded over four dozen comparative risk projects throughout the country. Those have all been documented and a lot was learned in that process. In those communities where that was done it has made a tremendous impact.

Two final items: One is that I think that the problems are going to become more difficult. Someone had a slide up here talking about the progress that analytic chemists were making. I understand that in one of the EPA laboratories someone has developed a machine that will be able to detect the shadow of where a dioxin molecule used to be. Once you have that, we figure we are on the brink of another breakthrough. With that somewhat facetious remark, let me just say that I think we are all in this together and I am very optimistic about what is happening now, partly because of this conference and conferences like these, and I just hope that all of our risks will be less than 10⁻⁶.

(Laughter)

DR. HOFFMAN: Thank you very much for stepping in at the last minute. Don has to leave for another meeting at 2:00 -- we were planning to do this, this morning, as you know, so we are going to have to say goodbye. Thank you.

Stu Sessions, however, is willing to pick up where Don left off, or somewhere else. Stu is a principal with Environomics. Stu was instrumental in some of those early studies that Don was talking about, doing comparative risk studies with EPA at a community level, and did some of the early comparative risk studies in the United States. He has since, while continuing to do that work in the United States, has also done it in several settings abroad, and brings quite a rich experience in comparative risk assessment to our discussion.

DR. SESSIONS: I am glad for Don's introduction. It is not clear to me whether most of the people in the audience knew what comparative risk analysis was or much about its history.

[Transparency]

Don went over essentially its genesis in EPA with the administrator's sense that what he was hearing from the career staff about the relative importance of the different environmental problems facing the country was rather different from, perhaps, the general public's perception and rather different from the directions from Congress in some cases.

As Don said, this technique has been applied extensively to survey environmental problems and to help in setting priorities with regard to them. It has been done at the national level several times in the United States by all EPA regions, by more than half the states, and by at least 20 communities. I have participated in many of those projects.

I then branched out of the comparative risk main stream and over the last, say, eight or 10 years, have essentially been working in a series of quite large projects in different cultures. My proposition is, perhaps -- I am not sure whether this is right or not -- that the experience in doing environmental comparative-risk assessment in diverse settings has some lessons to offer for the possibility of doing food safety comparative risk in this setting.

[Transparency]

A hallmark of comparative risk: Perhaps in contrast to some of the other techniques, QALYs and willingness to pay, comparative risk sets out to be as comprehensive as possible. The charge that it is given is to survey the whole range, in EPA's case, of environmental problems that the agency could potentially have something to do with, as opposed to many applications of benefit-cost analysis, which are to gain insight into particular issues or particular interventions, particular policies.

This is, by definition, holistic, and because it covers so much ground, because it, in the environmental sense, covers air pollution and pesticide residues and wetland loss, global warming and climate change, it has to deal with wildly incompatible sort of volumes of information and qualities of information, and often incompatible types of risks and objectives.

Don mentioned that EPA has had substantial problems in figuring out how to trade off against health risks,

ecological risks, and quality-of-life risks. That, presumably, would not be an issue in extending this methodology to food safety, which is fundamentally concerned with health, I would say.

At a lower level, there still are extremely important tradeoff issues. Within the health risk area, are you concerned with health impacts, severe impacts, to small populations that are disproportionately affected by some environmental problem, like living near hazardous waste sites, or are you concerned with total population risks that may not rise to the level for any individual as one of these smaller aggregate risk problems. Even if you are concerned with a single sort of risk, you have severe tradeoff issues having to do with population versus individual risk, for example.

In any case, it is holistic, and I should say the first objective is -- I should not say identifies "objectively" the most serious environmental problems, I should say identifies systematically ("objectively" implies less role for judgment than there is). Basically, a first objective is to rank the problems, the risks, rank the problems based on the risks they pose.

[Transparency]

A second objective is, typically, to understand the anatomy of how these risks come to be, in which particular settings does the risk associated with toxic air pollutants happen to be, in which settings are those risks higher, to which individuals, which pollutants, which sources, which mechanisms.

By understanding both the relative importance of the different risks and more about how they come to pass, people believe that these are key inputs for designing remedial measures. It is all done in the context of making choices, of priority setting: Which sorts of things should we do first, which program budget should we try to increase, which should we not?

In each of the international projects I have been involved in, they have been prompted by some periodic desire or some major desire to set priorities. In Thailand, a new prime minister was going to -- Thailand does five-year plans for many government policies -- develop a five-year plan.

In Cairo, Egypt, USAID, for many years, had -- USAID has funded, in fact, about half of the environmental protection efforts that go on in Egypt -- spent all this money, largely on water and sewer, and reached the point of deciding what next for U.S. environmental assistance to Egypt. It is aimed at helping make priority-setting choices.

[Transparency]

There will typically be two stages in comparative risk analysis. The first Don has talked about, which is the risk assessment, and I should say, also, risk evaluation, what are the most serious problems, and the product is a relative ranking of the different problems. It combines whatever scientific information is available with extensive judgment about how to trade off different attributes of these risks.

The second stage is risk management. Now that we have a relative ranking of the problems, what are we going to do about them? The objectives may range from simply generating some initiatives or action plans to deal with the highest risk problems to a broader objective, which is reorienting budgets.

[Transparency]

I will speak a little bit about the success that these projects have typically had on these several objectives. With regard to risk ranking, I think most participants and most observers feel that the projects have generally been successful. The participants in the projects have come to -- initially, it was a group of career staff in the first one at EPA, but as these projects have evolved, it has become agreed that it is important that the participants represent a very wide cross-section of technical experts, policy-level officials, community representatives, industry, environmental activists, stakeholders across the whole range of interest.

Rather surprisingly, these groups, this broad group of participants, have reached some consensus on what the results are. They think the risk ranking has captured something they agree on. They also typically believe or typically find that the ultimate risk ranking is somewhat different from the naïve risk ranking that has often been conducted in advance of the process. So they think they have captured something and they think they have learned something during the process.

There are typically also some important procedural benefits for the people who participate in this, people from different agencies, different programs, who do not usually talk to each other do as a result of this process. People who worry about health risks talk with people who worry about ecological risks. There is a lot of cross-fertilization. Typically, the consensus that is developed is then set to be carried forward by people in these different participating organizations, sort of implementing the results as they can in their own individual organizations.

I have said the results reflect judgment equally as much, if not more so, than science. It is highly important to make the judgments explicit but, even so, the product of the report, for example, from one of these

comparative risk projects, is often not fully persuasive itself. Part of the validation of the process is the people who participated in it, representing many groups, speak highly of it, so it is as much a procedural validation as a substantive validation.

There has been a second phase in most comparative risk projects. The second phase has typically been much less successful than the first. Oftentimes, the result is that you have a risk ranking and people then go back and say, well, let's do some more projects in the high-risk areas, let's up the budget for the high-risk areas.

In fact, the tie from this risk assessment into the ongoing planning, budgeting, and accountability procedures has often been very weak and these projects have typically been done on a sort of one-shot basis and establish an agenda or a blueprint that circulates for a couple of years that people refer to and then it sort of disappears. It is not updated on an ongoing basis, it is not integrated into the ongoing decision-making processes of the agencies that have supported these efforts.

There is also, I think, relatively little that the comparative risk people have been able to come up with in the way of analytical procedures to answer the question, not what risks will be reduced if we pursue X initiative but, the broader question, what will happen if we up the budget for this program area by 5 percent and reduce the budget for that program area by 5 percent?

What you do in a risk-management sense should have to do with not only the magnitude of the risk you are addressing but the feasibility and the cost effectiveness of reducing that risk. It is not at all clear that the feasibility and cost effectiveness of reducing risk in a high-risk area is necessarily better than it is in a low-risk area.

Efforts to investigate this in a broad way, looking at opportunities for risk reduction in a whole programmatic area, have not been particularly successful. It has been an analytical hope that people have had in doing these projects that has not been fulfilled -- they have not succeeded in doing that.

[Transparency]

I have a couple of speculations about what the experience in using this technique in the environmental arena might suggest for potentially using it for food safety. The initial one is that the risk-ranking process has been quite well embraced by stakeholders on opposite sides of the issues. It has worked to have industry and environmental activists together engaging on what are the largest risks that we face.

They are much more willing to engage and discuss on risk ranking, there is much more common ground, there is much more ability to reach consensus on the risk-ranking stage than there is in the risk-management stage. Different objectives and different desires often prevent consensus at the risk-management stage, but essentially splitting the process into these two separate steps, I think, has been interesting and useful. It is a sort of integrating procedure, at least with regard to risk assessment, across people with widely divergent interests.

A second issue -- I mention this quickly -- is that, ultimately, EPA has a lot of trouble in making final judgments and in making priority decisions because of tradeoffs and different impacts on health versus ecological versus other sorts of risks. Presumably that would not be such an issue in food safety.

A third concern is, as I have said, in trying to deal with risks across such a wide spectrum of program areas, you get widely different sorts of information bearing on risk in the different program areas and widely different qualities of information. A particularly vexing problem is when you get risk information that has embedded in it different degrees of conservatism.

My general sense is that oftentimes epidemiologically based evidence is essentially best estimate; it does not necessarily have conservatism or lack of conservatism, it is a best estimate, it is an expected value.

Oftentimes, incidence information is that way, also, you have to worry about undercounting and misdiagnosis, and things like that.

Typically, the toxicological risk information that is used has buried in it various degrees of conservatism. In many cases, for many chemicals, the best that happens in comparative risk analysis is assessing the number of people who are exposed to a particular toxin at levels above the reference dose or reference concentration, and you have no idea what that means with regard to the likelihood of suffering the actual health effect.

I think this sort of problem would probably, I would guess, be more endemic in food safety than in the environmental, in that so much of the environmental risk assessment is chemical-based and toxicological-based, and there are relatively few incidence data and not all that much epidemiologically based data.

I think this would be, perhaps, more of a problem in the food safety area. There have been a lot of interesting discussions about how to compensate and how to pull from risk estimates the conservatism that is embedded in them. It is doable to some degree, with effort, on exposure and that sort of thing, but pulling

that from dose-response information is extremely difficult.

A final observation is that, to some degree, many environmental risks are relatively constant across different people. Ambient air pollution: People breathe roughly the same amount of air per day -- there is some variation in the concentrations that you are exposed to. With regard to food, I expect that how you process and handle and cook the food and differences in intake of different foods across different people are substantially greater than the sorts of variations that, in general, come into play on the environmental side.

In comparing the risks associated with very large program areas, you have to take shortcuts and represent the risks associated with, say, hazardous waste sites by a series of scenarios, for example. The greater the variation across individuals in behaviors and exposures, the more scenarios you need to begin to capture some of the variation that really exists out there, and the more difficult the process becomes.

In summary, it has been a procedure that has produced some substantial benefits in the environmental area, far from all that people had hoped for it, but I think it could conceivably have some promise as well as some difficulties in food safety.

DR. HOFFMAN: We have some time for questions.

GENERAL DISCUSSION

DR. RODRICKS: Stu, I want to ask you a question about the ranking process itself and the role of judgment versus what you said was science. Is it possible to dissect in the ranking process, pull out, let's say, the influence of the sort of narrow technical view of risk assessment, which is trying to give a picture of probabilities and health consequences, as against other attributes of a risk that people may take into account in judging the impact of a risk? Is it possible to talk about, in the ranking exercise, how much is of the first as against the sort of judgment character, which I assume must play a fairly large role in the process?

DR. SESSIONS: Different projects have taken different views on the question you are posing. Some risk rankings have attempted to rank, including both the technical aspects of risk, or the scientific aspects, as well as the risk perception, or the "dread" qualities, whether it is voluntary or not. Other projects have attempted to give a more limited assessment of solely the technical risks. I think probably more have been the latter.

Interestingly, I think in the foreign settings people have been more willing to focus on only the tangible aspects of the risk, the "body count," or the numbers, as opposed to the qualitative attributes of it. I think there is more of a culture in willingness to rely on experts and looking for the technical judgment of the experts to guide priorities in the less-developed countries I have worked in relative to the U.S.

The way it has often been handled is there have been two separate committees participating in the risk ranking. There has been a technical committee that has been solely charged with trying to estimate quantitative risks as best they can. Then there has been a policy-level committee that putting a sheen on it having to do with equity and public perception and voluntariness, and that sort of thing.

In a couple of projects you have gotten even two sets of rankings, which is quite interesting, how people modify the technical ranking to reflect qualitative public concerns about different risks.

DR. RODRICKS: Thank you.

DR. HOFFMAN: If there are not any other questions, we are actually just about back on schedule, believe it or not. I think what I would like to do, since you had such a short lunch, is to take a break now, have the willingness-to-pay sessions, and then a general discussion following immediately on that.

[A brief recess was taken.]

USING WILLINGNESS TO PAY TO VALUE FOOD SAFETY

DR. HOFFMAN: Welcome back to starting this next session -- for me, I guess, I am saving the best for last, because this is the area of economics that interests me the most.

Our final session will be looking at willingness-to-pay studies as a way of attaching value, trying to figure out the tradeoffs that people make between reductions in risks to health or life and other goals that they have in life. Our speakers for this session will be Jim Hammitt, who is Associate Professor of Economics and Decision Science in the Harvard Public Health School. He is director of the program in environmental science and risk management.

Then we have, as discussants, Maureen Cropper, who is Professor of Economics at the University of Maryland, College Park, and who is also an economist with the World Bank and University Fellow here at RFF, and Kerry Smith, who is Professor of Agricultural Economics at North Carolina State University.

All three of our speakers have had extensive experience with both health and environmental valuation methods and I am looking forward to their comments on the kinds of adaptations and the roles that these

methods can play in informing priority setting across different types of health outcomes for food safety and different types of food safety programs.

DR. HAMMITT: I am going to give an overview of the concepts and methods for estimating willingness to pay to reduce health risks.

[Transparency]

In particular, what I want to talk about is, first, the theory related to three kinds of health end points: fatalities, chronic morbidity, and acute morbidity. Then I will talk a little bit about the empirical methods of estimating values, which are revealed preference, continued valuation, and then experiments, which I see as somewhat different.

Then I would like to conclude by comparing these with the quality-adjusted life year approach; in particular, get at that question of is there some dollar value for a QALY, to which I think the answer, from the individual utility perspective, is really "no."

[Transparency]

From a willingness-to-pay perspective, what we are trying to do is estimate the social value of reducing risks of some sort of health effect, and that involves three components. There is the private individual willingness to pay to reduce the risk, which would incorporate a version of the pain and suffering and utility loss from being sick, financial costs that are incurred by the individual who is sick, lost income or medical expenses. In addition, a lot of those financial costs are externalized, they are paid by health insurance, disability insurance, or something of that sort, but that is also a cost to the society. In addition, in some cases, there may be some altruistic value that should be added, people's willingness to pay to reduce the risk of other people suffering adverse effects.

A big difference between willingness to pay and the QALY perspective, as was discussed earlier, QALYs are very much focused on consequences, what the health effects are and how long they are suffered, whereas willingness to pay can, in principle (and probably in practice), incorporate elements of qualitative attributes having to do with voluntariness, controllability, dread, and many others.

[Transparency]

I would also like to make the point here that when we are valuing reductions in food-borne risk, a program to make many foods safer is probably more valuable than the sum of a lot of individual programs to make single foods safer, because if only one food is dangerous, it is relatively easy at low cost to avoid that by eating something else instead. If a wide range of foods is hazardous, it is much more difficult to substitute away from them, so it is this nonlinearity or nonadditivity in the value of a program that suggests we should be careful about estimating values of risk reduction on a food-specific or very narrow basis if we are really interested in evaluating broader programs.

[Transparency]

I will turn now to the theory of evaluating mortality risk, on which I will spend most of my effort. As I am sure everyone is aware, the main concept here that is used is the value for statistical life. The name is a sexy, catchy name, but it is a very misleading name. A fellow named Ron Howard, at Stanford, proposed an alternative name, value of a "micromort," that being one-in-a-million mortality risk, so if the VSL is \$5 million, the value of a micromort would be \$5.00.

That is what we are talking about, the willingness of an individual to trade his own money for small changes in his mortality risk -- Δw divided by Δp . You can think of that as a value for statistical life in a couple of senses. One is, if Δp is $1/n$, n being some large number, and you are envisioning a program to reduce the risk to n people by $1/n$ each, each of those people would be willing to pay Δw . The total amount of money collected, or total willingness to pay, essentially would be n times Δw . One fatality would be averted within a time period in expectation, so that gets you $n \Delta w$, which is, of course, algebraically equivalent to $\Delta w / \Delta p$. Typical estimates for the U.S. these days are on the order of \$5 million.

[Transparency]

This is an indifference curve relating probability of surviving for a specified time period, like the next year -- and, obviously the duration of this time period matters, although it often does not get much attention in the literature -- against wealth. Obviously, high-survival probability and high wealth is a good place to be. The opposite corner is a bad place to be, so you can draw indifference curves indicating relative preferences for survival probability and wealth.

VSL is willingness to give up a small amount of money to get a small increase in survival probability.

Under a model I will show you in a minute, we anticipate this indifference curve is downward sloping and is curved in the way I have shown. VSL is the slope of this indifference curve. The slope changes as you

move along the curve. That is one reason the VSL is not a constant for an individual, and I will show you many further reasons.

[Transparency]

This is the standard model of value for statistical life. It is the marginal substitution between wealth and mortality risk, dw divided by dp . The standard model here assumes a couple of things. We have w as wealth, or income, within this time period, the next year. $U_a(w)$ is utility, which depends on wealth conditional on surviving this year. U_d is utility, which depends on wealth conditional on not surviving this year, so U_d incorporates both the value of living for the part of the year you survive, if any, plus the value of additional wealth as a bequest to pass on to dependents or others.

Often it is omitted, but there is potentially a monetary consequence associated with death, which I have represented here by L -- it could be medical bills or it could be life insurance benefits, in which case it would be a negative loss. P is the probability of dying in the period.

The standard assumptions here, which I think are not very controversial, are U_a is bigger than U_d -- you would rather be alive than dead. The marginal utility of wealth conditional on survival exceeds the marginal utility of wealth conditional on dying in the period. That is a little bit less obvious, but I think it is pretty reasonable.

Then often we assume as well that the secondary derivatives of these functions are negative, which means risk aversion or diminishing marginal utility of additional wealth.

[Transparency]

If you make those assumptions, VSL -- what I did not say is the numerator here you can think of as a change in utility, so utility if you survive, less the utility if you die, so there is a difference in utility. The denominator is an expected marginal utility, p times the marginal utility if you die, plus $1 - p$ times the marginal utility if you survive.

Now we are going to look at how this ratio depends on the function of factors. One is p , the base-line risk. My buddy, Richard Zeckhauser, came up with the great name, "the dead anyway effect." If your chance of dying in the period is high, your VSL will be relatively high as well, so simply why not spend a lot of money for an incremental survival probability, what are you saving it for? You are likely to be dead anyway, there is no reason to save it, spend a lot for risk reduction, whereas if you are very likely to survive, there are a lot of other things you might want to do with that money, so you would be willing to spend less for the same risk reduction.

It is reasonable to think and empirically true that wealth increases VSL. The magnitude of how strong this effect is, is, I think, we do not really know very well. People often assume income elasticity ought to be about one, for lack of any better assumption. Most empirical estimates suggest that it is substantially lower than one, on the order of 0.3 to 0.5, within a study.

If you compare VSL across countries with very different income levels, there is some evidence suggesting the elasticity is bigger than one, but there are, of course, other reasons why VSL might differ across rich and poor countries. I think basically we do not know very accurately what income elasticity is.

[Transparency]

The effect of health here conditional on survival actually has an ambiguous effect on VSL. Consider two situations. One, if you survive, you will be in excellent health. Alternatively, if you survive, you will be in poor health. Clearly, the numerator term, the ΔU , is bigger if you survive in good health than in bad health.

But there are some empirical estimates and some reason to think that the marginal utility of wealth, if you are in poor health, could be lower than the marginal utility of wealth if you are in good health. There is a study by Frank Sloan and some colleagues estimating this ratio as somewhere between 0.1 and 0.7 in different circumstances, so suggesting quite a large effect.

That means if you are going to survive in poor health, the marginal value of money to you in poor health is not very high, so you might be willing to spend a lot of it to reduce your mortality risk. It is theoretically possible, at least, that VSL would be higher for someone who will survive in bad health than for someone who would survive in good health. It certainly seems counterintuitive and certainly contradicts the quality-adjusted life year perspective.

Another issue is how does competing mortality risk affect this? I might get sick and die from food poisoning, I might die from any other causes. Here you can show that VSL falls with the magnitude of competing risks, so if somebody is in a bad health state, has a short life expectancy, for example, because he faces some other disease, his VSL would tend to be smaller. What does it pay to reduce a food-borne risk, at least in theory?

[Transparency]

VSL also depends on financial risk. Wealth is, of course, uncertain, future earnings are uncertain. How VSL depends on this risk depends on characteristics of the utility function's risk aversion in the state of survival, in the state of death, and also a concept called "prudence," so a lot of things can happen, but it seems plausible for reasonable assumptions that greater financial risk will reduce VSL.

Age, or life expectancy, has an ambiguous effect on VSL so, from the QALY perspective, preventing the death of a young person is unambiguously better than preventing the death of an older person, because the younger person, presumably, has many more future QALYs at stake. That much is true from the willingness-to-pay perspective as well. If you have a large life expectancy, the gain in utility from surviving this year is larger, because you will go on to live many more years, presumably, but the opportunity cost of spending money to reduce current mortality risk may also be higher, because you will have a greater time period over which to spend that money.

You have a longer time horizon for investing that money, for example, so these two effects make the ratio ambiguous. There have been some life-cycle models to try to estimate the net effect. The general result is VSL rises with age and then falls, but at exactly what age it peaks, how much it goes up and down before that, are very sensitive to different modeling assumptions, and we do not have a lot of good empirical evidence on that, either.

The bottom line there is the value of reducing mortality risk can be sensitive to a large number of factors. Now I will talk about some of the other health effects.

[Transparency]

Chronic morbidity. The simplest thing to do is to reinterpret the model of willingness to pay for mortality risk reduction as a model of willingness to pay for reducing the risk of chronic morbidity, so here I have called it VSCM. Now I am going to let u_a represent utility for wealth conditional on being in good health, and u_d utility for wealth conditional on having this chronic disease. Let's see whether the standard assumptions still make sense in this context or not. Clearly, u_a healthy is better than disease -- there is no problem with that one. The second one, though, is potentially problematic, because, remember, we have this monetary loss, still, which could have to do with lost future earnings because you are ill, or high medical expenses because you are ill.

I think it is reasonable to think, at the same wealth level, the marginal utility of income would be higher in good health than in bad health, but if having chronic morbidity reduces your income as well, that would tend to push up the marginal utility of income in that state, so it is possible that this assumption would not be appropriate in the case of chronic morbidity. The risk aversion, I think, carries through with no problem. If these assumptions carry through, the qualitative effects I just talked about for mortality carry over directly to chronic morbidity. If the second assumption is not true, there are some changes that I think are not hugely important, but some of the things we talked about earlier would differ.

Another point that I think is important here is, remember, I started out talking about private willingness to pay, and there is the pain and suffering and utility loss of ill health or death, but there are also the monetary consequences. For fatality we tend to think that the utility is the big item; the fact that you lose future income is much less important, comparatively.

For morbidity, that may not be true, because there can be high continuing medical expenses and, because you continue to live, the health is not as bad as dying, so the monetary costs may be relatively more important.

[Transparency]

For acute morbidity, to my knowledge, there is not much theory developed on this point. We would anticipate, in accordance with common sense, that willingness to pay to reduce the risk of an acute morbidity incident would increase with the duration of the incident, how long you would be sick, and how severe the symptoms would be.

There are a few empirical estimates in this field that suggest that willingness to pay does increase with the duration of the illness to be prevented but not in proportion to duration; it increases at a rate much lower than the duration.

Willingness to pay seems to increase with the severity in terms of defined symptoms, and so forth, but how would you measure severity? One thing that could be done is to use the Q, the health-related quality-of-life weight that is used in quality-adjusted life year, take that as our measure of severity. There are good reasons to do that, as Milt was suggesting earlier. There is one study that compared willingness to pay to the Q here and found that willingness to pay increased for more severe illness, as measured by Q, but at a decreasing rate.

[Transparency]

Let me say a bit now about empirical methods. The two broad classes are revealed preference, where we observe what people do, assume people are making decisions that are in their own best interest, so we can figure out what their interest must be for them to have chosen the action they chose.

This requires, of course, that people are making choices where they are trading off health risks for other things. We observe those choices. Typically, we observe what people chose but we do not observe what they passed up, so we have to make assumptions about what the alternatives were, and that is potentially a problem if we make poor assumptions there. I talked a bit about cross-sectional versus time series, revealed preference estimates in food.

The other approach is continued valuation or other stated preference methods, which is incredibly flexible. You can ask people questions about anything and they will usually answer but we do not really know what to make of their responses.

[Transparency]

Revealed preference, cross-section. One set of studies that comes to mind is looking at willingness to pay for organically grown food rather than conventionally grown produce, for example. Here we need to find places where people are choosing safer and less safe foods and may be spending more for the safer foods. For example, a study I did a long time ago suggested, in Los Angeles, when organic produce was not so common, that it cost on the order of 50 percent more than conventionally grown produce. People who bought that were revealing a willingness to pay of that magnitude.

Now, the question is, what is it they were buying, what were they paying for? What did they think they were paying for? More recently, there is a recent study by Pamela Williams, who is a doctoral student at the Harvard School of Public Health, working with me. We did a survey of on the order of a thousand people in the Boston area and asked them about perceived risk and a lot of other attributes.

What they estimated is that organically grown produce was safer because there was less pesticide on it. These are mean estimates of the annual risk of dying from pesticides on produce among people who buy organically grown produce. Their mean estimate here is almost \$600 million for buying conventionally grown produce, so that is six in 10,000 -- that is probably close to people's total annual mortality risk.

Buying organic would reduce it to 100 in a million, so a pretty big reduction, although not to zero.

We also asked them about risk reductions from natural pesticides, naturally occurring toxic constituents in the plants and got microbial contaminants on the produce, and there they thought the conventional was more dangerous than the organic, but there was a much smaller difference between the two types of produce, which seems logical.

They also thought organically grown produce was much safer to the farm workers than conventionally grown produce -- there was a pretty big difference there. They talked about organically grown produce was better for the environment or ecosystems. Also, things having to do with a sort of political statement, supporting small family farms versus corporate agribusiness.

The extent to which these things are accurate is another question, but this is what many people believe, at least.

I suggest here that other kinds of cross-sectional studies that might be done would involve food produced with genetically modified organisms, or genetic engineering, more generally, irradiated versus -- you have to find something where the food is pretty similar but you can imagine it differing in risk. But then, as is obvious here, there will be other attributes of these technologies that influence people's behavior that we need to sort out. There have been a number of what I would consider to be time-series statements.

[Transparency]

Here is the alar-on-apples issue studied by Irene _____ and John Hain. There is a whole series of these, where there is some new information given out to the public about risks associated with some food and, typically, the demand for that food falls for some period of time and, from that, you can estimate willingness to pay.

It seems to me these kinds of studies have some big advantages over the cross-sectional ones in that apples are pretty similar. Once you find out they have alar on them, that is one important difference, potentially, but everything else about the apples is the same; it is not like organic versus conventional apples, which differ in many perceived attributes.

One limitation here is these are typically short-term studies. They tell us something about short-term willingness to pay, which possibly is an overreaction to new information -- maybe not. Also, the point about substitution, it is easy to stop eating apples for six months. It would be much harder to stop eating apples for the rest of your life, so the values for the short and the long run would be different.

[Transparency]

Let me talk a little bit about contingent valuation now, the other broad class of methods, where, typically, we conduct surveys. We explain to people some sort of intervention or different products that will reduce their risk of some specified health effect and we figure out -- maybe we state a price for it, if it is a private good, as would be reasonable in the food context -- whether people say they would buy it or not.

[Transparency]

As is well known, there has been a lot of criticism of C.V. Two of the issues are that the choices are hypothetical, which has a variety of implications. One is that respondents might strategically misrepresent their preferences in order to alter government policy in the desirable direction (I do not think that has proved to be an important problem in practice).

Perhaps more serious is that there is limited incentive for the respondent to really figure out what he would do. These are hard questions. How hard can you get people to work and how willing will they be to work to really figure out what their decision would be? For unfamiliar choices perhaps they do not even know. It may be unknowable to them what their behavior would be.

Another issue is that some results of C.V. studies seem to be inconsistent with pretty credible theoretical implications about how willingness to pay should depend on some factors and not on others, so I want to say something here about inadequate variation of willingness to pay with magnitude of risk reduction.

[Transparency]

For small changes in risk reduction I think it is reasonable to assert that willingness to pay ought to be almost exactly proportional to the magnitude of risk reduction. Remember, this is like the curve I had before. VSL is the slope of this curve. As long as the slope of this curve does not change radically over the interval where we are reducing risk, the slope will be roughly constant and, therefore, these two Δw increments will be roughly equal, so willingness to pay to increase survival from here to here should be twice the willingness to increase survival from here to here.

Let me point out that this is for different reductions in the probability of some specified health effect and to the extent that willingness to pay may depend on whether it is cancer versus auto accident or all these other attributes, I am not -- well, that is fine, that may well still be true. I am only talking about changes in the probability of dying from one specific cause, so holding all that other stuff constant.

[Transparency]

There is a substantial literature on C.V. studies on numerically specified health risk reductions, where, when John Graham and I reviewed, we could find 25 studies published in the 1980s and 1990s. Only 13 provide enough information to determine whether willingness to pay varied with the probability reduction and there we found that in most of the studies, indeed, willingness to pay was larger for a larger mortality or other health risk reduction, but in no study was willingness to pay ever proportional to the risk reduction. We characterized this as inadequate sensitivity to risk reduction.

That implies that the VSL we get, our estimate of the slope of that indifference curve will depend on which risk reduction is asked, and the choice of which risk reduction is asked about is essentially arbitrary. I think that is a serious problem.

[Transparency]

Another problem is that it can confound our estimates of willingness to pay varies with different health effects. This is a study by Henson -- I think in Scotland, actually. It is a C.V. study, so methodologically it is not fantastic (it was a mail survey, and so forth). But let me make the point anyway.

This was reducing the risk of Salmonella, either from eggs or chicken. He asked people's willingness to pay to reduce their risks of mild, moderate, and severe illness or death, with these stated risk reductions. He got these mean values, so 4.6 million pounds would be the VSL for Salmonella contamination associated with eggs.

If you just look at how these values per case line up with severity of illness, that is plausible. So the question is, it is plausible but is it right? You can figure out what the average willingness to pay for each one of these risk reductions was by multiplying the mean value by the risk reduction and you see that these are pretty similar, so it is possible.

Michael Jonesly and some of his colleagues did a follow-up study on this, where in focus groups they often find, when you ask people, "How do you decide your willingness to pay for something," "Well, I kind of think of how much I can spend that wouldn't really disrupt my budget very much," so it is a small amount. This is annual willingness to pay, so between about three-quarters of a pound and three pounds per year, so small amounts.

If what people are doing is saying, well, reducing the risk of illness is a good thing, I should pay something

for it, how much could I -- I mean, it is a good thing but it is not the most important thing in the world, so how much could I spend that would not really disrupt my budget? I come up with a couple pounds per year and that is what I report and then the analyst goes and divides this by risk reduction and he gets these hugely varying values per case that could be simply an artifact of the fact that the stated risk reductions varied tremendously across the health effects. It could have nothing at all to do with people's preferences for these health effects.

Another interesting point here is you see the values of reducing any one of these risks if you got the illness from eating chicken rather than eggs were somewhat higher. An alternative explanation for that is when people are figuring out what it is worth for this risk reduction, it is anchored to the price of the food. Chicken costs more than eggs, so willingness to pay is maybe a little bit higher there. That seems to conflict with what we think we are trying to measure here.

[Transparency]

Let me just say a big issue here is how do you communicate these small risk changes to people? I have done some work on that with John Graham and Theodore Corso, also a former doctoral student of ours, where we did a survey with split samples. This is an auto risk reduction and we used different visual aids for explaining the risk reduction.

This is a risk ladder, a risk scale. Things like this have been used in the literature for 15 to 20 years now. In the question I will show you we are asking about a side-impact airbag that would reduce your annual risk of dying in a car crash from either two-and-a-half or two in 10,000 down to one-and-a-half in 10,000, so it is the difference between one of these blue stars and the lower blue star.

This is a logarithmic scale with risks of dying from other things to provide people a context, to help them understand what these numerical measures mean. If they have some sense of how likely it is to die in a plane accident, from a bee sting, from lung cancer, the idea is to provide some context to help people translate the numerical information.

[Transparency]

We had another risk scale and then we had -- this is only half of it -- 25,000 dots, where we told them this risk reduction was going from, I do not know, six dots to four dots on this page, or something like that.

[Transparency]

This is comparisons between people, so no individual value is more than one risk reduction and saw more than one visual aid, so this is all between-subject comparisons. Here is the median willingness to pay for the small risk reduction and the large risk reduction. What you see here is that for the people who got no visual aid, the values are \$235.00 and \$250.00, virtually the same, even though the risk reductions differ by a factor of two. If you calculated VSL by dividing willingness to pay by the risk reduction, you get either \$4.7 million or \$2.5 million, depending on which risk reduction you used -- it is a totally arbitrary choice. Alternatively, the dots on the last one I showed, willingness to pay was \$165.00 or \$323.00. Those differ by a factor of almost two, so when you divide willingness to pay by risk reduction, you get the same answer for VSL. What this is showing is, as we might have expected, willingness to pay can be very sensitive to how the risk information is presented and at least some ways of presenting the risk reduction seemed to lead to estimates of willingness to pay that -- variation in willingness to pay with risk -- are in accord with our theory, whereas other methods, like no visual aids, lead to these kinds of failures that we have observed a lot in the literature.

[Transparency]

Let me say one word about experiments. What I have in mind here is studies by Jay Shogren, who is in the audience, and some of his colleagues, where they take a small number of people into some setting, offer them different kinds of food varying in risk, and force them to use real money to buy the safer stuff. It is like revealed preference in the sense that they are incurring the risk and they are using real money to reduce the risk, so it is not hypothetical. That is very nice. The experimenter has a lot of control in terms of determining what the foods are and what the characteristics are and what the information is presented to the experimental subjects. That is very nice.

One weakness, I think, is that because it is an experiment in a very novel setting for people, willingness to pay in this setting may not correspond too well to willingness to pay in the more routine decisions, but that is something we can learn about over time. I think that has potential.

[Transparency]

To wrap up, this willingness-to-pay measure I have been talking about is derived from and consistent with standard economic welfare theory. We do have empirical methods that allow us to get estimates. There are questions about how good those estimates are, but we can at least make the estimates.

Then let me just sum up by suggesting that the relative value of different health risks as ordered by QALYs and as ordered by willingness to pay could be substantially different. We talked about this a little bit. Baseline risk increases willingness to pay to reduce mortality risk -- this is all for mortality risk here -- whereas it has no effect on QALYs gained from risk reduction.

Higher wealth increases willingness to pay. Nominally, it has no effect on QALYs, although I think there are some small qualifications to that point. Poorer health unambiguously decreases the QALYs you get from reducing mortality risk but its effect on willingness to pay for that mortality risk is ambiguous.

Finally, we have one that lines up the same way, competing risk decreases willingness to pay and also increased QALYs gained from an intervention. Financial risk affects willingness to pay but not QALYs. Life expectancy, the effect on willingness to pay is ambiguous. With QALYs, higher life expectancy clearly means higher QALYs. By technology here, I really mean the qualitative attributes of the risk and the risk control measure, and all that. Irradiation reduces microbial risk a lot but people do not like it for other reasons, and that would affect their willingness to pay but would not affect the QALYs calculated, because QALYs depend on only the health consequences.

Thank you.

DR. CROPPER: I think that Jim did a really excellent job of summarizing the willingness-to-pay approach. I am not going to comment directly on what he said, but because I think one of the topics here of this conference is the choice, I would say, between willingness to pay and QALYs, I would like to really address that topic and, I guess, I think the kinds of questions that you would want to answer in deciding to do one approach versus the other.

I will say from the start that I am actually going to be very critical of the willingness-to-pay approach as a method of making decisions in a policy context.

[Transparency]

One question that Jim touched on a little bit, and I must say I am sorry I arrived here at the conference only at noon, is whether willingness to pay and, by extension, cost-benefit analysis, is superior on theoretical and/or philosophical grounds to the use of QALYs and cost-effectiveness analysis.

I will mention in the next slide why I think economists tend to favor willingness to pay and cost-benefit approaches -- I am not going to try to delve too much into that question.

Can willingness to pay for various health end points or various risks be measured with enough precision, let's say within an order of magnitude, so that you can use it for policy making? I think this is a very important question if you are going to try to using the willingness-to-pay approach to actually rank risks in the context of food safety, or any other policy decision, and I think willingness to pay has some real problems in this regard.

A related question has to do with if you have a wide variety of health end points, how many willingness-to-pay studies would you really need, assuming you could do valid ones, ones you had confidence in, to produce the estimates necessary to really rank a lot of different regulatory options? I do not know how much the different the kinds of health end points associated with food safety programs have been discussed, but I got a little flavor of that with Tanya's question this morning.

So these are the questions I am going to focus on.

[Transparency]

Why do economists like the willingness-to-pay approach? I think one of the reasons is really the desire to have, shall we say, regular people's, citizens', values count in making regulatory decisions. Certainly, in formulating QALY weights, one can go to a wide spectrum of individuals. There is not necessarily a big difference between the two. But if you are really going to use cost-effectiveness analysis, at some point somebody has to draw the line on what is too much to spend for increasing a QALY and, in that case, the people who, in some sense, are making this decision in a benefit-cost analysis are the individuals whose willingness to pay is being summed to do the study.

I am sorry I missed Milt's talk this morning, because of the change in the program. When we talk about assumptions that are used to justify the tradeoffs that people make to get the QALY weights, we do wind up with certain differences between this and the willingness-to-pay approach, which Jim actually hinted at. For example, if you ask people what they would pay for risk reduction, reduction in their risk of dying, you will not necessarily find that this willingness to pay is strictly proportional to remaining life expectancy or, as Jim also pointed out, it might vary according to qualitative characteristics of the risk.

Finally, as I think Clark Nardinelli was saying when I first came in, if you do use a willingness-to-pay approach, you can combine the value of reductions in risk on the health side with nonhealth benefits. That is definitely a plus.

[Transparency]

What I think are the real problems here, though, and, again, these are, in some cases, practical difficulties, but in some cases more fundamental difficulties, are these. When we do willingness to pay for health effects, at least most of the time (this is the approach Jim emphasized) we are asking people to trade dollars for quantitative risk reductions.

A question, in my mind, and we will talk about this in a little detail, is whether people really have well-defined dollar values for quantitative risk reductions, which is very different from saying we know people make explicit choices with regard to risks and we know that they pay for different sized risk reductions, but the question is, are these conscious choices, do people really have well-defined values here? We are asking people to make very different kinds of choices in getting QALY weights, where we are not really trading dollars for quantitative risk changes.

I think the difficulties in estimating the value of a statistical life are really substantial and I will talk a little bit about those. I know Kerry is going to actually present some empirical results on the value of statistical life, so I hope you will not take my criticisms personally, but I do think -- I mean, I have tried to do this for a long time, too, and I still think there are really big problems that make these numbers very hard to use in policy.

You definitely need different willingness-to-pay studies for different health end points and, in terms of the number of studies you might need in the food safety area, that is perhaps something we can discuss in the general discussion at the end of the day.

The reason I bring up as an example, I think, of the problems in using the willingness-to-pay approach is the series of studies that was done by the U.S. EPA. The EPA did a study of the benefits and costs of the Clean Air Act of 1970 and 1990 and another study of the benefits and costs of the Clean Air Act amendment, the 1990 amendment. These were willingness-to-pay studies mandated by Congress, where the majority of benefits were health benefits.

[Transparency]

In terms of monetary values, the majority of the benefits were reductions in premature mortality from reducing particulate matter in the air. In both studies the U.S. EPA used a value of statistical life of about \$5 million in 1990 dollars. What happened when both studies came out was that the agency was subjected to a tremendous amount of criticism, not just by academics but by other groups for the use of these very controversial numbers, which led to huge benefits associated with the Clean Air Act, benefits that, in the case of the 1970-1990 Clean Air Act, amounted to about \$16,000 for a family of four in the year 1990. What happened was, here was a study where a lot of very good work was done to look at exposures to pollution to try to come up with a sensible counterfactual to estimate the quantitative health impacts of the Clean Air Act, and a lot of work was done on the side of the cost of the regulations themselves.

One of the things that is never mentioned, and I am sure most people do not know this, because they have very much dismissed this study, is that the cost per statistical life saved by the Clean Air Act from 1970 to 1990 was about \$125,000, and the cost per life year saved, depending on what discount rate you use, is something in the neighborhood of \$20,000. When the studies were done, and I, along with other people here, were on the Science Advisory Board of EPA, we said, look, you really ought to be telling people about these numbers. One hundred twenty-five thousand dollars per statistical life saved is nothing to be ashamed of.

What happened was, no, this value was used of \$5 million. It led to huge controversy and a lot of the good work that was done, I think was just totally lost.

Do people have monetary values for risk changes? Part of the basis for estimating, empirically, willingness to pay for risk changes -- and we are talking mostly about changes in risk of dying, but we could talk about a change in the risk of contracting chronic bronchitis -- is that people are implicitly trading dollars for risks. It was suggested to me a couple of months ago that I have a base-line colonoscopy that I might have to pay for out of my own pocket, because there was not a good medical reason for it, and it would cost me \$2500. I went ahead and did it, because I thought this was the sort of thing that an educated, conservative person who tries to take care of herself does.

Did I actually compare the cost to myself for this with the quantitative risk reduction? No. The example here of mammograms, which is perhaps the more common procedure, and a cheaper one -- I had even read the studies by David Eddy on the cost effectiveness of mammograms annually for women between the ages of 50 and 60 and yet, I must say, when I make these decisions, I do not compare the quantitative risk reductions, even as estimated by Eddy, with the cost of the mammogram.

In focus groups that I have done with people in this room to try to estimate or to try to prepare surveys to

estimate what people will pay for quantitative risk reductions, it is really here that people do a lot to reduce their risk of dying and their risk of illness. But the problem here is actually making the comparison of the quantitative risk change with dollars.

I think you have to ask yourself seriously if people really do this, which, of course, then raises the question, if they really do not have, perhaps, very well-established values, what are you finding in revealed preference studies and what are you finding in stated preference studies?

[Transparency]

The largest category of studies that is done for estimating the value of statistical life are compensating wage studies, looking at tradeoffs in the labor market between what you are compensated and your risk of dying on the job. One possibility, and I admit I am being a little bit cynical here, although this is also partly based on analyses of data of this type that I, myself, did with people in this room, is that the estimates that you are getting depend very crucially on how you specify the equations that relate the wage you are getting to the risk of death on the job.

It is very important that you control for other sources of wage variability that are sometimes impossible to measure in a single cross-section if your risk data are also available, for example, at the industry level. This is a serious problem that I do not think people have adequately addressed.

I actually will say one thing in this regard. Recently, a well-known labor economist who shall remain nameless was asked by the U.S. EPA if he would reanalyze data from compensating wage studies. He is not somebody who is an environmental economist, he has no interest in valuing these sorts of things from the perspective of policy. The question was would he do this.

His concern was that if he did get involved in this procedure, the robustness of these estimates was just so fragile that they would really essentially disappear. As he put it, "It's not that the estimate is going to go down from \$5 million to \$2 million, it's going to disappear altogether," because it really depends very much on the econometrics of what you are doing. Jim mentioned that in most studies the assumption is that objectively measured risks of death on the job coincide with people's perceptions, which is what is really being valued here.

Finally, if you really do believe the results of revealed preference studies, you have to then go and say what possibly is happening in markets to produce these compensating differentials or these premia for safety, when in stated preference studies, as Jim mentioned, you have real problems in getting people to behave "rationally."

You can make the kind of argument we make to students in economics, which is that there are a few people in a market who really do understand the risk-dollar tradeoffs and they move the market. If those are the people whose preferences you are capturing, it goes a little bit against this idea that you are doing a willingness-to-pay study for the purposes of measuring what citizens in general are willing to pay for risk reductions.

[Transparency]

With regard to the stated preference studies, which Jim mentioned, they are encouraging but, again, you have this problem that, really, to me, at least, stands in contrast to what you see in the revealed preference studies. People repeatedly have difficulties in comprehending small probabilities.

The use of visual aids may help in this regard but, as you and John Graham found in a telephone survey, some large percent, like 40 percent, of the respondents really did not know which was the larger number, nine in 100,000 or one in 10,000. This is a problem.

(Laughter)

This problem of willingness to pay failing to increase the proportion to the size of the risk reduction for small risk reductions, I agree with Jim that this is a really serious problem. If people are perceiving these risks the way that we want them to -- and, again, they have to be perceiving them accurately for the interpretations we make of revealed preference studies to really be valid in a qualitative sense -- then this is a real problem.

This other problem, which Jim did not mention, is that in many studies, when you ask people to pay for risk reductions that, to people who work in the area of risk management, are large risk reductions, a one in 10,000 annual reduction in your risk of dying (which, at least in terms of environmental programs, is a very large risk reduction) -- in a study that I did with Allen Krupnick and coworkers, Natalie Simon and Ann Alberini, about a third of respondents would pay nothing for this size risk reduction. This is a serious problem. Presumably, risk reductions this small matter and question is, do people really value them?

[Transparency]

To end here, I understand people's desires to incorporate -- what I put here is preferences of average

citizens in setting priorities for regulation and for allocation of resources to what we might call broadly public health. But I guess the issue here is, is the best way to do this by estimating willingness to pay for quantitative risk reductions?

If you use figures that are subject to a lot of criticism, there is going to be a lot of trouble in terms of selling, I think, the results to the public and to politicians and that is why I gave the example of the benefits and costs of the Clean Air Act, because you do not have very many studies, actually, of willingness to pay for other health end points.

In the benefits and costs of the Clean Air Act, for example, there are only two health end points that have willingness-to-pay estimates behind them and when you look at the value of things like avoiding strokes or heart attacks, you use a cost-of-illness approach. If you wind up aggregating a willingness-to-pay figure, \$5 million for risk reductions, reductions in risk of death, with cost-of-illness estimates for other very serious illnesses, what kind of consistency do you get out of that?

Finally, this was actually something that came up earlier -- of course, I made up the slide yesterday -- I do think if you are taking the QALY approach, which is asking people to make tradeoffs, but not between dollars and risks, the idea of then taking something like a value of statistical life, chopping it up into a value for QALY seems totally inappropriate.

It seems to me the value of the QALY approach in cost effectiveness is that you make very clear what it is people are valuing when the decision has to be made at some point how much is enough or how much is too much when we look at the cost per QALY of some program. That is a decision that can be made in a variety of ways but I would really argue against trying to take the QALY approach and just convert it back into a cost-benefit analysis.

DR. SMITH: I have a slightly different take on all of this than Maureen. Jim and Maureen are always tough acts to follow and that is especially true today.

Let me say that I disagree with a lot of what Maureen says, but I do not disagree on fundamentals. I will try to explain why. We did not coordinate, I did not know what she was going to say until she said it, so this may seem a little bit broken at times, but I will try to go through it.

[Transparency]

I want to start with basic fundamentals in what we are doing in economic valuation. This is where one of my disagreements comes. An economic value, as everybody at this table and all the economists in the room here use it, is an extremely well-defined concept that takes a person's choice, which is all we ever observe, whether it is buying a sandwich at a sandwich store or accepting a job with certain attributes to it, and compares two situations, a base line, what you would do if you did not buy the sandwich, and everything that goes with that, and what you would get if you did buy the sandwich, and everything that goes with that, including paying for it.

Now, what economists do is they suggest that if a person buys the sandwich, it must have been worth more than the cost and they attribute that judgment to the transaction that took place, the handing of the sandwich from the person behind the counter to the customer on the other side.

Maybe the customer was a small child and I was not hungry, if I were the purchaser, but I saw that his income, his something else, was dependent on that, more was involved in the transaction than simply the sandwich, or think about a lemonade stand, where you go and buy a little cup of lemonade and you do not like it and you throw it out.

An economist looking at that transaction would say that the value of the lemonade was at least as much as what you paid for it. The point of that comment is that we want to think about what is going on as an object of choice, that there is an object of choice associated with the transaction and that translation from the choice to a value involves the economist doing a certain amount of algebra, making a whole set of assumptions about what happened if the person did not buy the good and what happened if the person did buy the good.

There is as much involved in that manipulation in the process as anything else, so when we think about how to do this, that is why I say we want to think about choices of models as well as modeling choices in determining how what we know from the value-of-statistical-life literature relates to the illness from food-borne disease.

For example, if I take the value-of-statistical-life work and I see someone accepting a job with a certain amount of risk, it does not mean that the person, in fact, attached a specific monetary value to the added risk associated with undertaking that job. It simply means he agreed to the job under the terms that were set and we do the translation.

The comparison has to be thought of fairly carefully. The implication of that -- there are lots of implications

that are very consistent with what Maureen said, because a lot of what we do with all this information really bears much less on the actual choices that people face. I am hoping that I can explain that clearly, and I will start with some three reasons for that and then go through the rest of these details.

[Transparency]

The first reason for starting with basics: When we are trying to use economic valuation information in a policy context, we are never valuing something that has ever actually been part of the choice, because if it were, we would not have to do the policy analysis in the first place. We are always talking about policy, of regulation, of change in labeling, a change in something else that does not exist, and we are trying to do an evaluation of it.

If it did exist, we probably would not be interested in an evaluation in the first place. Even the choice of organic versus inorganic food, we are not talking about mandating organic foods. We are simply saying does organic resemble the policy we deliver enough, not that we are going to get -- that is what I mean by a policy is always a constructed choice.

When we use measures of economic value that arise from different choices, we have to go to what were the base-line conditions, what was the change in those choices, how did the economists translate them to get the economic value, and what does that say about the process of using those numbers in another situation? That is what I am talking about, translating the policy into an objective choice.

Let's take three different modeling choices for commodities that involve probabilities. The first is the one that Jim presented, which is the standard gamble case in resembling and trying to explain, let's say --

[Stumbles on tape while going to the screen]

We have the utility of being sick, the utility of being well, and what we are talking about is a incremental reduction in this, so what we are saying is automatically we are converting the policy into a probability -- automatically. Now, is that the right thing to do in food-borne illness? I am not convinced of that.

On the one hand, I am disagreeing with Maureen's interpretation of willingness-to-pay evidence, but I am not disagreeing, you will notice, with her conclusion, that is, I am not clear that we know enough about the events, the object of choice, when we are delivering policies associated with food-borne illness to be able to translate that into an equivalent change in that ΔP or something about these functions that we just write down as representing utility in the two states.

[Transparency]

We could just as easily have modeled this another way. We could have said that it is about a choice among a discrete set of alternatives, that I will buy only one type of commodity, I buy only organic, I buy only pasteurized, I buy only something else, where, in this case, the P_j is the probability of some outcome associated with that commodity. This is the price of the commodity and a whole bunch of other characteristics, there are prices of other goods, and so forth. This is an entirely different structural model that will lead to different interpretations of how you use the probabilities. For one thing, you are not substituting among commodities, you are picking the best, and it will get you entirely different answers as to what to do with observed choices.

A third model, which has not actually come up, and I was asking Jim about this earlier, another person somewhere out there, Jay Shogren and Tom Crocker, have argued reasons for being concerned about the wage hedonics and about other revealed preference work, because, in effect, the risk itself that we think we are observing people are accepting, they are really not facing at all. They adjust, they undertake averting behavior, so they may think the job has a probability of P but, for them, they allocate resources, they wear safety helmets, they do other things, they take particular precautions, so the true probability is some function of the probability that would be faced if they did not take these actions, and a whole bunch of other things.

Unless we can fill in the other things, we can never figure out, reconstruct the value, that was associated with the choice, because, after all, we still observe the choice. The issue is going backwards, back to where all the action is. That is where I think most of Maureen's criticism comes from implicitly, that is, she is reluctant about whether we know enough to go backward.

[Transparency]

There are ways of going backward. I could take that same slide and fill in a bunch of details and it turns out that one can make assumptions that allow me to recover this. This is called rate and separability. If I impose that, I could actually go backward. That requires imposing a lot of structure. Again, you are taking the choices and, when you are going backward, you are adding information. Everything you get is a function of what you put in. Much of that is not observable.

The bottom line on all of this, to try to relate it to what Maureen said, is that we observe choices, people,

we do believe -- I firmly believe that people are rational, I firmly believe that they are cognizant of risks that are really important to them. When risks are explained to them --we have to really work at trying to explain. They do not think the way experts or professionals do about risk, and some professionals do not think about risk the same way, because of their training. The question is, how do we take that information and translate it into a single set of information for policy? That is the problem.

[Transparency]

Let me give my take on some of the literature and its relevance and then I will try to stop. The reality of trying to infer values of food-borne illness to models and to the policy of what we know, let's take some ERS examples. This is from a very nice volume that Tanya Roberts and a number of others edited.

[Transparency]

The events involved here resemble a fairly complex process. If you just take E. coli, for example, and look at the sequence of cases and the sorts of interventions that are either possible or not possible -- that is that pie that I had on the previous diagram -- determine what the actual outcomes are in terms of people dying or people half recovering or people fully recovering, this is the event set that we are trying to model.

Can we reduce it to a single probability where there are compound probabilities conditional on particular actions? My research suggest, and others suggest, that people do not think about compound probabilities the way the rest of us do. There is only a small bunch of us that --

(Laughter)

I do not consider academics and, in particular, economists, to be people. I am fully willing to admit that. There are more of these sorts of charts that we could lay out. The point, again, is what is the object of choice, how does it relate to the model that we are dealing with, and how does it allow us to use the choices we can observe to inform the policy about food-borne illness?

I am not optimistic about simple translations to existing studies. This is a plea for more research that looks at these dimensions and take it as it is given by someone who does research. There are some problems with that, but I am trying to be truthful.

[Transparency]

What about the studies that Jim mentioned, both revealed and stated preference? Event studies can inform us about effects but, again, we are observing choices and we are trying to reconstruct things from choices. I took a particular study I am interested in. Let's look at meat, the demand for meat, and what we know about the demand for meat and see if we can infer the value of safety from the information that is out here.

[Transparency]

This is the consumption of three different meat products just plotted as a time series. The thick one at the top is beef, the next one down is chicken, the next one below that is turkey and, at the bottom, is pork. This is consumption per capita as a function of time. This plots the average quarterly retail prices, so what we are going to try to do in this study is sort out the movements and all that junk and, of course, we have a signal. What else?

[Transparency]

Our signal is articles in Lexus Nexus that point out full-blown illness over that same time period for beef, pork, and poultry, so this is this other highly informative plot. We could have, alternatively, used the number of recalls of meat by meat types over time, and we have another really very informative signal here.

[Transparency]

What we are going to assume from all of this is that we can take this, and this, and infer what message consumers derive from that. That is what a revealed preference event study is about. My answer is not likely. We may be able to signal something very general, just an attitude more than anything else, but never use it to reconstruct a value.

[Transparency]

Should we believe people's risk perceptions, subjective beliefs? What do you think is the most important thing people face? Their lives. Death. I took people who died and looked backward at what they said. I am not kidding. The Hilton Retirement Survey allows me to take 12,000 people, look at them over eight years - - they were 51 to 61. I took a set of people who died by the fourth wave and I looked backward. I asked did they think they were going to live.

The first panel here is the people who died between the third and the fourth wave. They were interviewed in 1992. The fourth wave would be by 1998. I asked were the dead people answering the following question differently, the dead people, in wave four, than the people who are alive by the time of wave four? The question was: What do you think the odds are you will live to 75 or more?

There is a consistent decline here and, statistically, there is enormous evidence that this is a rational response, that people understand what is going on; that is, the mean level of a rescale on a zero-to-one scale of these folks is declining consistently, and you can see it in other things that are going on in their lives. The people who were alive at that point remain relatively stable in their answers to that question. I believe that people form these responses rationally. Do these responses show up in a wage equation? I was not going to talk about this, but I decided that I would, given what I have heard about older people. [Transparency]

This is the same set of individuals who work. This is a complicated overhead, because I just made it up here. The dependent variable is wages per hour that are paid to the this same group of older people, not the dead folks. These are the folks who worked by the hour. This was taken in 1994.

The dependent variables are a log of their wage rate and a whole series of independent variables out here. The one that you can focus on, this is a job list, the same sort of thing that is in the job studies all the time. I took that variable and I interacted it with age-specific variables to get at the issue of do people value risk differently as they age.

You have to look at the sum of this negative number plus each of these positive numbers to determine what their ultimate valuation is. I also have an independent measure of their attitudes toward risk, something that both Jim and Maureen talked about. This measure, as you go from 2 to 4, is an indicator of relative risk aversion. It moves the way one would think it would.

Taking all this, putting it together, three things arise. First, the results are enormously heterogeneous based on personal circumstances. People of different ages have very different values. They do not necessarily go down as a person ages.

Secondly, the people who are the most risk averse actually are willing to pay much more to reduce risk than the people who are less risk averse. It makes sense.

Thirdly, I took the same indicator that I had before and looked at the set of individuals in this sample who died after they made this job contract and asked if they knew it was coming, would they demand more compensation to accept more risk? Yes.

[Transparency]

Is this the last work on any of this? No. It suggests that economists can be reasonably constructive in the way they choose to use the data that are available. They use them to test specific hypotheses, and there is nothing wrong with that. If I were to take those same numbers and tell you I could value food safety for people 51 to 55, or 56 to 60, you should be concerned, because I am taking different choices and I am translating them into a policy outcome that is entirely different from what the persons did themselves.

The challenge that links both what Jim said and what Maureen said is reconstructing the elements that are in that choice that are relevant to the policy at hand and using it to convert -- I have more to say, but I am going to stop here in the interest of discussion.

Thank you.

DR. HOFFMAN: I know that I always find that I wish I could listen to the rest of what you have to say, but I think that it is time that we go on to take some general questions if we have any questions for this panel?

[No response]

And if we do not, I think what I will suggest is that we bring up -- I know that there were a lot of questions on risk assessment in the morning and I think there may be a few more questions on QALYs. I would ask the presenters and discussants who are still here to come to the front, and I would ask Jerry Gillespie if he would come up to moderate a general discussion.

Jerry Gillespie, as many of you know, is the director of the Joint Institute for Food Safety Research, which was just formed in the last year to help coordinate research across agencies, food safety research biological and chemical research and social science research across agencies that have responsibilities for food safety. We are delighted he could join us this afternoon and he will be moderating our general discussion panel.

WHERE DO WE GO FROM HERE?

GENERAL DISCUSSION

DR. GILLESPIE: Thank you very much, and congratulations to you and Mike for putting together, I think, an incredible session. I would like to just say a few words to maybe stimulate and direct a little bit our questions as we move to this final session.

I want to go back to what Mike Taylor said at both the beginning of yesterday's and today's sessions in terms of the roles of this conference. One of the sentences that he had was how do we improve our methods to make risk-based decisions. Then he asked four other questions that I believe are subsets of that.

One of those was how good is the information that we have and are using, how good are the evaluation

methods that we are using, what is good enough, and where are the methods going, or where are we going? I would like to maybe expand just a little bit on these questions as we think about them. How good is the information? I think one of the things that I have picked up from the conversation is that we are on the brink of a very important new discipline. I would put it to you that as you begin to emerge this new discipline, risk analysis, risk assessment, that you do so in a way that should be very much in the mind of being very rigorous and demanding, and that you always measure your success by the level that you contribute to new knowledge. Implicit in that is that you have rigor of process and you never stop looking for ways to improve the way you go about forming new information and knowledge.

How good are the valuation methods? I would reframe that and would ask the question, are we certain that the steps that we are now using are the best steps for the risk-assessment process? I think there is some value in continuing to look at that overall procedure and methodology at every step. I think there has emerged out of this session some interesting new demands in this last session that we had, in which the economists challenged themselves if their input is the right mode as we look at each step as we move across risk assessment as a whole but as we move from the assessment to the management through communication.

The next question that Michael put before us is what is good enough? Again, I would just restate, are we in the mood, or in the position, to continually investigate improvement? I ask this question because I am very curious in terms of what sorts of literature we are developing out there that looks at process, how rigorous is it, how well reviewed. I think that is a goal that we might have for this emerging discipline.

Finally, I think, very importantly, where are the methods of the process going? I think this could be a very important first step, or continuing step, as we move forward, but I do not think it should be the last. From the discussion that we now will have, I think some of what we ought to be asking ourselves is what Mike asked in a slightly different way, where do we go from here?

With that introduction, I would certainly encourage questions now to this distinguished group of participators and speakers.

DR. GOLAN: Hi, I am Elise Golan, Economic Research Service. I am a little slow on the uptake here -- this is actually a question for the group that just left.

Now we have to reconstruct the elements of the choices that are relevant when it comes to food, so are we anywhere close to that? What are the elements that are important? Is there "foodness"?

DR. SMITH: The best way to do that, it seems to me, is to talk to consumers first. Therein lies my problem, I think, with most of the simplistic approaches that we -- we economists -- use to take existing information and adapt it to a policy context. We have to learn more about how the consumer actually made the choice and what was important to him in the decision process.

I will give two examples that are relevant, examples that, actually, Maureen gave. A number of years ago, _____ and I looked at how consumers made the choice of foods that were labeled based on their cholesterol content and whether a particular bright-line approach to presenting the cholesterol reading that was necessary for men and women was important to the decisions that consumers stated they would make for low-cholesterol products.

Men, as nearly as we could tell, were completely immune. We had gone through focus groups and other things to those messages. They paid very, very little attention, bright line, no bright line, it did not make any difference. Women, on the other hand, were actually quite attuned to that distinction.

When we tried to analyze it further, the distinction rested not so much on gender as it did on who purchased the food, who actually made the food choices, something, by the way, that did not occur to either one of us when we were structuring this, which says something, I guess, about us.

(Laughter)

It probably says that we did not purchase the food.

The second example is mammograms, with a similar set of issues. How would you inform women about the risks associated with breast cancer and the appropriate strategies with respect to at what age to start getting mammograms. This was done a number of years ago. There were some conflicting messages out in the medical community about that age difference.

We picked one that seemed to be the consensus and, again, approached with a bright-line approach. In that case, the questions were asked of only women and we expected the response to be that women over the age that we had identified would immediately feel much more willing to get mammograms. That was not at all the response that we got. In that case, it was actually women who were younger than that age who changed their behavior as a result.

The message is, if we had spent more time talking to the subjects before we designed the labels, before we

designed the process, thinking that we were trying to fit this into our analytical structure, I think we would have had a heck of a lot more insight into that process. So we begin by talking to consumers. Then the challenge is translating those insights into the reconstruction of the choice process that is necessary to get at what valuation information we need.

DR. HAMMITT: I would like to comment on this area, too. It is sort of funny, we have at least two conflicting ideas. One is consumer sovereignty, figuring out how our government ought to behave. We ought to seriously inquire into what people's preferences are for tradeoffs, but we also have this idea that people do not handle risks or risk information well.

You can go into things like common Traversky [phonetic]-type work, showing systematic cognitive errors in dealing with probability. As was mentioned earlier, I guess it was Don Barnes talking about Profiles in Courage and stuff, we also have conflicting views of what our government ought to do. It ought to lead us in the direction that is in our interest, even if we do not recognize it is in our interest, but we also think it ought to do what we tell it to, we people.

All that stuff is wrapped up in this choice, where we know people systematically make a lot of mistakes. To some extent we want the government to set food safety standards, so we do not have to worry about these and we do not risk making mistakes ourselves, but we want it also not to make arbitrary rules or rules that conflict with what our preferences really are deep down.

DR. GOLAN: I am going to ask one more question just on that, just wondering. I think because we are uncomfortable with maybe each one of these measures, we say maybe we do not want to leave it all up to the people, because they could have some whims, we do not want to leave it all up to experts, we do not want to leave it all up to government, so what we do is we end up taking some cost of illness, some QALY, some willingness to pay, and we add them all up and we get some measure.

In doing that, we have lost the internal logic of each one of the measures and we end up with something we do not know what is. Do any of you have suggestions as to how we can get out of that problem?

DR. SMITH: I do not know how we can get out of that problem, but I do want to make a comment about what Jim just said, which is to say that economics does not assume that people are perfect or perfectly informed or expert in everything, or that people do not make mistakes. It assumes that people do the best they can with the information they have and the resources at their disposal, and that can mean that they are not well informed, they make bad choices, and then the question becomes how do we improve that?

With respect to the process of taking information from the QALYs, cost of illness, willingness to pay, whatever, and putting them together in a sort of stew, I think the problem there lies in the way we are presenting information as policy analysts to decision makers. We are trying to distill it down to a single attribute or number rather than explaining the dimensions that are associated with each one, and the challenge is to come up with a constructive presentation of the set as opposed to a distillation to something that is a bottom-line number that makes it easier to interpret. Decisions that are complex are not necessarily better if they are made on the basis of simple indexes.

DR. RODRICKS: May I just follow up on this last issue? I have a question for the economists on this issue of preference and choice. I am a little confused about -- I am not sure I can articulate this well -- there are choices people make over which they have almost complete control with respect to their risk. They voluntarily assume the risk. A lot of dietary choices are in that category, smoking, other kinds of things, versus those that are largely sort of involuntarily assumed risks.

Most of the risks we are talking about in the food safety conference, I think, are in the latter category. When you ask people about choice and preference, do they separate those two categories of risk? And are the results for those two different kinds of risks, if you accept that they exist, very different? In other words, if somebody else is imposing a risk, is there a view that someone else really ought to pay a lot, whereas I would be less willing to pay to reduce that imposed risk. I have not said that very well, but maybe you get the drift of it.

DR. CROPPER: One thing I would say is that if you are looking at revealed preference approaches, where you are looking at people's actual choices, they have, to some extent, to be voluntary choices. There have been studies that have looked at people's responses to the hazards associated with Superfund sites and you are making very definite assumptions about what people perceive these risks to be, which is basically that they are what they would be measured objectively to be.

If you really believe these studies, then, of course, what you are saying is that here was some risk in some sense imposed on people and they had some choice, but, of course, the folks who all of a sudden found the site in their back yards when the risks were announced, I suppose in some sense suffered a loss.

I would say that in revealed preference studies, the assumption is usually, yes, in fact, you are making a

choice, so there is a big controversy as to whether you can take estimates from a compensating wage study, such as the one Kerry presented, where, indeed, people are presumed to be making these choices voluntarily, although if you are really not knowledgeable about the risks, there is an issue there, and then applying them to a case where, in fact, the risks are not known to you and, therefore, it cannot be a voluntary risk.

DR. CASTILLO: Monica Castillo, with the Foreign Agricultural Service. This is a question, also, for the last group of panelists on the willingness-to-pay approach. I am wondering if this model could be enhanced by looking at the choices, not in terms of individual consumers but, rather, as family units.

I am thinking, just in my own experience, when I was single, before my husband and I had a family, my choices were quite different from the choices I make today in terms of whether I take my three-year-old son to a sushi bar, this sort of thing. I am wondering, is there anything in the literature, are people looking at this issue of choice in terms of family units as opposed to these individual indifference curves that were discussed?

DR. SMITH: There is a very large literature on how there are differences in the way in which households behave when there are more individuals involved in the choice. Is there a large literature on how households make decisions with respect to risk? The answer there is no, there is not. There are a limited number of studies. Bill Schultze was referred to yesterday as someone, I believe, who is working in this area -- I do not know who mentioned his name -- but he has been working in this area. I have done some work in this area.

His is revealed preference, mine was stated preference work, and there are about four or five other studies, but they do not span a sufficient set of experience either in the context of the samples studied or the risks studied to draw any clear inferences on how households behave with respect to risk-taking behavior in comparison to how individual behave.

DR. HAMMITT: When people talk about this, they often talk about individual willingness to pay, an individual risk, but most of the existing studies, certainly revealed preference studies, I think, have to be interpreted as household willingness to pay. Even these wage-differential studies, a worker takes a job that pays more or less, his change in wage influences the household resources, even though the risk of the health effect is suffered by only him or her. Obviously, the consequences of an adverse health outcome would also affect other household members.

DR. GILLESPIE: I wanted to pose a question and maybe start with Dr. Buchanan, if you would not mind, and then have responses from others on the panel. If we were to have another outbreak, and let's not try to define it other than it is a micro-food-borne pathogen that causes a serious episode similar to, let's say, the E. coli 015787 that started much of it, if you were to design a process in which the risk assessment would perhaps begin by looking at pathogens' risk for causing public health damage, then how would we weigh in, in the decision-making process, with social science, including economic weighing into the problem?

DR. BUCHANAN: Let me respond not just as a microbiologist but I am going to respond now as a regulator and as a risk manager -- I am going to put on my other hat, and try to take the conversations that I have heard this afternoon and put them into a bigger framework.

The regulatory agencies are charged by the people of the United States through Congress to regulate or safeguard them in specific areas. In the case of the Food and Drug Administration, we do food, drugs, cosmetics, et cetera. They also mandated that we do it through a particular process as outlined in the Procedures Act. It means a process where we go and we ask the people whom we regulate about the regulations we are going to put forward. We go through proposal and final rulemaking in this regard. In that process, we ask for opinions and we go to great effort to ask for those opinions. I think that we should look at today's subjects in that light. What we are doing in the process of doing risk assessments and doing cost-benefit analyses or choices, or whatever, is we are helping to take those questions we are looking for information on, that the regulatory agencies are charged with making the final decision, but we must do it in consultation with our stakeholders, is we are trying to separate the questions of "I know" from "I believe" from "I feel," so that we can deal with them individually.

All three of those are important. "I know" is our science base. "I believe" is our value judgment, our choices. "I feel" is the gut emotion that our consumers have about food safety. I think we need to view this and what we are doing here in that light. We are trying to help us have that discussion so that we can reach a final decision and set up a program that responds to the American public.

Would I say what happens in the next process? You need all that stuff, because sooner or later you develop new regulations because the old system no longer meets the demands of the new world. The world has changed, the old approaches for safeguarding the public do not work, so we have to develop new methods,

so we go out and, in order to get the best method that works, we do this through risk assessment, through research, through economic analysis, so that we can get one that meets the needs of the country. I know that is not exactly what you were looking for, but I thought it would also be helpful for setting the framework for what we are discussing.

DR. GILLESPIE: One great thing about ignorance of the field, I was not looking for anything. (Laughter)

DR. WACHSMUTH: I will react to your first word. When you said "outbreak," I would fall back to some of the things that Rob Tauxe said yesterday in terms of emergency response. That would be much sooner than we would be able to write a regulation or do anything in a regulatory sense, and we do have acts that cover that, the Public Health Act, and other ways that we can do it.

In epidemiology, some epidemiologists think that the application -- their field application of that science is, in fact, a type of risk assessment. If they are out there fast, they do the case-controlled studies, they get the statistics, it not only applies to that outbreak, but if they can identify the vehicle, then we have got -- it took a while with 15787. I think the first association epidemiologically was onions with McDonalds back in 1982.

Eventually, after a couple of months, it was hamburger. Then we got the pathogen, then we found out about the toxin. It still took a rather heroic effort on the parts of some people to do something about 15787. I think now we are still working with that on that basis, but the outbreak response and doing some things immediately and thinking of those in terms of a risk assessment is an existing approach, and we are refining it.

DR. WOTEKI: My name is Kathy Woteki. I would just like to add a footnote to that, too, because from the risk manager's perspective there have been many instances when there has been an outbreak-type of situation or a toxic substance that comes to our attention that is in a food product that is being fed through a population that we believe would be more vulnerable to the effects of that.

The industry argues, well, you cannot take any action, because you have not done a specific risk assessment. I would hate to see our public health system put in the position of having requirements, not only for risk assessment but also for detailed economic analysis before they can deal with those emergency types of situations that come up.

A lot of what we have been discussing, both yesterday and particularly today, are methods that are very good when you have a longer term policy development that you are dealing with. I think, Kaye, your presentation on S.E. risk assessment in eggs was a terrific example of that, where it was done prospectively, the policy development and now the rulemaking is going following after that. It just does not work in an emergency kind of situation and it should not be put as a roadblock.

MR. TAYLOR: This has been a really interesting discussion for me. Sandy knows that I do not really understand anything you have been talking about, but I have been hearing about these methods for a long time. First of all, I really appreciate the candor of the professionals in the economics field who have been talking to this interdisciplinary audience about their methods, and being really candid about strengths and weaknesses and raising, I think, very fundamentally, important questions about how these valuation methodologies can fit into food safety decision making and policy making.

I want to put a sort of hypothetical to the whole panel, including the economists but also the risk assessors, that gets at this sort of broad program design, system design, and resource allocation question that we started out talking about yesterday.

CDC says latest best information is we have 5000 deaths, 325,000 hospitalizations, and 76 million illnesses associated with food-borne pathogens. Of course, we know, I think it is, there are 28 pathogens that they are taking account of, and we know it is a great diversity of illnesses and reasons for hospitalization and the nature of the illnesses people have.

But I would like you to put yourself in the position of a senior adviser in some fictional food safety department. Someone has walked in the door with a charge to do the best he can with his billion dollars to reduce the risk of food-borne illnesses and you have until the end of the day to tell him what are the things he ought to be thinking about. What are the key bits of information he ought to have? What are the questions he ought to be asking to figure out how to make the best use of that billion dollars to reduce the risk of food-borne illness, as diverse and as difficult as it is?

You have some data about risk, you have some data about value, ERS has done a lot of work, you have your methods. What should the policy maker be thinking about, at least, in figuring out a plan, some approach, to using that money to reduce the risk of food-borne illness?

I would be interested in individual perspectives and, to the extent that the group wants to figure this out

collectively, you have 15 minutes.

(Laughter)

That is fruitful to think about. It is not that we are going to come up with a plan sitting here, but I hope it will focus a little bit on what we ought to be working on, kind of going out from here, in terms of what information we ought to be developing, data, risk-related information as well as economic information, and then focusing on really working on how we ought to be using the existing knowledge about valuation to set priorities. That is what it comes down to.

DR. WEINSTEIN: I will start. I would start with some sort of comparative risk analysis. There are only 28 pathogens, did you say? Well, that is not bad.

(Laughter)

I wish I had seen the tree that was attributed to Tanya that was presented by Kerry for E. coli, because that is exactly the kind of information that would enable a QALY assessment, the number of QALYs lost per case of E. coli. You have the different branches with probabilities assigned to the branches and you can attach the number of QALYs lost, so you can estimate that.

The only other piece of information you need was presented -- well, in principle, it was presented in Bob's talk -- the kind of risk assessment that gives you the probabilities. You take the probabilities, the consequences in terms of QALYs, and make a table. I am going to stop short of saying I would multiply the probabilities by the QALYs. I would multiply them, because you sort of cannot help yourself, but --

(Laughter)

I would also want to look at the separate data. I would not want to aggregate everything into a single number. To be quite honest, I was quite distressed to hear that that some agencies are not allowed to use cost-effectiveness analyses, so they have to turn them into cost-benefit analyses by multiplying the number of QALYs by dollars. I fully agree with Maureen, I think it defeats the purpose of doing QALYs, which is really to try to avoid a lot of the problems of putting a monetary value, but I would at least lay out a table. Where data do not exist, you do ranges, and you get experts and you say, okay, these are the top five, or 14, or whatever it is, and this is the next group, and this is the next group. Then you look more closely near the top of the list and maybe set aside the ones at the bottom of the list.

DR. BUCHANAN: Let me respond to that, because I think you also -- and I do not disagree with your process. But also, as you go through that decision-making process, you have to understand that the only thing we can be sure of in this process is change. As you go through this process, there has to be a portion of the resources kept aside in terms of what is going to happen next.

It has to be the emergencies that CDC has talked about. It has to be that if we are successful, for example, say, microbiological concerns bubble to the top -- surprise -- and we really put a concerted effort there. I will guarantee you that the pendulum will swing and dioxins will show up somewhere and we are going to have that problem, and then we are going to have another one.

There is a core structure that has to be maintained to be able to respond to the changing world. While we may want to move things around in response to risk, there also has to be that foundation that we can always fall back on, or at least readily reconstruct, because once you shut down a program, it is almost impossible to get it started again.

DR. WEINSTEIN: I could not agree more with that. In fact, I just came from another conference involving pharmaceuticals and assessments of new drugs, and so on, where one of the major criticisms of the kinds of risk assessment and risk-analysis models is that they are sometimes viewed as etched in stone. Once a study gets published in The New England Journal of Medicine, it is considered the 11th commandment.

None of these analyses should be viewed that way. This is really important. Whatever you do, it is good only when you did it and it may not be good tomorrow. If there is some new evidence, whether it is an outbreak, a new scientific study, whatever, it has got to be fluid.

DR. SMITH: It sounds to me as if there are three elements in the process that have been talked about. One, that actually gets at Jerry's earlier hypothetical question to the panel, which was the outbreak and some responses we heard. If you say I am going to wait until the outbreak is over to do anything, or until it has played itself out, because I have to do a benefit-cost analysis or risk assessment, there is no longer a decision, that is, the outbreak is gone, it is over with, there is not going to be a policy, because there has been a decision by doing nothing.

The first category that has been talked about here is the value of a quick response, rapid response, and the research on rapid response as a policy. In other words, what is the configuration of a team or set of teams that should be associated with rapid response? That is an allocation for the budget.

There is a second you might talk about, which is what Mil's comments were principally on, which is

refining the food delivery system; in other words, prioritizing today's problems and asking what are the most important problems as we know today's system and how do we make choices in that?

There is a middle block that is associated with that and how much resources should go into that? Probably the lion's share.

Then there is a third block that is associated with the longer term issues, that which is associated with emerging new science, that which is associated with the changes in our understanding of how people make choices, or how they are going to live their lives; that is, if one were to look at the fraction of meals that was eaten away from home 25 years ago as opposed to the fraction of meals that is eaten away from home now, and the conduits through which households get exposed to food-borne illness, they would be entirely different.

How does one, then, look ahead in thinking about making the system more responsive? There is that portfolio of policies to be thought about. I think the first one and the last one tend to get left out and we focus our attention on the middle.

DR. GILLESPIE: Is there another comment relative to Michael's proposition before we move to other questions?

DR. HOFFMAN: Since Laurian and Helen are not here, I would say the only thing that has just been overlooked -- I think we would all agree with this -- is that within any of these, but particularly within that second group that Kerry is identifying, refining the current food-delivery system, that we also have to think about the mechanisms we have for reducing risk and have to consider the relative costs and effectiveness of those different alternatives. That may be another piece of information we would want to have in that.

DR. HAMMITT: On the second block of stuff, what is most important is the change in risk resulting from different sorts of interventions we could consider. While we have talked a lot about valuation, because the four of us, at least, are fascinated by it, I think that is really of second-order importance, because a lot of these risks can be both fatal and nonfatal, so valuation helps us do two things.

One, it sets the bar for how cost effective must an intervention be for us to do it, and that is important, but benefit-cost analysis is not going to be determinative anyway, so whether we know exactly the value or not is not that important for the intervention that will be required or adopted.

The other thing valuation has to do with is when reducing risks, are there fatality versus various nonfatal conditions, or how important are the nonfatal conditions versus the fatalities. If there are fatalities that are particularly common in kids versus adults versus older people -- there is some reallocation at the margin, but all this stuff, I think, is within an order of magnitude or so.

My sense of our estimates of what the probabilities of harm are, and the change in probability of harm due to different interventions, those are uncertain by a couple of orders of magnitude in many cases. I would not let uncertainty about exactly how to value things slow us down too much on going after the big things, where we can do something to actually reduce the probabilities.

DR. GILLESPIE: Along those same lines, and I think pursuant to what Bob and Kerry said, we now know, looking in retrospect, how a pathogen like E. coli 0157 became a more serious pathogen by looking at how it emerged and evolved.

From that sort of information we might be able to begin to do more anticipatory sorts of research that could, in fact, help us direct policy ahead of catastrophe. I think we need to be very careful about directing resources without also having them be a part of the up front.

MS. CHITA: My name is Audrey Chita [phonetic]. I am a AAAS risk policy fellow. In regard to the rapid response to emergency, FEMA has a disaster cycle that it has, so it is always in a preventive, preparation, fast response, and then deconstruction afterward, what did we do right, what did we do wrong.

Maybe food safety could benefit from that sort of model, also. I would see risk assessment as being in the preparedness area, not in the disaster response immediately after an emergency has been identified. That is not the place for risk assessment. It is like a cycle. That is just a thought I had.

Another comment I had, and you can let me know what you think of this idea, or if anybody has done it, or how it is being done. Bob Buchanan showed a slide where he had various levels of pathogens in one column and then cases of food-borne illness in another column. It was like the lower the limit, the fewer the cases, which is intuitive, right?

To me, what was missing in that is a third column saying the cost, so that I could get some sense of the point where you reach diminishing returns, because it may be that to get down to that 0.1 people, I guess, out of a population size for illnesses, maybe that cost, in dollar amounts, 300 times more, or several orders of magnitude more than the step before it.

If you had some kind of gauge at what point your limit becomes -- well, you always hear this word

"feasible." To me, feasible and practical means at what point you reach diminishing returns, where you just keep putting in more and more money and you are never going to get that last percentage point in.

DR. WEINSTEIN: That is exactly what cost-effectiveness analysis is all about, and the idea is that you would want to spend on each hazard, on each risk, up to an approximately equivalent point. If you are spending a million dollars per QALY, or \$100,000 per QALY, or \$10,000 per QALY, you at least want to be in the right ballpark. You do not want to be spending up to \$10 million per QALY on one area and not at all in another area where there is a potential gain. Yes, I do agree with that comment.

DR. BUCHANAN: I do want to caution a little bit using the FEMA model for a new emerging food safety problem, because that is based on responding to an immediate problem. Emerging food safety concerns -- and too often we respond that way. The actual response to a food safety concern, a new emerging problem, is, one, taking care of the public health concerns of an immediate nature, dealing with the people who have gotten sick, stopping it from spreading, et cetera, but it does not stop there.

To truly be effective at controlling emerging concerns you then have to mobilize a research force. You need to identify a potential or, hopefully, a number of potential interventions that can be used. You are then probably going to have to develop new public policy as a result of that. It is only when you actually then have gotten to the point where that food safety concern, this new problem, is under control, can you then say the emergency is over, that you now have it as part of your normal delivery system. You have gone from step one to step two in the model that was presented.

It is only then that your job is done, so the FEMA model gets that you start it, but it does not really get you to where you have moved to stage two, making it institutionalized.

MR. HEPTON: Tony Hepton, Dole Food Company. Because food emergency, food safety outbreaks and emergencies are rarely agency-specific, I think it would be wise to have a mechanism that would allow, and maybe require, collaboration between agencies and, at the same time, a mechanism for coordination of efforts to deal with emergencies and, at the same time, a recognition that collaboration and involvement of the various stakeholders, particularly the industry, that may have a lot to add to the understanding of the problems.

We need to identify how we can bring together the diverse resources that we have across the nation under some coordinated effort, so that we truly do have a method of dealing with the problem but not leaving out vital pieces of information or, at the same time, coming up with potential causes -- and this has happened in the past -- where only half the information is really being used and we already are being told where the problem is, and sometimes this is a problem of states speaking up before they have all the information, or certain segments of the government speaking up before they have all of the information. That can be very damaging to large segments of the industry.

What I am looking for is there a way in which we can bring together the resources in a cooperative way, a coordinated way, so that we can truly focus on problems and not consider them agency-specific.

DR. WACHSMUTH: In the existing situation, not thinking of the ideal, that coordination is done as best we can with liaisons and constant contact within the federal agencies, but it is highly dependent on the recognition locally by the state and the state's appropriate response and then they are bumping it to a federal level, in most cases. Sometimes it takes a federal view to see something happening across states to recognize you have a problem, but most often, I think, it happens locally and then it builds up from there. There is definitely a need for that kind of network and an ongoing team.

I wanted to clarify, too. I think maybe my first comment was misunderstood about the epidemiology. That would be at the time it has hit the federal government, when the Band-aids have been applied locally, and someone comes in with some analytical skills and tries to tease out, as Bob said, the real problem, finding associations, and then acting on that. It may not be in a regulatory sense, but it would be much more far-reaching and it would have to be done in concert with industry.

We have a recall group within my group at the Food Safety and Inspection Service, but we do not do the recall, even with all the resources that FSIS has. We are totally dependent on the industry, which knows all of its conduits, knows the consignees, and can get back the product that might be out there, so they have to play a key role.

MR. DORSEY: I am David Dorsey. I work for Senator Kennedy on the Public Health Subcommittee of the Senate Health Committee. I take it that by way of disclosure -- I am a lawyer by training, I have worked at FDA and, before that, I was a mathematician, so I am the bookends, maybe, of the disciplines, a lawyer on one hand and then really kind of hard numbers on the other side.

It seems to me a question posed here has been whether we need a new food safety system, both in terms of the laws and organization, which is based on risk assessment. As I take it, that is the question posed here

and, at its most extreme, the way to pose the question is should Congress make a broad delegation to an agency or several agencies to make decisions solely based on risk assessments?

One way to pose the question is should Congress do that now? My sense, from having listened to the last two days, is that the discipline is much too immature for Congress to do that now. I would be interested in people's reactions.

The second question, then, is, should the disciplines mature, should Congress ever do that? For example, the statutes the Congress has now passed include safety standards in there, which are not necessarily risk-based safety standards. Food shall be safe. There is not a question of risk and benefit or weighing of harms. It is merely that things be safe.

The impetus here seems to be, perhaps, an illegitimate standard for Congress to impose on the agencies when enforcing the safety of the food supply. I would like people's reactions to that as well. Thanks.

DR. WEINSTEIN: I would like to comment on both points, actually. Except for the word "solely," I think that

-- I do not think there should ever be a law that says that any agency should rely solely on risk assessment, ever, or risk analysis. Should they have the information? Yes. Should there be a mandate that they develop the information? Maybe.

I think there is a difference between saying the agency shall do risk analyses than to say they shall make decisions based solely on those risk analyses. I am not a lawyer, so I do not know if that is tenable, but that is where I would be. I would say, yes, they should generate the information, but stop short of saying that should be the sole basis for decision making.

I am sorry, I forgot the second point, but I know I had a comment on it.

DR. CROPPER: I would like to add something for a second to what Milt said. I think the concern, at least on the part of economists, when you talk about the standard is making the food supply safe is that you are ignoring the costs of this. The idea of at least working toward cost effectiveness, even if you cannot do it fully, is that at least you are attempting to be transparent about what the physical benefits will be in terms of risk reductions, and you are also attempting to look at what it is going to cost you.

I think if legislation can, shall we say, at least encourage that regulatory decisions lay those sorts of things out, I think that is very important. One of our reasons for participating in panels such as this -- I include the economists here -- is really this idea that you want to have consistency in decision making across agencies, which is another reason why you would maybe like to urge this in the case of food safety but it would also be good to do it in other health and safety regulations.

You do not want to necessarily require, as Milt said, that this be the sole basis for regulation, but I think you are likely to get more efficient and certainly more transparent regulations if you try to promote the formulation, the gathering of this kind of information.

DR. WEINSTEIN: I remember the second point, because it struck a responsive chord in me. You were speaking of the maturity of the discipline and whether the discipline is ready to be used. I will put this in the form of a question, whether the biologists in the room would say that biology as a discipline was ready to be used for public policy 30, 40, 50 years ago when it started to be used.

We did not know anything about the human genome in those days. Now we know about the human genome. Looking back, was the discipline ready? Well, I do not know. What about physical science before quantum mechanics? It is a question of making the policy based on the best available science at the time that the policy has to be made.

I would say, yes, the discipline is ready to be used. Is it going to stay the same? No. It is going to get better, just as biology and physics and chemistry have matured over the years.

DR. BUCHANAN: I would like to reinforce those. We use the best that we have available and we have a system that is continually improving. I would urge you to go back and look at the food safety strategic plan that came out six months ago, because it laid out a framework of continual improvement. It laid out a framework for becoming more risk-based and science-based. It laid out a framework for going back and looking at the underlying legislation, as we heard from Professor Merrill.

It really does lay out a framework for how to take risk in those types of sciences and disciplines and apply it to a continually improving system of food safety, regardless of the organizational structure.

DR. HAMMITT: You seemed to me to be proposing as an alternative to a risk-based system a law that says food shall be safe. I wanted to flesh that out a little bit in that it is not clear exactly what safe means, but it is clear it does not mean without risk, because that is physically impossible in most settings.

Every kind of food we eat, any food, can be contaminated by bacteria. In that sense, none of our food is perfectly safe. A lot of the nutrients are harmful in some situations, and the like. So safe does not mean

without risk. It means safe enough in the context of our alternatives for making it safer, so I think it is very close implicit benefit-cost tradeoff.

MR. DORSEY: If I could respond, the first thing is, my question was not to suggest that these tools should not be used at all, which is the way it was interpreted, but I do think -- I mean, as we saw, for example, in the case before the Supreme Court pretty recently, there is an argument that the Clean Air Act should have been interpreted solely as a risk-based analysis or risk assessment or cost-benefit analysis standard.

There is a kind of ideological push to say this is the only input to these decisions, so that is a concern that I have, that it should not be the only input to the decision. There seem to be people who seem to be suggesting it should be the only one or a much more decisive one than I might like or others might choose. That is the first matter.

The second matter is, I think it is true that food additives, things put into food, the standard in the statute is "safe," which has been interpreted to mean reasonable certainty of no harm. That is not a risk-benefit analysis; it is just saying we will have uncertainty as to what we know about the harms that result, but the notion is that there no harm, period.

In fact, with some things that go into our food supply, it is not a risk-based analysis at all. That is just the way the law is and that is the way I understand it to be implemented by the Food and Drug Administration. Microbials, that is a completely different side of the equation. Those are not food additives, those are not regulated in that way.

There are parts of our food regulatory system that are very much a safe kind of standard.

DR. HAMMITT: You were mentioning about being more consistent. If you are suggesting that we cannot enforce reasonable certainty of no harm with regard to microbes, and so we do not write the law that way, we write that standard with regard to additives intentionally added, that would suggest we might be spending a lot of resources or giving up a lot of food production, or something, in order to meet this high safety standard on the additives, and that cost-effectiveness analysis would suggest that we could relax our efforts a little bit on the additive side, spend more on reducing microbes and, therefore, spend the same amount and be safer in total.

That relates to the question of do you actually care if you die from an anthropogenic additive or a natural constituent and how important is that difference? In the risk-assessment world we assume that difference cannot be huge, but that is an open question.

DR. WACHSMUTH: This is a very simplistic comment but it is always present in terms of risk-assessment discussions. I am not sure we have to consider a full-blown quantitative microbial risk assessment the only risk assessment. I think the things that were done -- and the HCCAP regulation is not a perfect document, probably, although it is close --

(Laughter)

I think that in the regulatory impact part of that the way that we approached looking at the reduction in illnesses, and the benefits and things that ERS was able to associate, that is still a risk assessment and, to me, a valid way to look at a public health problem when you do not have all of the numbers. To say that risk is only the formal qualitative thing is a little disturbing.

DR. BUCHANAN: I would also like to throw out the hypothesis that the Food, Drug and Cosmetic Act and the appropriate acts over it at FDA are very compatible with a risk-management approach to food safety.

Those terms like "reasonable certainty of no harm," the term "certainty" is a relative term and it has not been totally defined. It is a management-type decision to the level at which certainty is interpreted.

I do not think it is quite as clear-cut as most people would like -- they think is there. Certainly phrases like "may render" are, again, a risk-based approach to managing the risks that we have in the food supply. I find no problem with compatibility of a risk-management approach to food safety in even the current legal system.

DR. GILLESPIE: I want to thank the panel very much and I want to thank all of the participants for their very thoughtful questions and suggestions. I think at this time someone is going to sum up.

MR. TAYLOR: I just want to add my thanks, again, to this panel. This was extremely stimulating and informative for me, and thought-provoking, as I hope it was for everybody here, and thank all of you for the participation. For us this has been a really profitable couple of days. I have personally learned a lot. It has stimulated thinking that will lead to some further activities here at RFF on this subject. We need to figure out what they are, exactly. We are completely open to suggestions from all of you. We look forward to engaging you in what we do in the future as well. Again, thank you very much.

[Thereupon, at 4:45 p.m., the meeting was concluded.]

