DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

CENTER FOR DRUG EVALUATION AND RESEARCH

DRUG SAFETY AND RISK MANAGEMENT ADVISORY COMMITTEE (DSaRM)

Volume I

Wednesday, May 18, 2005

8:05 a.m.

Holiday Inn
8777 Georgia Avenue
Silver Spring, Maryland

PARTICIPANTS

Peter Gross, M.D.,Chair
Shalini Jain, PA-C, Executive Secretary

MEMBERS:

Stephanie Y. Crawford, Ph.D., MPH
Ruth S. Day, Ph.D.
Curt D. Furberg, M.D., Ph.D.
Jacqueline S. Gardner, Ph.D., MPH
Arthur A. Levin, MPH (Consumer Representative)
Henri R. Manasse, Jr., Ph.D.
Louis A. Morris, Ph.D.
Richard Platt, M.D., M.Sc.
Robyn S. Shapiro, J.D.
Annette Stemhagen, Dr.PH (Industry Representative)

FDA CONSULTANTS (VOTING):

Elizabeth Andrews, MPH, Ph.D.
Sean Hennessy, Pharm.D, Ph.D.
Allen Mitchell, M.D.

FDA STAFF:

Mark Avigan, M.D., C.M.
Gerald DalPan, M.D., MHS
Carol Holquist, R.Ph.
Paul Seligman, M.D., MPH
Anne Trontell, M.D., MPH

C O N T E N T S

C O N T E N T S (Continued)

P R O C E E D I N G S

Call to Order

DR. GROSS: Good morning, everyone. I am Peter Gross. I am Chairman of the DSaRM Advisory Committee. Today our job is to explore the issues related to FDA's risk assessment program for marketed drugs and I think it should be a very interesting meeting. Shalini Jain has a few comments.

Conflict of Interest Statement

MS. JAIN: Good morning, everyone. I am now going to read two conflict of interest statements. The following announcement addresses the issue of conflict of interest and is made part of the record to preclude even the appearance of such at this meeting.

Based on the submitted agenda, the advantages and disadvantages of the current system for safety signal detection and proposals for short- and long-term ways to improve the current system, and all financial interests reported by the committee participants, the agency had determined

that all interests in firms regulated by the Center for Drug Evaluation and Research present no potential for an appearance of a conflict of interest at this meeting, with the following exceptions: In accordance with 18 U.S.C. 208(b)(3), Dr. Richard Platt has been granted a waiver which permits him to participate in today's discussion.

A copy of this waiver may be obtained by submitting a written request to the agency's Freedom of Information Office, Room 12A-30 of the Parklawn Building.

In the event that the discussions involve any other products or firms not already on the agenda for which an FDA participant has a financial interest, the participants are aware that they need to exclude themselves from such involvement and their exclusion will be noted for the record.

With respect to FDA's invited industry representative, we would like to disclose that Dr. Annette Stemhagen is participating in this meeting as a non-voting industry representative acting on

behalf of regulated industry. Dr. Stemhagen's role on this committee is to represent industry's interests in general and not any one particular company. Dr. Stemhagen is employed by Covance, Inc.

With respect to all other participants, we ask in the interest of fairness that they address any current or previous financial involvement with any firm whose products they may wish to comment upon.

I will now read the second statement. The Food and Drug Administration has prepared general matter waivers for the following special government employees: Drs. Louis Morris, Peter Gross, Elizabeth Andrews, Ruth Day, Sean Hennessy and Allen Mitchell, who are participating in today's meeting of the Drug Safety and Risk Management Advisory Committee on the types of population-based studies that can be used to assess safety, for example, clinical trials for new indications registries, Phase IV postmarketing studies and epidemiological studies.

This meeting is being held by the Center for Drug Evaluation and Research. Unlike issues before a committee in which a particular product is

discussed, issues of broader applicability, such as the topic of today's meeting, involve many industrial sponsors and academic institutions. The committee members have been screened for their financial interests as they may apply to the general topic at hand. Because general topics impact so many institutions, it is not practical to recite all potential conflicts of interest as they apply to each member. FDA acknowledges that there may be potential conflicts of interest but, because of the general nature of the discussions before the committee, these potential conflicts are mitigated.

With respect to FDA's invited industry representative, we would like to disclose that Dr. Annette Stemhagen is participating in this meeting as a non-voting industry representative acting on behalf of regulated industry. Dr. Stemhagen's role on this committee is to represent industry interests in general and not any particular

company. Dr. Stemhagen is employed by Covance, Inc.

In the event that the discussions involve any other products or firms not already on the for which an FDA participant has a financial interest, the participants are aware that they need to exclude themselves from such involvement and their exclusion will be noted for the record.

With respect to all other participants, we ask in the interest of fairness that they address any current or previous financial involvement with any firm whose products they may wish to comment upon. Thank you.

DR. GROSS: Thank you, Shalini, we are already ahead of schedule! The opening remarks will be by Dr. Paul Seligman, who is Director of the Office of Pharmacoepidemiology and Statistical Science.

Opening Remarks

DR. SELIGMAN: Good morning. It is a pleasure to be with you this morning to welcome you to this two-day meeting of the Drug Safety and Risk

Management Advisory Committee. In addition to our advisory committee members, we are pleased to have Dr. Elizabeth Andrews, Sean Hennessy and Allen Mitchell, acting as special government employees, at the table to participate in this discussion.

The purpose of these sessions is to review FDA's postmarketing drug surveillance in epidemiology programs. By describing our current efforts, we are eager to get feedback from the committee and the public on ways to strengthen, improve and/or redirect our program to ensure that agency resources are best applied and that public health is bettered by ensuring that the benefits of medicines are maximized and their attendant risks are minimized.

In November, 2004, Acting Commissioner Crawford announced a series of initiatives to strengthen the agency's ongoing commitment to drug safety. One of these initiatives included more frequent public discussions on important safety issues. This meeting today and tomorrow is but one step in fulfilling that commitment to greater

transparency and openness in our processes and deliberations. The discussion of our postmarketing surveillance in epidemiology programs, warts and all, is entirely consistent with this commitment.

As part of the PDUFA III agreement in 2002, the agency committed to publishing guidance to the pharmaceutical industry on best practices and pre- and postmarketing risk assessment and risk management. After an extensive public process, concept papers, public meetings, draft guidances, we published our final guidances on pre-marketing risk assessment, on good pharmacovigilance and pharmacoepidemiologic assessment and on the development of risk minimization action plans. These documents reflect the importance that we place on developing high quality safety information throughout the clinical testing and subsequent marketing of a drug product, and on the additional educational and medication practice steps that can be taken to ensure that products are used safely and wisely once in general use.

In addition to guidance to industry, we

have recently completed guidance to our own FDA reviewers codifying good review practices, including a document focused solely on the conduct of clinical safety reviews of a new product application.

As you will be hearing today, we are actively exploring ways to improve our analyses and access to information through our electronic adverse event reporting system, or AERS system, through the use of data mining. In 2003 we completed a strategic plan to update our AERS system that we will implement over the coming years to make adverse event report data more readily analyzable to our own staff and are readily analyzable and accessible to the public and researchers outside the FDA.

All of these efforts are directed towards ensuring that we, at the FDA, and the American people have the best information upon which to base public health and regulatory actions on the use of medicinal products. Today and tomorrow we will be exploring issues related to FDA's risk assessment

program for marketed drugs.

DSaRM members and members of the public will be hearing all about the methods FDA uses to assess drug risks. These include the review and analysis of spontaneous reports of adverse events; medication error reports; drug use data; data from healthcare administrative data sets; the use of epidemiologic and observational studies; the use of clinical trials in the postmarketing arena, as well as active surveillance programs. You will also be hearing about various techniques such as data mining; about the use of registries, particularly in the context of pregnancy registries; and the role of Phase IV studies in the postmarketing arena.

We are posing a number of questions related to our use of both passive as well as active surveillance methods: What types of safety problems are most effectively addressed by the use of a surveillance system based on voluntary reports? Under what circumstances are such passive approaches to detect safety signals ineffective?

And, how can the use of our adverse event reporting system be improved to detect safety signals?

In addition, you will be hearing presentations on FDA's efforts on more active surveillance. How can new systems be used to augment or enhance our safety signal detection and risk characterization activities? What types of drug products, safety problems or situations are best suited to these more purposeful or active surveillance methods? We will discuss our efforts at monitoring drug utilization and ask questions about how best to use and where we might expand or improve the use of such data.

Finally, we are interested in hearing from the committee about what are significant data gaps, and how the Office of Drug Safety should go about filling this potholes. We are committed to continuous improvement of our programs.

If there was a perfect way to monitor drug safety that would work in the United States, we would have eagerly embrace it, but our healthcare and public systems are complex. To adequately

assemble the picture of drug safety requires a pallet with many colors and access to an array of brushes. We have a very full agenda these two days and many presentations and a full slate of questions for the advisory committee to discuss. Mr. Chairman, thank you and members of the committee for being here today and tomorrow. We look forward to the advice from both you and the members of the committee. Thank you.

Introductions

DR. GROSS: Paul, thank you for that preview of coming attractions. Before we go on, I would like to go around the table, starting to the right with Dr. Crawford, ff each of you would introduce yourselves, say where you are from and your main area of interest. Stephanie?

DR. CRAWFORD: Good morning. Stephanie Crawford, University of Illinois at Chicago College of Pharmacy, and my general interest safe medication systems from health services research perspective.

DR. MITCHELL: Allen Mitchell, Sloane

Epidemiology Center, Boston University. Interests are postmarketing studies, with a specific focus on specifically mounted epidemiology studies for that purpose.

DR. STEMHAGEN: I am Annette Stemhagen. I am an epidemiologist from Covance, and I am the industry representative on this committee.

DR. PLATT: I am Richard Platt, from Harvard Medication School, at Harvard Pilgrim Healthcare. The reason I was granted a waiver is that a fair part of my professional activity is involved in postmarketing safety surveillance activities.

DR. ANDREWS: Elizabeth Andrews, from Research Triangle Institute, an epidemiologist, primarily interested in epidemiologic studies of drug safety and risk management.

DR. DAY: Ruth Day, Duke University, Medication Cognition. I am interested in the ease with which people can find, understand, remember and use drug information in a safe and effective way and to laboratory studies on that.

MR. LEVIN: Arthur Levin, Center for Medication Consumers in New York. I am the consumer representative on the committee.

DR. SELIGMAN: Paul Seligman, FDA.

DR. TRONTELL: Anne Trontell, Deputy Director of the Office of Drug Safety.

DR. AVIGAN: I am Mark Avigan, Drug Risk Evaluation in the Office of Drug Safety.

DR. DALPAN: I am Gerald DalPan, Director of the Division of Surveillance, Research and Communication Support, in FDA's Office of Drug Safety.

DR. HOLQUIST: Carol Holquist. I am the Director for the Division of Medication Errors and Technical Support, in the Office of Drug Safety.

DR. GARDNER: Jacqueline Gardner. I am Professor of Pharmacy, University of Washington in Seattle. My interests are pharmacoepidemiology, drug risk management and pharmacy practice implications.

DR. MANASSE: I am Henry Manasse. I am with the American Society of Health-System

Pharmacists. My interests are in public policy and safe use systems for medications in hospitals and health systems.

MS. SHAPIRO: Robyn Shapiro. I am a professor of bioethics and Director of the Center for the Study of Bioethics at the Medical College of Wisconsin.

DR. MORRIS: I am Louis Morris, Louis Morris and Associates, interested in risk management and risk minimization.

DR. HENNESSY: Good morning. My name is Sean Hennessy. I am an epidemiologist and pharmacist at the University or Pennsylvania, and I do drug safety research.

DR. GROSS: Curt Furberg will be here momentarily. I am Dr. Peter Gross. I am Chair of Medicine in the Department of Internal Medicine, Hackensack University Medical Center, and Professor of Medicine in the New Jersey Medical School. My main interest is in quality improvement in health care.

MS. JAIN: Shalini Jain, health science

administrator, executive secretary for the committee.

DR. GROSS: The next speaker is Dr. Joyce Weaver, safety evaluator, Division of Drug Risk Evaluation, Office of Drug Safety.

Using the FDA's Adverse Event Reporting System (AERS) in Postmarketing Surveillance

DR. WEAVER: Good morning.

[Slide]

My name is Joyce Weaver and, as Dr. Gross said, I am a safety evaluator in the Division of Drug Risk Evaluation within the Office of Drug Safety at the FDA. I am going to talk about how the safety evaluators in our division use the FDA's adverse event reporting system to monitor the safety of drugs and biologics after the products are approved for marketing.

[Slide]

First I will describe how we conduct safety surveillance on products after the products have been approved, that is, the components of postmarketing surveillance, and I am going to focus

in AERS. I will describe the strengths and the weaknesses of AERS. I will talk about how safety evaluators use AERS, how we evaluate case reports and develop case series from the case reports sent to us.

Next, I will present four case studies with three drugs showing the use of AERS, and in this part I will be using information that is available in the public domain. Terfenadine is an example of a drug for which a safety signal was slow to emerge from AERS data. I will talk about that and tell you why this signal was hard to identify with postmarketing data. The second case study is salmeterol in asthma exacerbation. This safety issue is difficult to evaluate using AERS and I will explain why. The third drug is valdecoxib. I will present two case studies with valdecoxib. We quickly identified one safety signal for valdecoxib with AERS data, and I will describe a second safety issue with valdecoxib that was difficult to evaluate with AERS.

[Slide]

We cannot know everything that we need to know about the safety of drugs when the drugs are approved for marketing. A relatively small

population is studied that may not be truly reflective of the patients who will be using the drug after approval. Often fairly narrow indications are studied, and the indications may expand when the drug is used in practice. Finally, often the studies are of short duration and may not reflect chronic use of the product.

[Slide]

Our surveillance system is comprised of several components. First, the center of our surveillance system and the component we rely on the most right now to generate safety signals is the Adverse Event Reporting System, or AERS. Today I am presenting information about the central component of FDA's postmarketing surveillance system. But AERS is not the only database we use. Some of my colleagues in the Office of Drug Safety will present information to you about the other components of our surveillance system that you see

on this slide.

[Slide]

As I said, the piece of the puzzle I am going to talk about today is AIRS, the centerpiece of our surveillance system. AERS is a passive voluntary, spontaneous reporting system, and by voluntary, we mean it is voluntary for healthcare practitioners and for patients to report adverse events either to the pharmaceutical companies or directly to us. It is not voluntary for the pharmaceutical companies who are required to report these events that were voluntarily reported to them. The pharmaceutical companies are required to share these voluntary reports with the FDA.

AERS is a computerized database which began in another form in 1969. The database contains about three million reports for human drug and therapeutic biologic products. AERS does not contain reports on vaccines. That data is contained in a separate database.

[Slide]

So, what are the strengths of AERS? It

includes all U.S. marketed products. It gathers data relatively inexpensively to allow detection of events not seen in clinical trials. It is especially useful for detecting rare events with short latency, that is, rare events that occur fairly quickly after exposure to the drug. With AERS data we can develop case series to identify trends; to identify indications for products that may be problematic; and to identify populations that may be at increased risk.

[Slide]

AERS does have important limitations as a tool for detecting and exploring drug safety problems. We sometimes receive more than one report for the same event so when we are evaluating cases we have for a given event with a drug, we almost always must go through our cases looking for duplicate reports. Although we receive duplicate reports for some events, a bigger problem with AERS is underreporting. We don't know the true extent of underreporting but we know that it is vast. We know we receive only a small proportion of events

that occur with drugs including very serious, even fatal events. The quality of the reports is variable. The reports often lack data critical to the evaluation of the event. Remember that reporting is voluntary and busy practitioners are submitting the bulk of our reports.

There are reporting biases. Dr. Mary Willy, an epidemiologist in our office, will address some of these biases in her presentation. The actual number of events and the number of exposed patients in the population are not known so we cannot calculate incidence rates from AERS data. Also, it may be difficult to attribute events with a high background rate. For example, myocardial infraction may be difficult to attribute because it is such a common event.

[Slide]

Will the cases submitted to AERS reflected what happens when a drug product is introduced into clinical practice? The patient population may be more inclusive than the patient population that the product was tested on in clinical trials. The

patients may include more young and old patients, more women, more minorities, and patients with more complicated medical pictures. The product may be used for wider indications than studied. Use may be chronic. Because the patients who use the product may have complicated medical histories, the AERS cases tend to be complicated. In many cases the complicated nature of our postmarketing marketing cases pose significant challenges to us in attributing the adverse events from exposure to a particular drug product.

[Slide]

So, that is a brief overview of AERS. Now I want to turn to those of us who perform the front-line surveillance using AERS, and that is the safety evaluators. What do the safety evaluators do, and how are the case reports that I have been talking about dealt with?

The case reports are entered into AERS and each report of an unexpected event, that is, an event not in the labeling with a serious outcome, for example death or hospitalization, is triaged to

a safety evaluator. Selected other reports are sent to our in-boxes as well, including some reports that are sent as a result of increased surveillance on a known safety issue. The reports are sent to a virtual in-box on our computers. So far this year about 800 such reports each day have been forwarded to the safety evaluators. We can monitor our AERS in-box on a daily basis, reading the serious unexpected reports that have been submitted for the drug products.

In addition to the AERS in-box, we receive periodic safety reports on each drug product. These reports summarize all the case reports submitted to a manufacturer for a product. Our main mission is to identify and monitor safety signals for the drugs. We work closely with the epidemiologists in our office and with the medical officers in the Office of New Drugs.

[Slide]

I said we are looking for safety signals. What are the safety signals that we are looking for? First, we are looking for information we did

not previously know about a drug, that is, for new unlabeled serious adverse events, including drug interactions and drug-food interactions. We also look for information to add to our knowledge about adverse events that we already know about, perhaps increased severity of an expected adverse event, or more specificity, or information about some population subgroups that may be at increased risk.

With all the work that I have just described, you may wonder do we have a way to prioritize our work. How do we understand the forest if we are examining the forest tree by tree? We are developing a data-mining tool that uses Bayesian techniques to help us, and Dr. Carol Holquist will give you more information later today on this.

[Slide]

Data mining control through AERS using mathematical tools to identify higher than expected frequency of product-event combinations. Data mining is a tool for hepatocellular generation or support for further work on a hepatocellular. It

will supplement our manual in-box review and it won't replace expert clinical case review.

[Slide]

Now, how do we approach a safety issue? First, a safety issue is raised and this can occur in many ways. I already described our AERS in-box review, revising data mining results or periodic safety reports. The safety issue can be raised other ways as well--review of the results of a clinical study, perhaps in the medical literature; medical officer review of safety data in a new drug application; a question may be raised with us by a clinician, consumer or a member of Congress. So, the initial safety issue can come from many sources.

[Slide]

After the question is raised the safety evaluator can screen AERS for cases. We can analyze data mining information. We can look in the medical literature, and we can evaluate the cases we find. We can follow-up with reporters if we need more information on the cases.

[Slide]

When we search AERS for case reports addressing the safety question that has been raised

we use the medical terminology MedDRA, the Medical Dictionary for Regulatory Activities. The events reported to AERS are coded with MedDRA, and MedDRA consists of a hierarchy of terms. We can search AERS using this hierarchy of terms which consists of preferred terms, higher level terms, high level group terms or system organ class terms. Each successive grouping term broadens the search because the terms are less specific as one moves up the classification scheme. For example, a higher level term usually includes several preferred terms. We can search for additional cases in the medical literature. We can search for cases in other countries using the WHO database, and we can consult with our colleagues in other countries. We can use case definitions to help us with our search and to winnow down cases obtained in a search.

[House audio system malfunction]

[Slide]

When we gather cases from AERS we usually end up with a range of cases, including cases with very complete description and documentation and cases with very little information. Remember that most of these cases are sent to us by very busy practitioners who have carved a little time from a

busy day to communicate with us about an adverse reaction. We certainly appreciate that they have taken the time to do that. We especially appreciate it when the case reports contain complete information, including complete description of the event, what product the reporter thinks caused the event, other products the patient was taking, patient characteristics, medical conditions, risk factors, documentation of the diagnosis, information on whether the event abated when the product was discontinued and whether the event recurred if the product was reintroduced, that is, dechallenge and rechallenge information.

[Slide]

We evaluate the case reports we receive to try to determine if there are other factors that

explain the even or if, indeed, the drug product is the cause. For any individual case report it is rarely possible to know whether the event was caused by a drug product, but we do assess the strength of the evidence for causality. We look for the temporal relationship between the use of the drug and the event for drug-disease interaction. We look at other drugs the patient was taking that may have contributed to the event, and we look at supportive clinical and laboratory findings.

[Slide]

In addition to dechallenge and rechallenge, we look at biologic plausibility of the event, and we look for known drug class effects. We look to see if the event was observed in premarketing testing, and we look to see if there are other explanations for the event.

[Slide]

When we have a series of cases reporting the same event with a drug, we look for trends and for the pattern of events. We look at the cases

together for the ages of the patients, the sex, time to onset, whether there is a dose response relationship. We look for risk factors that may make patients more likely to experience an event, and we evaluate the strength of the evidence for causality and assess the clinical significance of the event.

[Slide]

So, we face a number of challenges in evaluating AERS data. First, it is difficult to attribute causality for cases with high background rates, for example myocardial infarction; events with long latency, for example cancer, may not be attributed to drug exposure. Cases are often confounded by other possible etiologies and often the reports do not include complete diagnostic information.

[Slide]

When we are exploring a safety concern we often develop case series from the cases we find from various sources, most importantly the cases in AERS.

[Slide]

We have developed case definitions for some events. The case definitions combine the

reality of how events are reported to AERS with the clinical information about the event. The challenges we face in evaluating the case report affect the application of a case definition. Do we require supporting diagnostic information in each case or, if the case was reported by a clinician, do we accept the diagnosis reported by the clinician even if the reporter did not include supporting information?

[Slide]

This is an example of a case definition that we use, and this is for aplastic anemia. We do accept the case as a case of aplastic anemia if a clinician reports it as such even without complete details about biopsy findings or blood work results. If the case was not reported by a clinician with a diagnosis of aplastic anemia we require a bone marrow biopsy, as well as additional laboratory findings.

[Slide]

The case definition also includes information on how to handle cases where bone marrow biopsy has not been done.

[Slide]

Finally, it includes suggestions on how to

locate the cases in AERS. The definitions searching AERS using a higher level term--marrow depression and hypoplastic anemia, as well as terms that describe the blood laboratory findings. I described before that we often need to broaden our AERS search from a particular preferred term, and this definition shows this. There is a term, aplastic anemia, in MedDRA, however, with this definition we suggest moving up one term above aplastic anemia, and the higher level term, marrow depression and hypoplastic anemia, includes a number of related preferred terms including aplastic anemia, bone marrow depression, hypoplastic anemia, pancytopenia, and so on.

[Slide]

Now let me pause here and just briefly

summarize the safety evaluator work. Cases are submitted to us from patients and from healthcare practitioners. Some come to us directly and some such reports come through the pharmaceutical companies. The safety evaluators routinely review our AERS in-box for safety issues. We search AERS and the medical literature for additional cases that may reveal a safety issue we have noted from our in-box review. We often search AERS using a broad search and then narrow the cases, often using a case definition. We look for good supportive cases. Remember that we receive a broad range and quality in our case reports. We contact reports for additional information in some cases. We develop case series from the cases and we may consult with the epidemiologists in our office to put our findings into context to help quantify risk. Dr. May Willy will give additional information on the role of the epidemiologist in our office.

[Slide]

Now I want to turn away from what we, as

safety evaluators, do on a daily basis and step back to look at how well a spontaneous reporting system like AERS has handled safety issues, and I will look at four safety issues for three different drugs and examine how well or poorly the AERS safety data supported the discovery and then exploring these safety issues.

[Slide]

Terfenadine was approved in 1985 and was withdrawn from the market in 1998 because of its role in causing cardiac arrhythmias. We usually receive the most reports for a drug within the first few years of marketing, and this pattern will be described by Dr. Willy in her presentation. However, the reports we received of QT prolongation and Torsades with terfenadine did not follow this pattern, with the most reports received in 1992, seven years after the drug's approval.

[Slide]

How can we understand what happened here? What were the problems with dealing with this safety issue? We think the difficulty was in

diagnosis and attribution of the event. The first step in reporting a safety issue is diagnosis and this was not easily accomplished in the case of terfenadine-induced arrhythmias because patients taking terfenadine generally were not being monitored with EKGs. Even when a cardiac event occurred, that is, when a rhythm abnormality became clinically apparent, we think the diagnosis of the event was still difficult. In fact, we have reports of syncope for terfenadine that probably were cases of Torsades. So, even when the event was clinically apparent, this still may not have been diagnosed as to the actual cardiac problem that was occurring.

Secondly, once diagnosed, the event must be attributed to exposure to a drug, and finally the event must be reported to us. We think in this case the diagnosis of the event and attribution to terfenadine were problems. Because of this, this signal was slow to emerge. Salmeterol was approved in 1994. Early in the marketing of salmeterol we received postmarketing reports of asthma

exacerbation with the use of the product. So, what do we make of those reports? Of course, we expect some asthma patients to experience exacerbation of asthma even while on therapy. These cases reported for almeterol are confounded by the indication for the use of the product. But are we seeing something in these reports that is out of line or what we expect not only with the drug but with the disease? The real question is whether there is a differential rate of exacerbation of asthma that is attributable to the drug and AERS cannot answer this question.

[Slide]

Because AERS could not answer the safety question regarding whether salmeterol causes more asthma exacerbations, a large safety study was undertaken to answer the question. This was the salmeterol multi-center research trial, or SMART. The trial was stopped about two years ago after partially answering this question. I am not going to go into the findings of SMART today but you can go to the URL at the bottom of the slide for more

information on the study.

[Slide]

Now I want to move to the last drug. Valdecoxib is a COX-2 selective NSAID. It was approved in 2001 and was first marketed in early 2002. Early on in marketing we received a number of cases of Steven-Johnson syndrome and toxic epidermal necrolysis temporally related to the use of the product. This information was quickly incorporated into the labeling.

[Slide]

What were the characteristics of this issue that led to exploration via AERS? This is exactly the type of issue that AERS handles very well. First, the skin events are rare, easily diagnosed events, with short latency from the time the drug was started. Secondly, the practitioners attributed the reactions to valdecoxib and, thirdly, the practitioners reported the events to the manufacturer and to us. Finally, we were able to compare reporting rates for this event for valdecoxib and other drugs in the class at a

similar point in marketing. With this comparison, we were convinced that these events are a real problem with the drug. We use reporting rates very cautiously but this is one case where we could compare the rates. Dr. Willy will talk some more about reporting rates in her presentation. So, we were able to recognize this easily and incorporate this information about serious skin reactions in the labeling.

[Slide]

A second safety issue with valdecoxib is an example of an issue that is not a good issue for AERS. Valdecoxib is associated with thromboembolic events but these are fairly common events. Because of the high background rate, it is difficult to explore this using AERS data. For both salmeterol asthma case study in which the cases were confounded by indication, and the thromboembolic valdecoxib case study where there is a high background rate of the event in the general population, AERS is of limited use.

[Slide]

To summarize, AERS reflects use of products in clinical practice so these data are different from clinical trial data. AERS is a

passive surveillance system that relies on diagnosis, attribution and reporting by healthcare providers. The cases reported to us often lack complete diagnostic information.

[Slide]

AERS is especially useful for rare, easily diagnosed events with short latency. AERS is less useful for attribution of events with high background rates, confounded by indication, confounded by other etiologies, or with long latency following drug exposure. And, AERS cannot establish frequency of events in the population.

[Slide]

Finally, I would like to acknowledge my fellow safety evaluators in the Division of Drug Risk Evaluation. There are 27 of us who do the front-line work that I described in this presentation, and their names are on this slide.

DR. GROSS: Thank you very much. That was

very instructive. Dr. May Willy, the epidemiology team leader from the Office of Drug Safety, will talk about epidemiologic analysis of spontaneous adverse reports.

Epidemiologic Analysis of Spontaneous Adverse Reports

DR. WILLY: Good morning.

[Slide]

I am an epidemiologist in the Office of Drug Safety, and I am going to talk to you today about the epidemiological considerations in the analysis of spontaneous adverse event data that was just described by Dr. Weaver.

[Slide]

I plan to describe the limitations in using passive surveillance data for an epidemiological analysis, and the methods used to identify and analyze rare adverse drug events.

[Slide]

The limitations that must be considered when analyzing passive surveillance data include the following, and will be described further in my

next slide: First, the underreporting of adverse drug events is a well-known problem. Rarely no more than an estimated 10 percent of serious events, and between 2-4 percent of non-serious events are reported. Second, reporting rates vary over the length of marketed drug, something called the Weber effect. Third, reporting of adverse drug events has changed over time. Fourth, analyses are often limited because adverse event definitions are not consistent among reporters. Lastly, the best epidemiological analysis of passive surveillance data cannot yield an incidence rate.

[Slide]

The underreporting of adverse drug events has been studied in a number of settings. A Canadian study of reporting of toxic epidermal necrolysis, or TEN, to the Canadian regulatory agency was described in 2004. In this study the investigators decided to study the reporting of TEN because it is almost always a drug-related event. The investigators found that of 250 cases of TEN that were admitted to burn centers, only 25 were

reported to the Health Canada surveillance program.

[Slide]

A second study of underreporting was published in 1990. This study showed that the reporting of adverse events increased 17-fold following 2 years of physician exposure to a pilot program that encouraged reporting.

[Slide]

There are a number of reasons for underreporting, including the nature of the adverse event; the type of drug product and its indication; the length of time on the market; the extent and quality of the manufacturer's surveillance system; whether the product is prescription or over-the-counter; and the perception of liability by the potential reporter.

[Slide]

Reporting of adverse drug events for a newly marketed drug can be influenced by the Weber effect. The Weber effect was described some time ago by Dr. Weber when studying the reporting of adverse reactions for seven nonsteroidal

anti-inflammatory drugs. Dr. Weber observed that the peak of adverse event reporting occurred at the end of the second year of marketing despite the lack of declining usage. The Weber effect has been observed to occur with other drugs and should be considered when comparing reporting rates of several drugs since the length of marketing may vary.

[Slide]

Another consideration for analyzing adverse events is that the reporting of adverse drug events has increased over time as there has been more emphasis on training healthcare processionals to report to the FDA. Comparing newer drugs to older drugs becomes difficult because of the differences in reporting to the FDA.

[Slide]

This graph summarizes the number of AERS reports received by FDA since 1990. The counts have increased from roughly 50,000 in 1990 to over 400,000 in 2004.

[Slide]

The analysis of potential signals can be difficult when an increase in reporting is observed but you know that there has been media attention.

The news about a signal may not only affect the drug in the media report but also the drug in the drug class. High profile safety concerns may lead to increased reporting from lawyers and patients. A high profile news report about just one unusual case of some adverse event may stimulate others to report as well.

[Slide]

The definitions used for an epidemiologic analysis can be difficult since, as Dr. Weaver described, reporters can use different terms in reporting the same syndrome. Cases may be misclassified by the analyst if reporters are using different criteria for diagnosing a disease. In addition, it can be difficult to determine how to define a case when there is missing information.

[Slide]

Now let me review some of the epidemiological methods used to analyze passive

surveillance data. Many of the methods used by epidemiologists in the Office of Drug Safety are similar to those used by safety evaluators, and often a safety analysis will involve a team of epidemiologists and safety evaluators.

[Slide]

Two common ways that a drug safety signal using passive surveillance data is evaluated are either by descriptive analysis of a case series or an analysis that involves calculating reporting rate. Many times both methods are used.

The reporting rate calculators use the number of reported cases for the numerator and some measure of drug exposure in the population for the denominator. Calculating reporting rates can be difficult because the available numerators and denominators are often limited. As described previously, determining the numerator can be difficult because the definition used to describe a case is often not consistent among reporters. Additional, search terms used to identify cases may miss reports that do not include the chosen terms.

The denominator can be difficult to determine because data on patient exposure are rarely available. Exposure can be continuous, short-term or intermittent and all can be difficult to measure. Since exposure data are not available, we use prescription or sales data as proxies for exposure.

[Slide]

Reporting rates can be used in different ways when analyzing passive surveillance data. They may be compared to background rates, to drugs within the same class, or compared to drugs for similar indications. Each of these comparisons is limited since background rates may be difficult to obtain, drugs in the same class may have different years of marketing, and drugs for similar indications may serve different populations.

[Slide]

There is no standard methodology used in analyzing passive surveillance data. Some of the factors considered when interpreting reporting rates that may help define a signal include finding

a notable difference between the reporting rate for the drug of interest and the background rate, or between the drug of interest and reporting rates for similar drugs. Consistency in reporting can also help. For example, observing that there is an increase in reporting for less severe forms of the event of interest can be helpful. Premarketing clinical trial data that suggests a possible risk may also help define the drug signal. And, literature reports may provide support of a possible signal.

[Slide]

Her is one example of how reporting rates were used in an analysis. This comes from a 2001 publication that compared three glitazones and their reporting rates for acute liver failure. Troglitazone was found to have a reporting rate that was notably higher than the other glitazone reporting rates.

[Slide]

Comparing the reporting rates for fatal rhabdomyolysis for the statins was reported in

2002. This table sums the total number of fatal cases, in the first row; the number of prescriptions listed in thousands, in the second row; and the reporting rate, in the third row. The calculated reporting rates show some small differences among most of the statins but the final column provides the reporting rate for cerivastatin, which is a magnitude higher than the other drugs.

[Slide]

Background rates used for comparisons might come from literature summaries or national data, such as hospital discharge data, mortality data and other types of relevant data. Often background rates are not available though, particularly for special subgroups like childhood syndromes. If there is a rate available though, the comparison can help determine if there is a drug safety signal.

[Slide]

here is one example of how background rates were used in an analysis. This comes from

the same 2001 publication described earlier and it compares the background rate for acute liver failure to troglitazone, the background rate being 1/million person years.

[Slide]

Although the reporting rate can define a drug safety signal, they are not incidence rates. The suffer from incomplete numerators and crude estimates of denominators. Reporting rates can be helpful but almost always must be considered a crude estimate.

[Slide]

In conclusion the analysis of passive surveillance data can identify potential drug signals. Epidemiology tools are helpful to quantitate the significant but limitations with the system can make final conclusions difficult, especially since there is no standard metric for calling a finding real.

[Slide]

In my final slide I would like to acknowledge the epidemiologists that work in the

Division of Drug Risk Evaluation. Thank you.

DR. GROSS: Thank you, Dr. Willy. The next speaker is Carol Holquist, Director, Division of Medication Errors and Technical Support for the Office of Drug Safety. She will talk on using FDA's AERS in postmarketing surveillance for medication errors.

Using FDA's AERS in Postmarketing Surveillance for Medication Errors

DR. HOLQUIST: Good morning.

[Slide]

Today I will provide an overview of how AERS is used in postmarketing surveillance of medication errors.

[Slide]

However, before I begin discussing the surveillance of medication errors, I think it is important to provide you with a definition of a medication error. Secondly, I will briefly describe the detection and evaluation of medication error safety signals and, finally, I will provide three examples of postmarketing medication error

cases that help illustrate the types of issues the Division of Medication Errors and Technical Support encounters.

[Slide]

How do we define an error? An error can be defined as any preventable event that may cause or lead to inappropriate medication use or patient harm while the medication is in the control of the healthcare professional, patient or consumer. The key to this definition is the term "preventable event." These events are not intrinsic to the drug product and are generally multifactorial in nature.

[Slide]

Also included in this definition is where these preventable events can occur. They can happen at any point in the medication use system. FDA adopted this definition in 1999 from the National Coordinating Council for Medication Error Reporting and Prevention, better known as NCCMERP. NCCMERP was founded in 1995 and is comprised of national healthcare professional organizations, regulatory bodies and industry. If you wish to

find out further information about NCCMERP, I refer you to their web page at www.nccmerp.org.

[Slide]

When assessing errors, FDA uses the NCCMERP medication error index. The index categorizes an error according to the severity of the outcome and considers factors such as whether the error reached the patient, and if the patient was harmed and, lastly, to what degree. This categorization helps us track errors in a consistent and systematic manner.

All errors can be groups into two very broad categories, potential and actual. Potential errors represent category A, which is the blue area on this pie chart. Potential errors are what we refer to as product complaints. There is no patient involvement at this point. Complaints concerning the sound alike potential between two name pairs, as well as complaints involving the similar appearance of product labeling, would be typically classified in this category.

Categories B-I represent errors that have

actually occurred. However, the level of patient involvement varies. For instance, errors classified in category B, this first orange piece of the pie, are typically referred to as near misses. These are the types of errors that are caught prior to reaching the patient. For example, a technician fills a prescription with the wrong strength or the wrong drug but the pharmacist performs a double check and realizes the error before dispensing the medication to the patient.

Categories C-I represent errors in which the patient received the wrong medication. However, the variable here is the level of harm and the patient outcome depending on the medication involved in the error and whether or not the medication was actually administered to the patient. So, within this categorization you can go from no harm, which is the orange area, to more serious harm, which is the yellow area, to the worst-case scenario, which is death. It is important to note that at FDA we treat potential errors as seriously as actual errors. Potential

errors are early warning signals to tomorrow's actual events.

[Slide]

In March of 1999 medication error surveillance became a full-time function of the Office of Drug Safety. It was at this time that all medication error reports began to be entered into the AERS database regardless of patient harm. Error reports can be submitted from various sources such as the manufacturer, consumers and healthcare providers including pharmacists, physicians and nurses. Consumers and healthcare providers can voluntarily report to either the FDA MedWatch or the USP/ISMP Medication Error Reporting Program. Manufacturers report directly to FDA. Collectively, we receive approximately 300 reports per month from these reporting programs.

[Slide]

The Division of Medication Errors and Technical Support and safety evaluators are the individuals responsible for conducting postmarketing surveillance of these errors. The

staff is comprised of clinical pharmacists, nurses and one physician. DMETS safety evaluators have a dual mission. They are not only responsible for monitoring postmarketing medication error safety signals but provide premarketing safety reviews of proprietary names, product labels and labeling and packaging in an effort to prevent medication errors.

DMETS collaborates with the Office of New Drugs medical officers and chemists because these are the individuals responsible for review of the labels and labeling at FDA. We also collaborate with the Office of Generic Drugs and the Office of compliance on postmarketing medication error safety signals. Our safety evaluators do not have virtual errors in-boxes like our colleagues in DDRE so on a monthly basis all reports coded as a medication error are reviewed manually.

[Slide]

DMEDS typically focuses on errors relating to medication errors involving product name confusion; error prone labels; labeling, packaging

and devices because these are the areas in which FDA has oversight.

[Slide]

A medication error safety signal can be identified as a result of a voluntary report, a question from CDER OND, or other centers such as biologics and devices. Signals can also emerge from an article published in the medical literature or an outside inquiry from a practitioner, consumer or even a patient safety organization. Following this initial inquiry, AERS is searched for all similar cases.

[Slide]

Until March of this year only one MedDRA preferred term was available to search medication error reports within AERS. Thus, DMETS safety evaluators typically began any investigation with a narrow focus and then broadened the search depending on the initial findings. Oftentimes reports are not coded correctly as a medication error and, thus, depending on the nature of the error we typically would not uncover these cases

until we broadened the search to overdose, accidental exposure, product problem or use another PT term related to the adverse event associated with the error. We really recognized the need to expand our terminology not only to facilitate the retrieval of cases, but to help automate some of the categorization that safety evaluators were doing manually. Thus, new medication error terminology was introduced in the 8.0 version of MedDRA which was released in March of 2005.

Additionally, we may see an increase in the number of medication error reports if the SADR rule is finalized as proposed. The proposed rule will require manufacturers to submit all medication error reports as expedited. Currently, there is no reporting requirement for medication errors other than the existing postmarketing reporting of serious adverse events.

A good quality medication error report not only contains the items highlighted in Dr. Weaver's presentation but would also contain information such as the specific brand or generic name involved

in the error, the manufacturer, dosage form and concentration. Most errors are product specific and, thus, this information is critical in the assessment of any error.

We also need to have a clear and complete understanding of all circumstances and events that led up to the error, including the contributing factors, work environment, location of the error and, if possible, the personnel involved in the error. All this information helps us categorize and piece together the system's failures that occurred and caused the error. Once we have some semblance of the total picture, we can begin to see where FDA might have a role in preventing further events.

[Slide]

In order to ensure all cases of a specific error identified by safety evaluators, we not only have to search AERS but also the medical literature, including safety alerts and bulletins published by patient safety organizations such as ISMP and USP. Once the reports are duplicated, we

attempt to follow-up with the reporter in order to obtain a clearer picture as to how the error occurred.

Following receipt of any additional information, the safety evaluators can then begin to categorize the events by type and cause using the NCCMERP taxonomy. We find the taxonomy to be a useful tool in analyzing medical patient error reports as it provides a standard language that describes type and causality of medication errors. Following this drill-down of reports we have a better understanding of the risks and/or contributing factors that led to the error.

[Slide]

One of the biggest challenges we have in assessing causality is getting a complete understanding of how and why the error occurred. Many individuals report anonymously and, thus, no follow-up can be made. We then rely on our postmarketing lessons learned from other products that have similar problems. When an adverse event is reported with a medication error too must

evaluate the strength of the evidence for causal relationship to the error and adverse event. Other confounding factors may be present such as the disease state which can diminish the probability that the adverse event was directly related to the error. We allopurinol also evaluate the label and labeling or packing of a device described in the report to determine if, one, what is being reported factual and, two, if additional error prone features identified by the reporter are present.

[Slide]

As I stated, the biggest challenge in assessing the root cause of any error is the absence of a complete description of how the error occurred and the actual product involved in the error. We also have a limited number of reported cases, some of which can be confounded by other etiologies. Thus, it can be difficult to drive regulatory change with such small numbers.

It can also be challenging to acquire the labels and labeling of the products used in error, especially with respect to grandfather product

distributors and those subject to OTC monograph products.

[Slide]

Having briefly described the analysis of medication errors, I would now like to provide you with three examples reflecting the types of errors we encounter with poor label design, poor packaging design, and product name confusion. AERS was used for detecting the signal from the first two cases. However, in the last case AERS was not particularly helpful in answering the specific safety concern with concomitant administration of the same drug product, from the same manufacturer having two different proprietary names.

[Slide]

In this first example I wish to illustrate the poor design of the label and labeling which contributed to confusion and error. Shortly after Temodar's approval in 1999 we began receiving complaints with respect to the placement of the net quantity in relation to the product strength. Additionally, we also received complaints that the

100 mg strength was difficult to read due to insufficient color contrast. The sponsor uses a black print on a blue-turquoise background.

Within one year of receipt of these product complaints, or what we refer to as potential errors, we had a small number of reported fatalities. When we initially reviewed the labels we noted that, in fact, the net quantity did appear in juxtaposed position to the product strength and in a more prominent fashion where you eye is drawn to this number rather than the product strength. To further complicate matters, two of the product strengths overlap with available commercial package sizes of 5 and 20.

One of the fatalities involved misinterpretation of this net quantity 20 capsules as the product strength. The patient was ordered 20 mg daily for 5 days. However, the pharmacy dispensed a higher strength rather than the intended 20 mg.

[Slide]

The second examples is a good illustration

of an error prone packaging design. We received four cases involving inadvertent administration of this topical corticosteroid in the eyes and ears. Upon review of the bottle and label presentation, it was very clear how these errors were occurring. Not only does the shape and size of the bottle resemble those used by otic, ophthalmic and nasal products, but the tip used on the bottle is the type of design used in the packaging of these products as well. Additionally, the route of administration "for topical use only" was not prominently displayed. The statement is hidden in the crowded text that appears with the same prominence at the bottom of the label.

[Slide]

This final case was one in which AERS gave us a few but very important cases showing the concomitant administration of Zyban and Wellbutrin. Bupropion, the active ingredient in both Zyban and Wellbutrin, is associated with a dose-related risk of seizures. Because bupropion is available under two different proprietary names with very different

indications of use, smoking cessation and depression, we were concerned that patients might theoretically visit different healthcare providers and receive prescriptions for each product and administer these medications concomitantly, not knowing that Zyban and Wellbutrin contain the same active ingredient.

We did, in fact, receive a small number of reports of seizures, which supported our hypothesis. However, the limited number of cases didn't allow us to demonstrate the extent of the problem. Even using claims data, it was difficult to discern concomitant administration versus the cases in which a patient was switched from one agent to the other.

[Slide]

In summary, AERS can be useful in detecting safety signals concerning medication errors. It can also be reflective of the use of products in clinical practice. However, it is not a good indicator of the magnitude of the product problem, primarily due to the limited number of

cases reported to the agency.

[Slide]

With respect to medication errors, we only see the tip of the iceberg.

[Slide]

In closing, I would like to acknowledge the 16 safety evaluators currently employed in DMETS. Thank you.

DR. GROSS: Thank you very much. The next speaker is Dr. Judy Staffa, epidemiology team leader, the Division of Surveillance, Research and Communication Support, from the Office of Drug Safety. She will talk about available types of national drug use data.

Available Types of National Drug Use Data

DR. STAFFA: Good morning.

[Slide]

My name is Judy Staffa. I am an epidemiology team leader in DSRCS, Division of Surveillance, Research and Communication Support, in FDA's Office of Drug Safety, and we are the home for many of the external contracts to purchase drug

use data from outside the agency.

I am going to provide you a very brief overview of the available types of drug use data, and I have underscored "national" in the title because, although there are many different avenues and sources from which one can learn about the use of products, FDA is in a rather unique position where we really are mandated to look at national patterns and to focus on sources where we can obtain those kinds of estimates.

[Slide]

Just to give you an idea of where I am going here, I am going to very briefly reiterate some of what the other speakers have said about the applications for drug use data in the area of drug safety. I am also going to walk through some of the typical questions we want to answer with relationship to drug use when a drug safety question emerges. I am going to walk through some of the challenges we face in trying to obtain these kinds of data. Then, I am going to give you a lightning tour of the available types of data, but

doing it by question, trying to look at what types of data might be available to answer different questions by setting of care, which I hope will become obvious to you why I need to do that. Then, in summarizing, I hope to convey some of the future challenges we are dealing with in trying to expand the amount of data and the type of data we access.

[Slide]

Just briefly, Drs. Weaver and Willy described some of the applications for what we do with drug use data. One of the classic ways we use the data is to provide denominators for AERS cases. So, this is to provide a context and that is done through calculating reporting rates.

In addition to that, we also do a lot of descriptive work with these data. We try to describe patterns of use. We want to understand who is prescribing drugs and to whom, and what are the characteristics of the patients receiving those drug with regard to age and gender, and also what types of illnesses they may have. We also use these data to gain insight into how long patients

remain on drug treatment, and whether or not they are exposed to concomitant therapies that they may or may not be advised to take together.

More recently, we have begun to use these data to look into the impact of risk management strategies that involve restrict the use of drug products. We are beginning to learn to do those kinds of analyses. And, as Carol alluded to, we try to use the data to assess the impact of potential medication errors when possible.

[Slide]

So, what are some of the questions that come up that we really want to use drug use data to answer? Well, "the $64,000 question" that always comes up is how many patients in the United States are taking a product? When a drug safety signal emerges that is really one of the first questions that emerges. Even though it is a very simple question, I hope I will convince you by the end of my talk that it is not quite as easy as it sounds.

[Slide]

Some other common questions really stem

from some of the applications I talked about earlier. We want to understand what do these patients look like. How old are they? What gender are they? Are they pregnant? We want to know how long they stay therapy; what other therapies they take. And, we want to know for what indications are drugs prescribed and by whom.

[Slide]

Some of the challenges in answering these questions are pretty formidable. The first and the most important is what I call the fragmentation of the U.S. healthcare system. Unlike some other countries, as you know, we don't have a national healthcare system. We can't follow patients across different care settings and identify all the drugs to which they are exposed and all the disease conditions they experience. So, what happens in our healthcare system is that we wind up with what I have termed "pockets" of use. We can look at drugs being dispensed or used in different pockets. Those pockets are really defined predominantly--what I will focus on today--on

settings of care, which I will elaborate a little more on in the next slide. They can also be defined by pairs. For example, different insurance companies are managed care organizations who are paying at least some piece of what patients are paying for drug therapy. Or, a pocket can be defined by the buyer. Typically, hospitals will be in together and purchase drugs through a buyer.

So, what happens if you are trying to get a national picture of drug use is that you end up looking at drugs in each of these different pockets, trying to understand what that pocket looks like at a national level; take the data you have in a pocket, try to weight it somehow to project it to a national level, and then sum data across pockets. So, as you start to think about that you realize how many errors, how many leaps of faith and how many difficulties there are in each of those steps to try to result in what we want, which is typically one number of understanding the extent of use.

[Slide]

Just to elaborate a little bit on settings, because that is the kind of pocket I am really going to focus on today, depending on how a

drug is used and in what setting it is typically used will define how well we can characterize its use. So, to start out with the most common way patients receive drugs, it is really the outpatient pharmacy setting. Most patients receive prescription drugs through a prescription they pick up at their pharmacy, or they receive it through mail-order. Patients can also be dispensed or administered drugs in a physician's office, or they can receive a product in the setting of a clinic, whether it is free-standing or a clinic associated with a hospital. Another setting would be the inpatient setting where a patient goes into a hospital for a stay and received different drug therapies while there. Finally, patients can self-prescribe or purchase over-the-counter products which are regulated by FDA, and we certainly have an interest in their safety. Patients will typically buy these products, either

on their own or on the recommendation of their physician, and use them in whatever way they choose. So, these are the settings that really define the areas in which we can look at drug use.

[Slide]

In addition to the fragmentation or pocket type limitation to what we can do, we also have other challenges in the kinds of data that are available. There are very few sources out there that are really specifically collecting information on patient exposure to drugs. So, what we end up doing is using other data sources, collected for other purposes, and se can secondarily use them to look at patient drug exposure.

Two types of data that we often use for those purposes are administrative or billing data, which come into existence because drugs are a product and they are paid for and the tracking of that payment for the product is captured in large data systems. We can also use marketing data which, although not collected for epidemiologic purposes, are clearly collected for the purposes of

tracking and monitoring and understanding the distribution of drug products. So, the good news is we can actually use these data sources. They exist and we can benefit from them. However, we always have to bear in mind the limitations to these data sources for our purposes since that was really not the purpose for which they were originally collected.

I will mention here, as I will talk about a little bit more, there are some newer data sources coming on the landscape where they are actually linking data across different data streams even within settings of care. For example, in the outpatient setting we are seeing data sources now that are linking across different kinds of administrative billing systems and that provides more flexibility but, of course, as we begin to learn about these databases I am sure that they will also have their own set of limitations that we need to bear in mind.

[Slide]

So, let's start our tour. I have outlined

some of the major questions we want to answer. I want to talk about the different settings and how well can we actually estimate the answers to these questions by setting. We will start with our "$64,000 question" of how many patients take a particular drug.

[Slide]

I am going to start with the outpatient setting and I am going to start with a subset of the outpatient setting. I am going to look at pharmacies because this is typically where we have the best data. Traditionally, probably for 30 or 40 years, we have been able to get national estimates of the number of prescriptions dispensed for products that are dispensed out of pharmacies or out of mail-order and, to some degree, long-term care. These data are obtained from a sample of pharmacies. Since we can collect all that information from our samples and since we can readily pretty much figure out how many pharmacies are out there and how many mail-order houses are out there, we can weight those estimates and make

national projections and that gives us the total number of prescriptions dispensed.

However, traditionally there has not always been patient age and gender information in those prescription data. Sometimes, when it is there, it is often missing or incomplete. So, we are the prescription level but we still haven't gotten to the patient level in these kinds of data sources.

More recently, as I have talked about, there are newer data sources that are beginning to emerge in which we can also get these national estimates of dispensed prescriptions but, in addition to that, we can also link those to the patients associated with them and weight those estimates and actually get the answer to the "$64,000 question," which is how many patients nationally have been exposed to this drug.

This is accomplished because these data systems actually link across data streams so they are taking information from pharmacies but also adding some information from pharmacy benefit

managers and from insurers to try to supplement that information and gain information on the patients behind those prescriptions. By doing this, we add additional information on patients' age and gender so we can be a little bit more confident about the characteristics of those patients. These re relatively new data systems which we are actually pursuing at this time.

The limitations with both these systems are that they are still not covering all outpatient settings. So, for drugs that are not classically dispensed out of a retail pharmacy these data systems really don't capture that. Again, we always have to remember that just because a prescription is dispensed certainly doesn't mean that the patient has actually brought it home and taken it.

[Slide]

Now let's move into some of the other outpatient settings. Let's start with physician offices. There are many drugs that are administered to patients by physicians in their

offices, and we have several different sources to look and try to count persons. We can look at convenience samples of office visits that are derived from audits or surveys of several thousand physicians. What happens is that these physicians collect information on all patient encounters during a certain period of time. The data they collect are, again, weighted and projected to try to represent a national picture.

We can also look at data collected by the National Center of Health Statistics, and I have listed the National Ambulatory Medication Care Survey. It is a very similar design for collecting information from office-based physicians.

The problem with these is that, number one, we are still working at the visit level. If you remember, I described that these are collecting information on visits so we can look at the number of times a drug is mentioned in a visit, but we still can't get at the patient level because patients, of course, could make multiple visits to their physicians.

For the convenience samples, we find that when you get down to asking questions about specific drugs we can run into some very small

sample sizes and at that point the projections can become very unstable, and we get a little nervous about making national estimates based on very small numbers.

We often wonder too how generalizable these data sources are. Are the physicians that choose to participate in these kinds of data sources really representative of what all physicians out there are doing? For the NAMCS data, although they are probably stronger in the area of making national projections because of their methodology, we do find that for new drugs the data are often not timely enough to allow us to look at brand-new drugs as they emerge on the market.

The other source I talked about was clinics. Patients increasingly, in different disease areas, receive medications in a clinic. We have very little data available to us to be able to

count how many patients actually receive medications in this way. We are pursuing some medication claims data to try to perhaps tap into what is called J-codes, which are certain procedure codes that can be drug or product specific and for which physicians will bill the administration of a particular product but, again, it takes time to receive those J-codes. They are not comprehensive. But we are planning to explore whether or not they can be helpful to us.

So, for right now we really rely on sales data, which is how many products are sold into the back door of the clinic. That is not patient level data and at this point we really don't have any way of understanding what the clinics then do with those drugs and how many patients actually receive them. So, it is an area where we really don't have much data.

[Slide]

Moving into the inpatient setting, how well can we count patients that are receiving drugs in an inpatient setting? If you have been in this

field for a while, you will know that traditionally we have never had any ability to look into hospitals and to understand how patients use drugs. More recently, however, we have gained access to data where we can at least look into a sample of hospitals and look at billing data from those hospitals to understand how many discharges there were in which that drug was billed. Then, we can take those discharges, since we pretty much understand the universe of hospitals in the U.S., and try to project that to reflect what is happening in that particular pocket.

Along with the drug information available in these discharges, we can also get information about the age and gender of the patient, and also all of their discharge diagnoses and any procedures they had done during their stay.

However, there are a lot of limitations still to these data even though it is much better than what we had before. Clearly, we are at the discharge level. We are counting discharges. We are still not counting patients, and if patients

come in and out of a hospital more than once we are clearly double counting.

Because these are billing data, we have lists of discharges; we have lists of billed procedures; and we have lists of drugs billed; but there is no linkage in between. So, we really couldn't say what the drug was used for so that can be difficult. Again, as dispensed prescriptions don't necessarily represent patients taking the drug, there are a lot of nuances with billing data so even if a drug is billed during a hospital stay there are various reasons why the patient may not have received it. And, although these data systems can provide a lot of information on drugs billed when a patient is on a ward or an ICU, there are certain areas in a hospital where we still can't really get a good view of what happened and those areas are really most importantly the operating room, as well as radiology services, because the way billing is done, many of the drugs used are bundled still within the charges representing other types of care in that setting. So, it is still

hard to see those settings.

Finally, even though we feel fairly comfortable at this point in making some general projections from these acute short hospital stays, when we get into subpopulations, for example pediatrics, we still wonder--there hasn't really been a clear definition, for example, of what the universe is of pediatric inpatient care in this country is. So, we don't feel very comfortable in making those projections from these types of data, and we are still working on trying to figure that out.

[Slide]

Finally, how well can we count patients in the over-the-counter drug scenario? Well, traditionally we have had very little information for over-the-counter. Again, we are in the realm of sales. We can understand how many bottles of acetaminophen are sold through retail pharmacies or hospitals, but we have no idea how many patients are buying that acetaminophen and what they are doing with it. So, again, that is not at the

patient level.

More recently we have learned about different efforts to collect household survey data, to actually select households and interview consumers within the households to understand better how they use over-the-counter products. Some of these efforts are actually designed to try to make national projections from these data. So, we are doing some work to try to understand these data better but at this point we are really not strong what the strengths and limitations of these might be.

[Slide]

Moving on to the more longitudinal questions, we want to understand how long patients stay on drugs or how often they take multiple drug therapies. What is available to us in the different settings that we have defined?

[Slide]

In the outpatient world, what we have been using for the past several years are longitudinal patient level insurance claims data. This is very

useful. We can actually look at patients over time for as long as they are within their insurance plan, and we can look at prescription refills and actually link them together and create an episode of care so that we can understand how long patients are staying on therapy.

The limitations to this are that it is really not possible to make national estimates based on these data, even though it is very helpful to be able to look at them, because we discovered that it is very difficult to identify the universe of people in the United States who have prescription drug coverage. That can be a very difficult animal to define. Therefore, until we can define that we can't really project this our nationally.

We also wonder about the generalizability given that we are looking at patients with insurance and, clearly, depending on patients' ability to be able to pay for prescription drugs, they may stay on drugs for different lengths of time. Not all drugs are covered by insurance

systems. There are many restrictions on formularies or preferred methods of payment. We have to take that into account depending on the drug we are looking at. Finally, we run in to the same problem where, just because it was dispensed, it doesn't mean it was taken.

More recently, as I described, the databases we are pursuing we are hoping will increase our ability to look beyond the insurance setting with regard to our analyses of duration over time because, by going outside of insurance and linking across these data streams, we are hoping to be able to also look at patients who pay cash for their presentation and follow them over time to see how long they stay on products. Again, we are in pursuit of these data but, again, it is not clear yet what the limitations will be.

[Slide]

In the inpatient setting we really don't have too much ability to be able to look at duration although we can do a little bit. The billing data I described actually allows us to go

down to the day of stay billing detail. For example, if a patient is in the hospital for a 5-day stay we can look at each day of that stay and determine on which days a particular drug was billed. If we decide we would like to add those up, we could possibly imagine that that might be the duration of therapy for that patient on that product.

But, again, there are many limitations to this. It is not possible to get national estimates. The nuances of billing make us a little nervous, and we still can't link it to the discharge or to the diagnoses because those are all assigned at discharge so we can't really tell when a diagnosis occurred during the hospital stay and link it with any kind of drug treatment. So, our ability to look at duration in the inpatient setting is rather limited.

[Slide]

Finally, what kind of data do we have to be able to look at prescribing practices? What are physicians prescribing products for?

[Slide]

In the outpatient setting, basically the types of data I referred to before, we can access

these audits of looking at physician visits to outpatient or office-based visits either from convenience samples or from NAMCS. The limitations here are that, again, we can't actually get counts. We can get visits; we can get patients from these data, and the same limitations apply. We worry about the selection factors. We worry about sample size. The estimates can be rather unstable but we can get some picture, particularly for drugs that are used quite prevalently.

More recently you see more and more pockets of medical records. These are actually more prevalent, and Dr. Graham will be talking about tomorrow about our efforts to get these data in other countries, but in the U.S. we see pockets of these occurring as well. As these become developed, as more physicians begin using electronic medical records, as well as e-prescribing or hand-held tools for capturing

prescribing information electronically--as those become research-ready, we are hoping to look into whether they might be sources for identifying prescribing practices. Although again, as with the others, we will always question the representativeness of these types of data to all physicians.

[Slide]

We really have very little information on understanding who is prescribing drugs for what in the inpatient setting. Again, with the billing data we can look at the specialty of attending and consulting physicians and we can access all the diagnoses, however, we don't have the key linkages between which prescriber, which physician might have prescribed which drug, and we certainly can't link between the drug and the indication. So, we really can't do much with understanding prescribing practices at this time in hospitals.

[Slide]

In summary, I hope I have made the case to you that our knowledge of drug use is really

largely setting specific at this time and, as you look across those settings, the types of data that we have available to us vary tremendously depending on those settings. I have tried to lay them out here from the most detail, which is in the outpatient pharmacy world where we are actually on the verge of exploring patient level data, all the way down to outpatient clinic, where we really have no patient level data at all. So, depending on where a product is used, in which setting, really determines how well we can answer the key questions about understanding a product's use.

[Slide]

Some of our future challenges are clearly to try to pursue data to fill the holes that I have described. We would like to increase coverage of the settings that we have; some of the holes in inpatient data. A high priority is pursuing outpatient clinic data because, particularly in areas such as oncology, this can be the mainstay of product use. In home healthcare and long-term care, as they emerge as more and more important

settings of care, we would like to be able to understand better how products are used.

We would also like to increase coverage of certain populations where we know that drug use is highly prevalent and that safety issues are of great concern. The elderly is clearly one of those populations, and we have been in conversations with colleagues at CMS. As the Medicare Part D data become available, we are trying to understand whether we can access that data and tap into a greater understanding of how elderly people use drug products. We have been working for the last couple of years with our colleagues at FDA, as well as at NIH, to better understand the use of drugs in children. So, that is an ongoing effort, as well as the use of drugs in pregnant women.

Finally, the HIV-infected population is of great interest. There is a variety of drug products that this patient population takes but, given that they often receive their drugs through channels that are not often covered by our distribution systems-- for example specialty

pharmacies or outpatient clinics, we find that our data sources are not really adequate to understand patterns of drug use in that disease group.

Finally, if I am going to describe the nirvana of drug use data, we will love to see efforts, as the technology evolves, to be able to link patients, to follow their drug use in an outpatient setting, follow them into the hospital and see what drugs they take, and then follow them as they come back out. If those data systems emerge, we would be very interested in knowing more about them.

[Slide]

Finally, I would like to acknowledge the folks back at the ranch. This is the group of pharmacists, epidemiologists and contract specialists who are responsible for obtaining, acquiring, analyzing and interpreting the drug use data that is available to the agency. Thank you.

Question and Answer Period

DR. GROSS: Thank you very much, Dr. Staffa. At this time, we will have a question and

answer period for the people sitting around the table, and the question should be related to the material covered this morning. Are there any questions? Yes, Henri?

DR. MANASSE: I will just address this question generally to the speakers this morning. I think for a number of years we have talked about spontaneous reporting and the fact that we are perhaps only picking up ten percent of all the reports. Have there been any studies looking at the 90 percent, that is, sampling perhaps or in other ways investigating what is not being reported? It may sound like kind of a strange issue but, bottom line, the spontaneous system isn't picking up the proportion of signals that we would like to pick up I suspect.

DR. GROSS: Any other speakers this morning like to address that question? Joyce?

DR. WEAVER: Just to say that I don't think we know anything about that 90 percent. We don't know a whole lot about what is being reported to us, let alone the other portion.

DR. GROSS: Annette Stemhagen, you had a question? DR. STEMHAGEN: My question is for Dr. Staffa. The industry also struggles with

the calculation of denominators but in periodic reports there are some data presented by the sponsor for a particular product. You didn't mention that as a source of data. Do you look at that, and is it useful, and what kind of conclusions can you provide on that?

DR. STAFF: Well, I think that is clearly looked at but I guess I look at it as industry is dealing basically with the same marketplace as the FDA is dealing with, and when we become aware of data sources, typically the industry does as well and we often find that we are often using the same data sources to try to estimate but I find that they are running into the same issues.

DR. STEMHAGEN: I guess my question was whether there were anything different based on some other methods that they might have.

DR. STAFF: Typically, what we see in the periodic reports is often what we can ourselves

generate too. In fact, we look to see if industry has been able to find any source of use data that we can then explore as well.

DR. GROSS: While you are here, Dr. Staff, what is Medicare Part D?

DR. STAFFA: I apologize. Medicare Part D is actually going to be the provision where Medicare pays for prescription drugs for the elderly, where currently they do not but by doing a comprehensive payment and having it all, hopefully, in one data warehouse we would be able to look at the patients more clearly.

DR. GROSS: Robyn Shapiro is next.

MS. SHAPIRO: I guess I am still unclear about the nature of the analysis that is done with the AERS data. For example, I was involved with a DSMB in industry that identified a particular risk for a particular subpopulation, and I am wondering if that type of question is looked at, or can be looked at by the government using AERS data or not.

DR. WEAVER: Yes, we do look at subpopulations. When we do a case series we are

looking for populations that may be at increased risk. I am not sure exactly what it is that you are talking about so much.

DR. GROSS: Richard Platt?

DR. PLATT: First I want to congratulate all the speakers. They were terrific particulars. I have four questions. Shall we go in turn and I will ask the first?

DR. GROSS: Why don't we start with one? Could the speakers from this morning perhaps stand up at the podium? It would be easier for you to answer questions as they come up.

DR. PLATT: My first question is maybe best addressed to Dr. Seligman or Dr. Trontell but anyone who can answer it, can you give us an idea about the magnitude of the resources that are allocated to each of the activities that we heard about this morning?

DR. SELIGMAN: In terms of the AERS system, we probably devote in the neighborhood of about 5.5 million dollars a year towards not only maintaining the database but also most of those

resources go to actually coding reports that we receive. As you can imagine, as was pointed out in the presentation, when we get up to, you know, the neighborhood of 300,000 to 400,000 reports a year it is a lot of coding, and a report that has to be coded, you know, can cost us in the range of $25-30 a report to get into the system.

We are addressing that particular concern through current efforts to encourage, and ultimately to require electronic reporting of all AERS reports which will, hopefully, mitigate our need to code all those.

In the area of drug utilization databases, Judy, do you want to say something? I would probably estimate we probably spend in the range of about--well, I will let you give the precise numbers but I am going to say 2.5-3 million dollars a year.

DR. STAFF: I think that is about right. Of course, for us trying to answer questions, we would like that to be a lot more as we identify holes in the data. That is actually where the

challenge is, to try to prioritize. When you are trying to buy data to answer any question that might come up next, where do you put most of your resources? You want to put those in the areas where you can answer the most questions and get the biggest bang for your buck. But it has been a challenge and I think we have been trying to prioritize that.

DR. SELIGMAN: So, in terms of the speakers this morning, I think those are the primary resources that we spend.

DR. GROSS: Richard, we will get back to you with your other questions. Stephanie Crawford?

DR. CRAWFORD: Thank you. I also would like to thank the speakers. You answered three of my five questions and, hearing what Dr. Platt just went through, I will just ask one of the two remaining right now. For any of the speakers, since we are thinking about drug safety and risk management, what signals, in addition to adverse events in general--what signals might indicate threats to the integrity of the pharmaceutical drug

supply, such as counterfeiting, adulteration, misbranding, etc.?

DR. WEAVER: One of the toughest things that we deal with is reports of lack of efficacy. Lack of efficacy is really the most frequently reported event to us where a patient is complaining or a physician is complaining that the product is not working as they expected. The difficulty, of course, is teasing out when it is something such as what you have stated. We have had signals come up in that way.

DR. STAFF: I would also just say that I believe there is also another reporting system at FDA that deals with issues of product quality. DQRS is what it is called.

DR. HOLQUIST: Drug Quality Reporting System, and that is run by the Office of Compliance.

DR. GROSS: The next question comes from Lou Morris.

DR. TRONTELL: Can I just add one comment for Dr. Crawford? In those instances where

counterfeiting or product quality results in patient harm, we do pick up those adverse events in our reporting system but the Drug Quality Reporting System that Dr. Holquist described is maintained by our Office of Compliance, and includes not only instances of patient harm where they share the same reports we do, but complaints of a product not working or issues where the product may crumble or discoveries of counterfeit products.

DR. MORRIS: Also for Judy Staffa, Judy, I notice you didn't mention any names of any sources. Is there any reason why you can't mention different sources?

DR. STAFF: Well, some of the data sources that we are pursuing we are not really allow, just by federal law, to mention.

DR. MORRIS: I am interested in learning what sources you have in-house, what sources can you pursue under some outside contract, and just knowing what your resources really are. I mean, how much of the two or three million goes to IMS sources? Which ones do you purchase? Are there

other contracts? What do you have?

DR. STAFFA: Currently, we actually purchase a number of audits from IMS Health, I think Katrina Gary is out there in the audience and she can correct me. I think it is in the realm of a million dollars a year to access a number of different audits on prescribing data, on outpatient prescription data.

DR. GROSS: Let me comment on that. Shalini keeps us on the straight and narrow here and, because of conflict of interest issues, those specific answers can't be made. Lou, sorry about that. The next question comes from Jackie Gardner.

DR. GARDNER: Actually, I was just going to ask Henri to follow-up on his question to the staff, did you have a suggestion for what might be sampled that would give them more information about underreporting?

DR. MANASSE: Well, let me begin by saying that from a patient perspective and from a broader population safety perspective, a spontaneous reporting system that only gives us ten percent of

the reports I think is a serious issue. It seems to me that we need to somehow figure out what is going on in those 90 percent of cases were there may be adverse effects ongoing and determine that.

I don't have a particular methodological suggestion, other than to say that I think there needs to be perhaps some further dialogue, either in this committee or collaborating with some other committees, looking at the entire policy question should we continue with spontaneous reporting. I think part of the answer to that needs to be driven by this other 90 percent which none of us has a very good handle on. I think, Jackie, that methodological experts, coupled with patients, coupled with prescribers, coupled with pharmacists should be able to get their heads together to get a better handle on this.

DR. GROSS: Anne?

DR. TRONTELL: I will try and volunteer a partial answer. It is very hard to know what we don't see or observer. However, some of the discussions you will hear later today and tomorrow

will talk about data systems which we are actively exploring to detect safety signals, as well as to quantify them. So, I might ask you to consider at the end of our discussions how you might rephrase your question to pursue it better.

DR. SELIGMAN: Part of the challenge today is to ask the committee to put their thinking caps on because this is one of the questions we have for the committee this afternoon which is, indeed, how to improve or modify or in some cases maybe even abandon the Adverse Event Reporting System if there are, indeed, better ways to do this kind of surveillance work.

DR. GROSS: The next question comes from Ruth Day.

DR. DAY: I am concerned about the perceived reliability of spontaneous reports. It is often said that physicians reports are more reliable, say, than consumers and in general that is probably the case. But I am wondering if studies have been done retrospectively, after safety signals have been recognized, to go back in

and look at the reliability of reports from physicians, from pharmacists, patients, other consumers because it may well be that some kinds of signals are detected better or worse within the individual groups. I was just wondering if any efforts have been directed towards that.

DR. GROSS: Would anyone like to answer that question?

DR. DAY: And if not, whether you think it would be useful to do it some day?

DR. GROSS: Evidently a very interesting question, Ruth. Perhaps later on in the session we will have some other information. Art Levin has the next question.

MR. LEVIN: I sort of have, I guess, a comment and then a question or two questions but, like Richard, I guess I will do one at a time. The comment is, of course, on Medicare Part D which strikes me as a wonderful opportunity to have a rather large captured population that is going to stay in that scheme until they get carried out. Unlike the other problems with populations and

people going in and out of plans and in and out of coverage, that is a very stable population that I think offers, you know, a really unique opportunity to follow people over time and gather a lot of information. So, I would hope that there is more than conversation since we are looking at a 2006 initiation of Medicare Part D. I would hope that those conversations are well along and that people are really thinking seriously about how to use Medicare Part D data for these purposes.

My question is sort of like "a what happened to question." Following the IOM report in '99. "To Err is Human," Secretary Thompson announced that an effort to try to get a few agencies, including FDA, CDC and it might have been the VA and somebody else, to begin to share data with an eye towards improving patient safety. I am just wondering, at least to my mind, that effort has sort of gone not very far, if anywhere, and I am just wondering where that is.

DR. STAFF: Well, I can speak to some of that. I think there have been efforts to

collaborate across agencies and we have tried very hard to see what kinds of data other agencies have and try to work with them when specific safety issues arise. We have done studies with the VA using their data when circumstances indicated that a drug we were looking at was very prevalently used in that population. We have had conversations with DoD about using their data as well.

We have run into some logistical difficulties. What we have been trying to do is establish standing memoranda of understanding so that the chutes are greased, if you will, because when a safety issue emerges you really need to act rather quickly and you often don't have time to put the legal arrangements in place. So, there have been some efforts to do that, but it can be difficult because there are just levels in both agencies. But we have made efforts to work in that direction.

MR. LEVIN: I think the intent was to put it all in a database that everybody could use, to have the benefit of having all the government

agencies that receive information relative to pat safety contribute that to a single database that would be available for searching by any one of those agencies. What I am hearing is that that really hasn't progressed.

DR. STAFFA: Anne, you might be able to comment.

DR. GROSS: Anne Trontell?

DR. TRONTELL: The logistics of what some of us as epidemiologists think of as sort of the "grail," the one database that we can all access and query certainly has formidable logistical challenges. Most of the efforts in that arena are talking about how different data systems can communicate with each other--health level 7, adoption of standard terminology. So, those move but probably at a slower pace than some of us might wish could occur.

I wanted to add to what Dr. Staffa said. We have a specific collaboration with the Centers for Disease Control that you will hear about. I think it offers an example where existing data

infrastructures, in this case an emergency department surveillance system, can be incrementally reconfigured to pick up adverse drug events. In that context, the discussions with CMS now really relate to how the data collection for the Part D benefit can best be configured so that it would allow FDA to ask important questions. In that regard, I think we are greatly assisted by the fact the current commissioner of CMS is a former FDA commissioner and is highly dedicated to making those data systems available for use not only for drug safety but for drug efficacy and other issues.

DR. GROSS: Does anyone else have further questions besides Drs. Platt and Crawford? Yes, Sean?

DR. HENNESSY: My question I think is for Dr. Weaver. We learned that FDA receives about 400,000 anecdotes a year having to do with adverse events--and I use "anecdotes" in the fondest sense. I think that we can learn a lot from those. I am wondering about the relative value of treating those as individual case reports versus trying to

treat them as if they were epidemiologic measures. In particular, of all the resources that go into handling the spontaneous reports, is there much follow-up with the reporters to find out the crucial details that weren't included in the individual reports? If not, is staffing the issue, or are there other issues that prevent follow-up with the reporters?

DR. WEAVER: With the number of reports that we receive, and you saw the number of safety evaluators we have, obviously, staffing is an issue. We do follow-up on cases where we have an important safety issue that we are pursuing. So, we can do follow-up; we do follow-up; but do we follow-up each of those 400,000 reports to make sure each report is complete? No, we don't.

DR. TRONTELL: Sorry to keep jumping up. I wanted to make Dr. Hennessy aware that more than 90 percent, upwards of 95 percent of the adverse events that come into the system come to us by manufacturers. Again, in this regard they have regulatory requirements to actually do the

follow-up. There are attempts to get the most complete information. Again, sometimes we run up against the fact that clinicians don't necessarily have the time or inclination to bring extensive background information to a case. But there is a substantial effort that industry undertakes, in addition to what our own safety evaluators might take, in a particularly compelling case where missing information might make a big difference.

DR. GROSS: We are almost out of time. Richard, one more question.

DR. PLATT: This is a follow-up, Anne, to your answer to Arthur's question about your conversations with CMS. When you say things are going well--this is a comment disguised as a question--when you say the conversations are going well in terms of FDA's ability to use that information, is it going well in terms of having access to a fully linked system that will link the prescription drug use to other information that Medicare has about those individuals--their diagnoses, their procedures, their demographic

information?

DR. TRONTELL: I think that is the intent with the CMS data. Right now, such data are only linked for the proportion of Medicare beneficiaries who are also entitled to Medicaid, so-called "dual eligibles." I actually worked at CMS when it was called HCFA. With the size of the data system, with more than 40 million inpatient and outpatient records there will need to be linkages. I don't know to what extent they will be standing linkages. They may, as Dr. Staffa said, need to be constructed on an as-needed basis to ask specific questions. We are not yet at that level. Right now, the outpatient drug utilization data is not benefit and the data system to be in place is being constructed but, certainly, CMS has information on beneficiaries. They follow them through their HEC number which is a variation of their social security number. So, the potential for linkage is there and that is really the great power that that data system might offer to us when the benefit data is added to it.

DR. PLATT: I just want to be sure--

DR. GROSS: I think we are going to have to stop. I am sorry. Shalini Jain has a couple of

points of for us all.

MS. JAIN: I am sorry to take away from our meeting time but I have two quick announcements. The first is if there is anyone in the audience that had registered for the open public hearing session process as a presenter, if you could please check in at the front desk and make sure that we are aware that you are here. We don't want to miss your ability to present later on today. In addition, if anyone has a cell phone in the room, if they could please either put it on vibrate or turn it off, and if you have to take call please step outside the room. Thanks.

DR. GROSS: Thank you, Shalini. We will now take a 15-minute break and reconvene at 10:15.

[Brief recess]

DR. GROSS: If I could have your attention please, please sit down. Dr. Gould, would you come up to the podium please? Dr. Gould? At this

particular point, Dr. Larry Gould is Senior Director, Scientific Staff for Biostatistics and Research Decision Sciences at Merck Research Laboratories. He will address issues in the practical application of data mining techniques for pharmacovigilance.

Issues in the Practical Application of Data Mining Techniques to Pharmacovigilance

DR. GOULD: Thank you, Dr. Gross. Before I start, I should like to make two comments, first to thank the FDA for inviting me to make this presentation. It is, indeed, a privilege. Secondly, I should like to commend the FDA speakers this morning for their really outstanding and lucid description of many of the difficulties and issues that are faced in trying to do pharmacovigilance based on spontaneous reporting databases. This makes my job a great deal easier.

[Slide]

As was pointed out, clinical trial safety information is incomplete. There are few patients in clinical trials relative to the universe of

patients that one would see once the drug is out on the market, and rare events are likely to be missed simply because they are rare. In fact, the patient population that you see in clinical trials really doesn't necessarily mimic the real world. It is a relatively restricted patient population.

So, what one needs to do to get a better idea of the potential safety and liability or toxicity of a product is to get information from postmarketing surveillance and spontaneous reports, and this is pharmacovigilance.

Now, this traditionally is carried out by skilled clinicians and epidemiologists, and that is true at the FDA and it is true in every company that maintains repositories of information about its products. There is a long history of research on this particular issue, and it goes back to something over 30 years. So, it is not like spontaneous reports are something that have recently become of interest. This is something that has concerned people from a practical and technical point of view for a very long time. The

references that I have here by Dr. J. Finney and Dick Royall are statisticians and so this is actually an issue that has occupied the attention of the statistical community for a very long time as well.

[Slide]

Now, I have here this really complicated diagram of how signals get generated in the traditional way. You have patient exposure, you know, the usual way patients take the drug and there are some potential signals that might be generated. One has to get some information from marketing folks to get an idea relative to risk of exposure to the extent to who is being exposed. But, as was pointed out this morning, that is not a trivial task. One may see a simple suspicious case or a cluster, and this perhaps could lead to identifying potential signals. We have some evaluation of accumulated data. It is easy to think it is attractive. It is seductive to think of having various databases available for managing information--you go to this database, that database

and the other database--until you actually have to go and do it and you find out that the databases are not done in a manner that lets you gather information from different databases very easily. This is a non-trivial information process and task and so programmers have to be consulted--what integrates the information; gets refined signals; what is the comparative data from other reports; what is the background incidence, basically consulting the literature. After doing all of this, you may come to a course of action that would include perhaps a recommendation for a change in the label.

[Slide]

So, what are some limitations of the traditional approach? Well, you have to remember these are incomplete reports of events or things that happened. They are not necessarily reactions; they are events. We have a difficulty. How do we compute the magnitude of an effect, assuming we know what an effect is? As has been pointed, in databases many events are reported and many drugs

are reported. This is a multiplicity issue of no small magnitude. There is bias and noise to a system. That has also been pointed out this morning. It is difficult to estimate incidence. It is actually impossible to estimate incidence, or damn near, because the number of patients at risk and the patient-years of exposure are seldom reliable. Even if one wanted to estimate it, as the question was raised this morning, if you wanted to estimate something about the 90 percent that you don't see, the difficulty you have there is that in a statistical context when you are doing sampling you have to identify the frame, that is to say, the potential population from which you wish to do your sampling. It is not at all clear what the frame is.

I will tell you my own persona opinion, very, very strongly held, it is inappropriate to consider incidence using only spontaneous reports. You need to know something about exposure and the spontaneous reporting system simply does not give that to you.

[Slide]

So, here is the pharmacovigilance process. We have the traditional methods. We have data

mining. Both of these give you a way to detect signals. These generate hypotheses. This is the key point. You then have to go through a very considerable exercise to refute or verify these. The point is the real value here is generating the hypotheses, just focusing your efforts on what you want to chase down, follow-up and get detailed information about. It is a way to use your skill resources efficiently. You may get insight from outliers. You might find out whether this is mechanism based, this type A, or idiosyncratic, type B. That is one of the things you would find out in chasing these things down. There may be estimated incidence. Consider what the public health impact benefit and risk might be and act.

What might you do? You can inform folks about the potential of this particular relationship. You could change the label, which often is done. Or, you could restrict use and

withdraw the drug from the market. So, the process itself looks pretty complicated but actually it is very logical and there is a reasonable sequence of steps.

[Slide]

Now I am going to talk about data mining for a while, and I am actually going to try to be non-technical about it mainly because the concepts that are really important don't really have to do with the technical details or the algebra. What might one use data mining for? Well, you might identify subtle associations that could exist in large databases that you might not otherwise find. You could perhaps have early identification of potential toxicities. You might identify complex relationships that are not apparent by simple summarization.

For example, you might consider how two or more drugs interact in producing--I hesitate to use the word "producing" because it implies causality--in being associated with the occurrence of some kind of an adverse event because it may be

there is a kind of synergy between drugs in this regard. So, spontaneous report databases, by virtue of their very size, give you the opportunity to investigate possibilities of this sort; likewise, possibilities of constellations of adverse events that might reflect some effect on a body system. It is a screening tool to identify potential associations for follow-up, as I pointed out before.

Now, there is more to pharmacovigilance than data mining. Data mining is a refinement to discover subtleties. You still need the initial case review. This is really important, and the FDA does it and every industry does it. Case reports, as they are received, are reviewed. If there are really important events, sometimes called sentinel events, then what happens is these are chased down, no matter what. You are not going to waste time worrying about incidence, or data mining, or anything of that sort. Stevens-Johnson syndrome, agranulocytosis, anaphylactic shock, death--these are going to be followed up, period. Even if you

see one case, it is going to be followed up and it has nothing to do with data mining.

It is necessary--I mean, I say that simply as a caveat simply because people think, by golly, are you going to be losing something by the time required to do this, and the answer is no. Now, once you have done all of this, it is really essential to go for clinical, biological and epidemiologic verification of the apparent associations. What you have here is associations. That is all you have; no causality. So, in looking at this one of the things one needs to think about is how one might use data mining most effectively in improving pharmacovigilance practice. It is considered as an intellectual activity; considered as an idea that has been around for a long time; considered as something that is taken on faith; considered as something that is required by the regulations. It seems to make sense. The question really comes as to how can we use this most effectively:

[Slide]

A little bit about statistical methodology, just to put us in the frame of how these things work, first of all, the statistical

methodology is not the key issue. Mostly you are going to use variations of a two-way table. In many two-way tables, clearly, you have many target drugs and target adverse events but you have a target drug, a target adverse events and then you can simply say, well, all other drugs, all other events, and you have a simple table like this. The basic idea of all pharmacovigilance is that you are going to flag something when the number of events that you see, divided by the expected number under some assumption about no association between the drug event and the reporting is large. If you, in fact, see many more events than you might expect to see if there were no association you might say, aha, maybe we need to chase this down.

Now, there are some possibilities--reporting ratio, proportional reporting ratio and odds ration all are variations on this theme. If you see this occasionally

described in the literature, even in the drug safety literature, in effect, there is nothing remarkably different about these things. There are minor technical differences but, in fact, they are all ways of doing this.

[Slide]

In more recent times, within the past six or seven years, people started considering the application of variability. You might, for example, have three events reported on a drug and those are the only three events you see in the database. Well, if you consider the denominator that is going to give you a whopping large value of the number of events divided by its expectation because its expectation is going to be very, very small, much less than three. It may be about 1/100. So, you are going to get a whopping value of that ratio but you only have three events. So, clearly, there has to be a great deal of uncertainty associated with that, much more so than if you happened, say, to have 150 events and you find that that number of 150 is considerably larger

than you might expect under the assumption of no association. That being the case, one needs to bring in the statistical ideas that can help deal with the variation.

Ways people do this are the usual chi-square statistic, Bayesian methods and empirical Bayesian methods. In fact, if you have more than a few reports they all give you pretty much the same results. So, it is of technical interest for those of us who are interested in these technical things but in fact, from a practical operational point of view, it makes relatively little difference.

One big problem with all of this is that we go through all of this and you see something that lights up and you say, aha, it is a signal. Well, maybe it is and maybe it isn't. That is why the follow-up is needed because there really isn't any gold standard for this. We don't know independently of the mechanism used to identify a signal whether that is really a signal or not. If we see something that looks like a signal and we

can chase it down by obtaining a great deal of additional information and find out whether that was real or not, but you are not going to know what signals you missed if you never see the signal; if you never are motivated to go chase that down either because there are not a whole lot of them that occurred, or not many occurred than you might expect or, even worse, because none occurred at all in your database.

There is an ordering of how these statistics go but it really doesn't mean that any one identifies real associations. The point is how do you know that you haven't got a false positive or a false negative? The answer is you don't. That is a technical issue.

What are some limitations? Well, you have heard these before and you will hear them again; reality is worth repeating. There is significant underreporting depending upon the seriousness or novelty of the event, the newness of the drug and the intensity of the monitoring. There are different regulatory reporting requirements in the

U.S. and overseas. This reflects reporting practice, not incidence. We are not talking about incidence here; we are not talking about risk. There are synonyms for drugs and events and this is a sensitivity loss; this is a coding issue. Much duplication of reports, especially in the AERS database. That has already been mentioned. We don't know what the exposure rate is because we don't know who is taking what, and Dr. Staffa pointed out this problem. And, more to the point, in any given report the patient report may mention several events, several adverse events the patient was experiencing, several symptoms or signs and the patient may be taking several drugs at the same time and you don't have any way of knowing that any particular drug caused any particular reaction. That is why the follow-up is important. There is a lot of duplication here.

[Slide]

This is a major limitation that is often ignored, that the cumulative reports cannot be used to calculate incidence or estimate drug risk. This

is probably the fourth or fifth time that I have said it but I have to repeat it because it is really an important point. This is absolutely vital. More to the point, comparisons between drugs cannot be made from these data, simply cannot. Unfortunately, this occasionally still is done and there is literature on this. The difficulty from a really practical point of view is that the disclaimers that say, well, we know we can't do this but we are going to do it anyway are sort of like giving testimony in court that the judge disallows. Well, the damage is done; you dropped the bomb and people don't necessarily read the fine print. So, it is just inappropriate to do that.

It is easy to show differences with data mining techniques. This is a really easy thing to do but it is impossible to make valid inferences about causality. In fact, these relationships that are brought forth may, in fact, be quite misleading.

[Slide]

So, what are some implementation issues? Well, portfolio bias in company databases can lead to inaccurate estimates of relative reporting

risks. Merck is going to have a lot of reports for Merck drugs. Pfizer is going to have a lot of reports of Pfizer's drugs, not as many as Merck's drugs because they have to report to Pfizer. Ditto, Glaxo, Avantis and everybody else. So, every company in their databases has a bias weighted towards their products.

Bit question in terms of implementation--does the public health benefit justify the cost of following up signals detected by routine data mining methods? We don't know the answer to that question, by the way. It is an issue still under investigation.

Another consideration is that there are different ways of doing this. We have variations in tools and databases amongst regulators. So, we can wind up having to do an awful lot of work and carry an awful lot of cost. Regulatory agencies can wind up incurring an awful lot of cost and, as

Dr. Seligman pointed out, one of the considerations that the committee is actually charged with thinking about is do we really need to do this at all? So, we don't know what the benefit is.

There is certainly a literature out there that shows that frequency-based signal detection methods such as data mining could be useful in identifying subtle associations in terms of reporting of adverse events and, therefore, possible and previously undetected toxicity relationships. No question about that. The question is do they have value in an industry setting, and we have to think about that. We are investigating that as well at Merck and I know other companies are too.

What we need are examples of situations where the computerized approach failed to identify important issues and where signals were created by publicity or reporting odd effects. See, the problem here is, when all is said and done, what you need to know is does this stuff work well as a diagnostic or screening tool. Until you know how

it performs as a screening tool it is an interesting exercise but its value still needs to be proven.

[Slide]

What are some other considerations? Well, what is the data mining activity intended to accomplish? This kind of gets around to what this is all about. What are the questions that need to be answered from a clinical, epidemiologic and regulatory point of view? We need to address the impact of various factors such as evolution over time, association with key demographic factors, and so forth, and this has been pointed out earlier this morning as well.

[Slide]

Some more issues--the composition of the database may be important. Important associations of a new drug could be cloaked by events associated with an old drug with a similar mechanism of action. For example, if you wanted to find out what the reporting relationships might have been for angiotensin-2 antagonists, A2A antagonists, you

might think about older drugs such as captopril or analopril that have a similar mechanism of action, whose occurrence may in fact obscure, to some extent, the potential association you see with A2As. Individual company databases tend to be comprehensive about the company but not the general spectrum of drugs. I am going to say this again, databases contain reports mentioning drugs, not demonstrations of causality. If there is a take-away message at all that comes out of this, that is it.

[Slide]

Finally some discussion, what we have found when people go through this exercise is that most apparent associations represent known problems. We know about these. Good. It is kind of an internal validity issue. About 25 percent may represent signals about previously unknown associations. That doesn't mean previously unknown toxicity, just associations.

I have to put a pitch in here after all, statistical involvement in implementation and

interpretation is important. These are situations where issues of bias, error, validity come into play, and in order to make legitimate inferences and draw legitimate conclusions about what you see statistical considerations have to be taken into account. The actual false-positive rate is unknown, as are the legal and resource implications of using the data mining. That is another issue I am not going to get into but, in fact, when you do data mining and find associations, do you in fact have something that is discoverable and, secondly, might this be an issue that pertains to potential litigation? I don't know.

[Slide]

What next? Well, PhARMA-FDA working group is considering ways to address the issue. I have to tell you that, having been heavily involved with that particular project, I cannot be anything but completely complimentary about how people in industry and FDA have worked together. It was an exemplary example of cooperation between regulators and industry in outlining and delineating what is

known about a particularly important technical issue that affects both. I have to say that the group has done a magnificent job. When you all see the paper come out I think you will agree.

It may be worthwhile--one of the issues that people had commented about on the AERS database--to construct and maintain a clean canonical database from AERS to proved a common resource for checking data mining findings based on individual company proprietary databases. This is not a trivial task. It is quite difficult. But, in fact, if everybody is working from the same database and the same resource to evaluate associations, then we come to very similar conclusions. This is a very worthwhile thing to do I think. That is it so I thank you very much for your attention.

DR. GROSS: Thank you very much, Larry, for simplifying the statistical mysteries of life. Shalini has an announcement.

MS. JAIN: I just wanted to let everyone know at the table and in the audience that Dr.

Gould's slides will be posted on the meeting web site so you will have access to those. In addition, if you need a copy please contact me on my e-mail that is listed on the Federal Register, and I will be happy to help you with that. Thanks.

DR. GROSS: Thank you, Shalini. The next speaker is Dr. Carolyn McCloskey, an epidemiologist with the Division of Drug Risk Evaluation, from the Office of Drug Safety. She will talk about data mining with AERS and the FDA's spontaneous adverse event reporting system.

Data Mining AERS, FDA's (Spontaneous) Adverse Event Reporting System

DR. MCCLOSKEY: Good morning.

[Slide]

My name is Carolyn McCloskey, and I will present the data mining activities in our division, the Division of Drug Risk Evaluation, in the Office of Drug Safety. Most of the epidemiological data mining work in our division and the data mining findings in this presentation heavily represent Dr. Rita Ouellet-Hellstrom's work.

[Slide]

The objectives of this presentation are to provide a brief history of data mining activity at

the FDA leading to the current use in the Division of Drug Risk Evaluation under CRADA, or Cooperative Research and Development Agreement. The CRADA involved development of the webVDME application to mine the AERS database and a corresponding pilot evaluation, in addition to some other activities. Lastly, I will cover the future directions expected for data mining and the pharmacovigilance activities of our division. [Slide]

FDA's data mining activities of the AERS database were initiated in 1998 by Ana Szarfman using a grant by the Office of Women's Health. These efforts led to a CRADA in March, 2003, which continues to the present. Data mining research development continues in Ana's group on drug-drug interaction, logistic regression modeling and other areas.

[Slide]

Currently the main data mining activities

in DDRE include various projects related to the CRADA. Through this agreement, FDA collaborates with Lincoln Technologies, Inc., a private software firm, to develop web-based data mining software that can be used in DDRE for pharmacovigilance. Neither party contributes funds. FDA has contributed AERS data, technical knowledge in the areas of epidemiology, biostatistics and information technology and staffing resources. Lincoln has developed the data mining environment and user interface and has also contributed computer hardware and staffing resources.

The CRADA objectives are to further development user-friendly data mining application in the web-based environment. It involved performance evaluations by the safety evaluator and epidemiology user groups and training followed by continued development and refinement of the application. Dr. Gould presented some statistical methods used for data mining. Here, at FDA, data mining is applied to the AERS database to screen for new safety signals.

[Slide]

The webVDME application uses the multi-item Gamma Poisson shrinker statistics, or

MGPS, which calculates the empirical Bayes geometric mean, or EBGM, and the confidence intervals. The mean is an observed to expected score which adjusts for sampling variation such as sample size, but it does not adjust for reporting bias. The application also allows calculation of EBGM scores using various strata. We, in DDRE, traditionally calculate the EBGM score using the standard stratification of gender, age and year or receipt, but we can customize the data mining to use other strata.

[Slide]

The confidence interval calculated around the EBGM mean represents a 90 percent probability that EBGM will occur between EB05, the lower bound and EB95, the upper bound.

[Slide]

This slide is an example of EBGM's adjustment of sampling variability for a particular

drug that has adverse events of myalgia, which has a large number of counts, and sinal osteoarthritis, an adverse event with a small count. The MGPS calculates an expected value for these PT terms within the database. Looking at the observed to expected ratio or reporting ratio, the RR in the box, we have a score of 4.99 for myalgia and 6.16 for spinal osteoarthritis. As a result of the adjustment calculation, the EBGM or mean or 4.97 that is under that second arrow is very close to the reporting ratio because of the large numbers of reports. In contrast, the EBGM score of 4.54 for spinal osteoarthritis is smaller than the observed to expected ratio of 6.16 because the statistical tool had more influence in adjusting variation for small numbers of reports. Both scenarios represent possible data mining signals even though the spinal osteoarthritis EBGM score has a wider confidence interval, the lower boundary of EB05 of 3.03 is still above 2.0, a level frequently considered a threshold for a data mining signal.

[Slide]

A pre-pilot performance Bayes evaluated the webVDME record retrieval with AERS case retrieval. This included evaluating the

nomenclature of multiple trade and ingredients and the allocation of the drug to the suspect versus concomitant category. The logic for removal of duplicate reports was extensively reviewed. At the same time, the Office of Information Technology was also evaluating the system performance.

[Slide]

Once the performance evaluation was acceptable, we then started a pilot phase to learn more about the strengths and weaknesses of this tool to enhance our safety evaluation process. Medical safety evaluators in our division who routinely review safety reports evaluated data mining scores for drugs and biologics to evaluate the application performance with respect to the listed criteria.

[Slide]

The epidemiologists in our division conducted their own tests as part of the pilot

study. The temporal trends analysis found signal scores were affected by clustering, litigation, switching to the new application for the AERS database, and changes in MedDRA terms. The effect on the EBGM signal of using the trade versus ingredient name stratification by gender, age and receipt date and use of the suspect category versus suspect and concomitant categories were evaluated. The epidemiologists also evaluated the complexity of the application and its use.

[Slide]

The next two slides are examples of the data from the epidemiologists pilot studies of the differences between a new and an old drug. They show that the EBGM rankings and confidence intervals can change over time, indicating that factors within the AERS database and factors outside AERS can influence data mining findings.

[Slide]

This is an example of a newly marketed prescription drug about one year after approval. The EBGM score, or adjusted observed to expected

mean of the drug event combination, is charted for each of the adverse event preferred terms or PTs. The PTs are hard to read but I am only going to cover a few data mining points.

Note the large confidence intervals, the blue lines, for the two preferred terms or PTs of dysthymic disorder and ejaculation disorder NOS. These are indications for the drug but also in AERS as an adverse event associated with the drug. Their EBGM scores are above the horizontal red line, the usual EB05 threshold of 2.0, for consideration of their meaning as a possible signal. Those PTs with wide confidence intervals may indicate a small number of reports. The rest of the PTs have much smaller confidence intervals around smaller EBGM scores. Note that the three boxed PTs all include a description of a convulsion. The reports representing these scores may be evaluated individually or another data mining run could be done using grouped PTs such as the high level MedDRA term of adverse event codes.

[Slide]

This is an example of an older drug that has been on the U.S. market for more than ten years. Here again are EBGM scores charted for

various PTs reported with this older drug. Note that the EBGM scores for the PTs are above the red line, EB05 of 2.0 and are ranked in descending order of the EB05. The score for the first PT, fear of disease, is very high relative to the other PTs, including the second boxed PT, Torsade de pointes.

A chart review showed that fear of disease cases were in litigation and that Torsade de pointes cases were a signal of its association to the drug that was not overly affected by factors outside of AERS.

[Slide]

Data mining is a useful tool for evaluating disproportionalities in large databases. Some of the conclusions from the CRADA pilot are that data mining is a statistical tool that assists in identifying the usual patterns in AERS data of drug-event combinations. But the patterns must be

interpreted in light of all that is known about that drug and event and about the reporting of that drug-event combination.

[Slide]

It is absolutely necessary to understand the data that is used for data mining so that the interpretation of the data mining information is as accurate as possible. In addition to understanding the pharmacologic, metabolic and clinical characteristics of the drug and the adverse event, it is also critical to understand the external forces that influence the reporting disproportionalities which may be reflected in data mining scores. The data mining results must be interpreted in the context of the drug-event combination. This includes the differences with other drugs in the drug class; factors that affect drug usage; and adverse event reporting.

[Slide]

The AERS database limitations have already been discussed by Joyce Weaver, Mary Willy and Larry Gould so I will not go over these again.

These limitations of the database may be reflected in the data mining scores and should be considered when evaluating data mining results. Both false negatives and false positives are possible when using this tool for signal detection, and the importance of clinical case review and review of reporting biases for risk evaluation is not diminished.

[Slide]

CRADA activities and conclusion of the pilot are providing training and access to WebVDME. Technical problems and refining user customization of the application are being addressed.

[Slide]

In summary, the data mining initiative in our division has shown that data mining signals assist but are only one avenue in prioritizing the work load of case series evaluations. Data mining signals identify associations that are greater than expected within the database but are not an indication of causality or degree of risk. Data mining signals also reflect limitations of the

data.

[Slide]

The lower threshold used for data mining, such as the EB05, is a compromise between sensitivity or false positives and specificity or false negatives. Therefore, the absence of a data mining signal does not mean there is an absence of a drug-event association, and the magnitude of the data mining score does not necessarily correlate with the magnitude of risk. As long as we rely on the existence of a spontaneous reporting data base not only is clinical evaluation necessary in assessing the potential signal, but review of the reports is also required to assess reporting biases.

[Slide]

The future directions of data mining in DDRE include pursuing prospective signal detection and parallel use of data mining with current pharmacovigilance methods. In other groups within the Office Drug Safety there is continued research in more advanced methodology, especially in the

areas of drug-drug interaction and in logistic regression modeling.

[Slide]

I would like to acknowledge the efforts of all members of DDRE who have participated in the data mining initiative and in the preparation of this presentation. I especially appreciate the data mining work of Dr. Ouellet-Hellstrom which I presented this morning. Thank you.

DR. GROSS: Thank you very much, Dr. McCloskey. At this particular point, it is almost eleven o'clock and the open public hearing must go be on time. So, we are going to postpone the question and answer period for later this afternoon or mid-afternoon. At this particular point we will begin the open public hearing. Before the hearing beings I have to read this statement:

Both the Food and Drug Administration and the public believe in a transparent process for information gathering and decision-making. To ensure such transparency at the open public hearing session of the advisory committee meeting, the FDA

believes that it is important to understand the context of an individual's presentation. For this reason, the FDA encourages you, the open public hearing speaker, at the beginning of your written or oral statement to advise the committee of any financial relationship that you may have with any company or any group that is likely to be impacted by the topic of this meeting. For example, the financial information may include a company's or group's payment of your travel, lodging or other expenses in connection with your attendance at this meeting. Likewise, FDA encourages you at the beginning of your statement to advise the committee if you do not have any such financial relationships. If you choose not to address this issue of financial relationships at the beginning of your statement, it will not preclude you from speaking. Thank you. The first speaker, please.

Open Public Hearing

Consumers Union

MS. KENNEY: Thank you. My name is Jeannine Kenney. I am standard Consumers Union.

We are a non-profit publisher of Consumer Reports magazine.

In the interest of disclosure, I will tell you that we receive no commercial support whatsoever. We accept no advertising in our publication. We are fully funded by the subscriptions to the Consumer Reports, as well as our on-line services. So, we have no financial interest in the outcomes of this meeting.

I will tell you a little bit about what we do. Our magazine ranks seventh nationally in terms of subscriptions. We have 4.5 million magazine subscribers and one million on-line subscribers. We also have been involved in health advocacy for many, many years. We publish Consumer Reports on Health, a monthly newsletter with 400,000 subscribers.

We also recently launched a project called Consumer Reports Best Buy Drugs which uses data on safety and efficacy of the drug effectiveness review project, out in Oregon, and then puts that information in the context of cost to give

consumers information on what we would call "best buy" drugs, the safest, most effective drugs that are also the most affordable. This project's goal is to counteract some of the drug detailing conducted by the drug manufacturers to doctors and some of the direct-to-consumer advertisements that consumers are bombarded with that don't necessarily provide the full story on safety and efficacy.

We became involved in drug safety related issues back in the 1930s, before the first FFDCA was first enacted when our predecessor organization published the book "100 Million Guinea Pigs."

Our comments, which you have in written form which are much more detailed, really focus on broader policy issues. The issues that FDA is presenting to you today really focus on how to improve the AERS data system; how to improve postmarket monitoring really from a technical and scientific standpoint. But we ask you, as you are looking at those data safety trees, to also look at the policy forest that sets the context for this debate and, certainly, that has been largely

impacted by some of the safety crises that FDA has faced over the last 12 months.

Some of those policy issues in question are fundamentally does FDA have the authority to get the data that they need to identify safety signals, and take action to manage risk when those signals become pretty significant? Do they have the resources to conduct postmarket drug safety? And, are there structural changes that FDA needs to address, what we think are some inherent conflict of interest involved in this process?

This morning Dr. Weaver and Dr. Willy presented the limitations of the AERS data to you and the epi. analysis that is conducted on those data. This is one of the few sources of information available to FDA. Obviously, we are talking about larger linked databases as well.

One of the questions you are supposed to answer is whether or not, or under what circumstances, different types of studies would be appropriate for different types of drugs--epi. studies, clinical trials, drug registries and so

forth. It is important that you understand that, regardless of how you answer that question, FDA has very limited ability to gather that information. It can't require additional clinical trials once a drug has been approved. It can ask for them; it can negotiate with the drug companies but it can't require them, and I think most consumers would be stunned to know that.

Its resources are extremely limited. As you heard, five million bucks spent for analysis of the AERS data. Put that in the context of the Office of Drug Safety's overall budget which is less than 30 million dollars annually, and then put that in the context of the amount that is spent on the drug approval process which is about ten times greater. So, the folks presenting to you here this week are doing their jobs with very limited resources and we are seeking greater resources. Obviously, that is a congressional issue. It is also an agency policy question, how they allocate resources.

Then, the third question is, regardless of

whether data sources signals some safety questions, what is FDA's ability to act on that? And, can it act timely? I will give an example. We heard this morning that the AERS data identified the fatal skin reaction problem associated with Bextra, prompting a label change in 2002. It wasn't until late 2004, however, that a black box warning was put on its label which is obviously going to be flagged more readily by consumers.

So, our recommendations to you today are to focus on these broader policy questions. FDA is asking you much more specific questions. I don't think your answers need to be limited to those questions and a signal from you that some fundamental issues need to be addressed would go far.

I have the red light on so let me summarize very quickly. We need improved pre-approval practices that will better inform postmarket safety actions. Those are highlighted in the written comments. And, we would like to see greater authority, resources and independence for

the Office of Drug Safety so that they are better empowered to conduct their postmarket safety surveillance.

Finally, just a process question, since public commentors are required to submit our comments to FDA about a week and a half in advance of the meeting, it might be helpful if those comments are provided to this committee prior to the meeting so that they have a chance to read them and evaluate them. Thank you.

DR. GROSS: Thank you, Miss Kenney. Speaker number two, please?

Public Citizen

DR. WOLFE: Thank you. There was a meeting just about three years ago, minus three or four days, on the same kind of topic. I think that meeting was useful; I hope this meeting will be even more useful.

I have found it useful to me to think about the various stages of health regulation much in the same way that Dr. Donald Beattie, of the University of Michigan, the sort of father of

assessing quality, thought about the stages of the quality of healthcare. He identified the structure; the resources; organizational arrangements, and so forth, the process; the evaluation of providers' performance, particularly if it is shown to improve the outcome; and, finally, the outcome. The point he made over and over again--I sat with him on a government advisory committee for a quite a while--was that unless you link process to outcome and can show that problems with outcome are due to problems in process, or conversely, what you perceive to be problems in the process are resulting in a worse outcome you don't get anywhere.

I have just transferred some of those principles to government health and safety regulation so that in the case of the FDA the structure is obviously the bricks. The more people that are there. The process is laws, regulations, policies, the processes, which we have heard a lot about this morning, of evaluating data. And, the outcome is what is done with all of this.

In the case of drug risk management it is how successful is the structure and process of collecting and evaluating data in terms of causing

change, taking a drug off the market when it is supposed to be; not approving a drug if it shouldn't be approved, and so forth.

Underlying the problems with this process are results of surveys done by us in 1998 and the FDA survey in 2001 and an Inspector General survey in 2003, all focusing on pathology, I suppose, of the process. I will just read for a minute the results of the FDA's study that was done because they were losing a lot of good people. This was done in 2001, and it was in the context of what then CDER Director, Janet Woodcock called a sweatshop environment created in the wake of PDUFA:

The FDA survey found that a third of the respondents--this is in CDER--did not feel comfortable expressing their differing scientific opinions. Over one-third felt that decisions, such as holds, refusal to file actions and non-approvals are stigmatized in the agency. Over one-third felt

that their work has more impact on a product's labeling and marketability than it does on public health.

As I said, we have found similar things before and the Inspector General found similar things afterwards. So, if there is something wrong with the process, namely open debate is not encouraged; people are dissed, so to speak, if they are doing things that might reflect unfavorably on a drug, like testifying in an FDA advisory committee meeting and being told not to because it might influence the vote in the wrong direction, we have some serious problems.

I am going to focus just briefly on the issues of the structure and process, and certainly the FDA could do a better job cheerleading, which is really I think the right phrase, to get maybe 20 percent of adverse reactions instead of 10. The FDA paid for and conducted a very successful experiment in Rhode Island, back 20 years, ago which resulted in a 17-fold increase in adverse reaction reports submitted annually from Rhode

Island compared with the yearly average before. Similar increases were not experienced in the rest of the country--the control. But as soon as the intervention stopped reporting dropped back down.

Now, certainly spontaneous reports, as you heard this morning and know from previous meetings, and so forth, are only one part but they are an important part because they can provide quickly a certain amount of detail that, at the very least, can be hypothesis generating. You can get some detail but not as much from retrospective surveys of Kaiser, and so forth, of patients. So, certainly, that is one area of improvement. There is a paper published recently showing that whereas there are some advantages of data mining, as you have just heard, overall it didn't do significantly any better than the old-fashioned way of looking at spontaneous adverse reaction reports.

I want to use a couple of case examples though and encourage you on this advisory committee to try and get from the FDA the data that would allow you to evaluate a much larger number of case

examples, and these are what goes wrong. Now, there is what goes right also. Certainly, for a number of drugs that were taken off the market or subject to a black box warning the processes occurred because of findings in the spontaneous system, sometimes randomized, controlled trials just slightly after approval, Vigor and the CLASS study. But there are a disturbing number, and they keep going on, of instances where the outcome, namely taking a proper action to protect the public health, is thwarted somehow despite the evidence for doing something. I mean, plan B is something in reverse where at a scientific, medical level the evidence is pretty clear that this post-coital contraceptive should be approved for over-the-counter use but, because of things not medical, not scientific but political, it has gone wrong.

I will cite a couple of instances where the reverse has happened. The first is trovafloxicin. Trovafloxicin was approved in December, 1997. It was I think the eighth

fluoroquinolone antibiotic. Prior to its approval, in a study of prostatitis, 140 men, 10 percent of them, had significantly elevated liver function tests. So, one could argue that perhaps it shouldn't have been approved because none of the other fluoroquinolones had anything like that. Shortly after it came on the market, an update safety memo from FDA said it became apparent that trovafloxicin had the most reported liver-associated adverse events--these are postmarketing events now, in addition to the randomized trial I described--within the first six months after approval of any of the approved and on the market quinolone antibiotics. This serves to illustrate the magnitude of trovafloxicin-related hepatotoxicity as compared to other approved fluoroquinolones.

We filed a petition to ban the drug in 1999, at which time there were eight cases of liver failure, five deaths and three liver transplants. It was taken off the market fairly promptly around that time in Europe and in Canada, but the FDA

decided to leave it on the market and "limit its used to just hospitalized patients or patients in nursing homes."

Shortly before we filed our petition I called Dr. Jerry Mandel, someone I have known for a long time, the author of "Principles and Practice of Infectious Disease," and I said, what would be the downside, if any, if trovafloxicin were taken off the market? He said, absolutely none. There is no unique advantage at all. So, a drug with no unique advantage but unique risk--why is it still on the market? It was left on the market by the FDA and by the end of last year, even though the company quietly "discontinued" the use there were 18,000 prescriptions filled last year. By the end of last year there were 58 cases of liver failure, including 29 deaths and 9 people requiring liver transplants. So, at the very least, shortly, within months after it came on the market, the FDA should have done something such as taking it off the market. No unique benefit, unique risk should mean off the market.

I am just briefly going to go over another example, which is Geodon, another atypical antipsychotic drug. Its initial approval was held

up because it had significantly elevated QTc intervals compared with other drugs. The company, Pfizer, was asked to do a randomized study and look at it in context of a number of other approved antipsychotic drugs, and it turned out to have, other than thioridazine, Mellaril, the longest QT interval. No unique advantage in terms of efficacy; it has some advantage in terms of not putting on weight for some people but a lot of people don't put on weight on the other drugs.

One of the handouts for this meeting was a position paper by FDA on how do you handle QTc prolongation. Clearly, this drug has more QTc prolongation than a number of other drugs, yet it doesn't even have a black box warning. Thioridazine does have a black box warning. This drug, Geodon, does not.

We just took a look as of now, as of the end of last year, at how many adverse reactions, as

ventricular arrhythmias, had occurred with this drug, and in the three-plus years it has been on the market there were 33 reports of ventricular arrhythmia, of which 26 were ventricular fibrillation or ventricular tachycardia. We don't have prescribing data on this drug but we know that there are about six times more prescriptions for another antipsychotic, olanzapine, which also has some of these cases but, when adjusted for what I believe will be the prescribing data on both drugs, it will turn out that the rate of ventricular arrhythmias, tachycardia and fibrillation, is much, much higher.

In summary, I would just like to urge all of you to request from the FDA and evaluate on your own more case examples of what goes wrong. I think it is important to present, as is being done at this meeting, the process for identifying things but after they are identified, if nothing is done about them at all or promptly or during the process of approval, if something is found and not acted on, then the process is seriously flawed.

And, I think there were not enough autopsies done. Back in 1945 about 40 percent of the people in this country who died had an autopsy.

It is down to about 8 percent. The FDA has not, to my knowledge--certainly it hasn't been made public--made a series of autopsies on the serious mistakes that have been made in approving drugs and not promptly taking them off the market. I think this advisory committee would do well to make recommendations and become involved in this process.

Dr. Furberg and Dr. Sady did such autopsy or postmortem on Baycol using some company data as well as the FDA data, and even if you didn't have access to company data but just FDA numbers, and so forth, I think it we be very enlightening and we could really improve the outcome. Thank you.

Northwestern University School of Medicine

DR. BENNETT: Charles Bennett, hematologist-oncologist at Northwestern University and the VA. No conflicts of interest to acknowledge.

I want to report today on the RADAR project that was published last week in JAMA, research and adverse drug events and reports. It is our independent work that we are doing to supplement some of the work done by the FDA.

We use a limited definition for adverse

drug reactions. We look at things that result in death, severe organ failure or requiring intubation, cardioversion, blood transfusion or organ transplant so we are looking at what we think is the tip of the iceberg of very, very serious adverse events in the RADAR program.

Reasons for the RADAR, clearly, Sidney Wolfe's paper in JAMA years ago said that clearly it takes seven more years before adverse drug reactions will be described in black boxes. Lots of clinical trials are just too small. We know that. And the adverse event reporting systems--we have heard the limitations.

So, what we do in RADAR? What is different? We evaluate initial reports or previously unrecognized but serious adverse drug

reactions. Where do we get our reports from? I will give you a little schematic. Basically, because we are clinically located, we get reports often from the academic environment which we work in. We get them from academic collaborators. We get them from people calling us up now; we get them from lawyers questioning us about issues and associations. We get them almost anywhere we can find these initial reports.

After we get the reports we question the FDA databases through the MedWatch program. We used to have a little bit more easier access to the adverse reports. As Sydney mentioned, they are available on-line. We used to get a little bit more on-line in the past. It is a little bit slower now but it is still pretty good.

We develop hypotheses for mechanistic pathways, and that is an important part of it. Rather than the statistical reporting methods you heard up front, it is very important, we think, to have hypothesis-driven mechanistic pathways up front to have an idea of where you are going. If

you don't know where you are going I don't know if you can find out where you want to be.

We evaluate a lot of laboratory and pathologic findings and we understand that at the end the last step is almost controversial for us and it is our most difficult, we have some method of reporting incidence rate estimates but it is the weakest part of our program. We said it up front, again, at the end of the day we are not going to be able to give you confidence about reporting rates and incidence rates. We give you confidence about the signals we detect.

Where does that money come from? None of that money comes from pharmaceutical support. Our money comes from National Heart, Lung and Blood, National Cancer Institute, Veterans Affairs, and we do not accept pharmaceutical support for our work.

We have 25 core investigators. We are lucky, we are one of the NCI-designated competency cancer centers. There are 41 in the United States. A competency cancer center from NCI is a place we are identifying as a particular place of expertise

and should be considered as a collaboration with the FDA. What is specific about a competency cancer center? We have funded by the NCI core technologies that relate to geriatrics, cardiology, neurology, dermatology, hematology, oncology, pharmacology, epi., stats., and pharmacy. These are people on the budgets of the competency cancer center. We don't have to pay for full FTEs of these people. There are 25 FTE type people so if you put those internally into the FDA there would be a significant budget implication. For us, we buy 5-10 percent of each of these individuals through the cancer center.

We have weekly conference calls. We have members in Chicago and we also have affiliated members in Utah, Albuquerque, around the country and now overseas. We have meeting minutes and agendas which circulate prior to the calls, and we have afterwards interviews and follow-up meetings as well.

How do we disseminate our information? We submit our articles to medical journals. If the

article was outstanding, like the one we have pure red cell aplasia with erythropoietin we are lucky enough to get it into Erythropoietin. If the article was not so outstanding we usually get rejected by the journals because it usually has conflicts of interest in many cases so we have been rejected seven, eight times. We revise package inserts. We work collaboratively with the FDA. We send material down to the FDA. We ask for package inserts, "dear doctor" letters and sometimes we get them and sometimes we don't. We present our data at national medical conferences. We meet with the FDA. We have been fortunate enough to meet with Anne Trontell and others and it has been very, very helpful for us to go on our own to meet with the FDA. We present our data on pharmacovigilance programs. We share with them our data. We present to them our manuscripts oftentimes before we submit them to the journal and we ask for comments.

How does the flow go? You probably won't be able to see this but the flow is that we have a signal generated, which is that anybody in the

country can bring a signal to us and say we have an interesting adverse event, is it possibly association? We work with that. We have a signal generated by our core team. We say is it really unique? Is it novel? Has it been looked at before? If it is a "go" we move forward and look for MedWatch follow-up data.

If we get some more follow-up from MedWatch we generate a case report form. Once we get the case report form we look for ancillary sources of data. This can be clinical trials; surveys of physicians; clinical trial reports from various cooperative groups; literature searches; and sometimes we have to go out on our own to generate our own data sets. After all that, we generate a report. We report it back to the FDA. We report to the pharmaceutical companies and we attempt to publish our findings.

Reporting rates, as I mentioned, is very difficult. We understand that and we heard the presentation. Total number of users, clearly, that is a very gross exaggeration. Every paper we have,

every review we get says the reporting rate we have is wrong and, therefore, your paper is probably not going to get through the system. We understand that up front but if we tell you it is one in ten percent, if it is ten percent we may be wrong. It might be five percent but it is not one in a million. And, we get a lot of flack on those incidence estimates. We tell you it is one in a million; they say, no, it is 1/100,000. We understand we might be correct but we can give you a ball park.

Here is what we have done to date. We have done 17 drugs. The majority of our material comes from the FDA, which is in this first column. We get a lot of reports from the MedWatch database. We have had great access to the FDA database. It has been very important to us. But we don't stop there. We look at publications. Our RADAR group, because we are so well connected around the country, we have been using a lot of RADAR sources. More recently, we are getting access to data from attorneys. Attorneys have given us good support in

terms of some cases that they have looked at. We also look at patients, and we ask doctors and, finally, clinical trial reports. Clinical trial reports, as you know, are going to be open more and more. We find that Phase I and Phase II clinical trial reports are very helpful. The Phase III clinical trial reports are going to be out there. We are asking that Phase I and Phase II should be out there as well.

What do we do with our findings? We found, again, in many of these findings that percent in a database can range from zero percent in the FDA database to 100 percent in the FDA database. So, it can be a wide range, from 0-100 percent. The ones that are zero percent, how do we find those? For instance, we found these three cases from a lawyer related to a drug that never made it to market. But the 360 volunteers that took the drug for $5,000 a piece, 13 of them have persistent thrombocytopenia and antibodies. Three of those patients have lymphomas five years out from the trial. We haven't been able to generate a

look-back study for those 360 volunteers. The question raised is if three of those people have lymphoma, do any of the other 357 patients have lymphoma after being in the trial?

We have looked at Lovenox. At our hospital alone we had a significant amount of bleeds after patients got Lovenox following cardiac procedures. We raised a concern that this is an issue around the country. It took a while before we could get it out. We haven't been able to put it into the warning label for a variety of reasons.

For instance, nevirapine--nevirapine is a controversial drug, very important drug. Our nurse took nevirapine after needle stick and became sick as a dog. Liver function went up to five-fold abnormalities. We called around the city and asked if anybody else had a nurse taking nevirapine after needle stick. In fact, we found a phlebotomist at Illinois Masonic. That phlebotomist required a liver transplant. Fortunately, somebody had a car crash and was able to donate the woman a liver and she was able to survive. It is unbelievable that

we found six people in the City of Chicago after needle sticks, and they were taking nevirapine off-label for needle stick exposures.

Plavix--we reported as soon as the drug came on the market that the drug was going to have TTP. We said it is going to happen because it happened with Tyclid. We went out actively and requested every plasmapheresis center to look for Plavix cases and, in fact, we found them.

We have been able to get "dear doctor" letters. We have been able to get it in the package inserts and black box warnings, precautions and adverse events in the FDA labels.

What are some strengths and weaknesses? Clinical trial reports are very helpful. We like them. They are very complete. However, they have only a small number of reports. You heard the limitations of the MedWatch database. Physician queries, they are very complete for us because I have never had a physician say no to me yet. They may say no to some other people but for us, we have been very persistent. And, the pharmaceutical

company, if you work pretty hard you may get some information.

Reporting rates--clearly, we have a problem but if it is frequent, we tell you it is frequent; if it is rare, we tell you it is rare. We can't tell you how rare or how frequent in many cases.

The legal system has been very helpful and the State Attorney General, Richard Blumenthal, has filed a citizen petition with the FDA last week, saying that one of the adverse events we identified should be included in the package insert.

Finally, what refinements do we recommend? The MedWatch is accessible on-line. We have had better access in the past and we would like to get back some of that access that we had and, finally, systems to address prospectively identified persons with ADRs. We want to look at those. Particularly, we would like to look for TTP cases and, finally, update the programs in place. The STEPS program, developed by Allen Mitchell, we would like to extend it to other side effects. And

we would like to collaborate with the NCI and the FDA.

DR. GROSS: Thank you very much, Dr. Bennett. The next speaker is speaker number four.

Ingenix

DR. WALKER: Dr. Gross, members of the committee, I am Alex Walker. I work for i3, which is a part of Ingenix which is, in turn, part of United Health Group. In the division that I supervise about 60 percent of our business consists of postmarketing surveillance studies mandated by the FDA and paid for by the various pharmaceutical manufacturers. About a quarter of our business is working with United Health Care, a sister company on analysis of drug issues for that managed care organization.

I would like to tell the committee about an initiative that we are undertaking to look at drug safety issues within United Health Care database. The need that we are trying to respond to is a routine and comprehensive system for drug safety evaluation that looks at "all drugs and all

outcomes"--I have those in quotes because it is not possible literally to look at everything, but that is the goal--capable of generating signals which could be verified elsewhere or tested internally, and it should be capable of testing signals arising from other sources. The goal of this activity is to serve all the stakeholders in the drug marketplace from patients and doctors through payers, managed care, regulators and manufacturers.

We are actually doing--we are not proposing, we have already started work on an active drug safety surveillance program using the full data assets of United Health group and principally within United Health Care, within which we now have 11 million people whose insurance information can be retrieved, and we are trying to change our system so that we can push close and closer to real-time surveillance.

The plan of the activity is as follows: For every new molecular entity which is introduced we will bring it into follow-up where there is a critical number of users. We have identified about

a thousand users within United Health Care data as a benchmark for that for statistical reasons. I mentioned the data source. This is important in that it is an open formulary, which is to say new drugs come into United Health Care Data very quickly and we can get experience early. In addition to the 11 million that I have mentioned, United Health Care and related companies within United Health Care group are growing very rapidly. You may have heard of recent acquisitions of Oxford and MAMSI, and Americhoice--Americhoice is important because it is a Medicaid provider--that are working their way to get included. I haven't included those numbers in the 11 million.

In thinking about what we are proposing, I want you to bear in mind that these are insurance claims data, no better; no worse, although I should say that through our work we have had a lot of experience in sorting, thinning and extracting medical data from those, and one of the main parts of our research activities is verifying claims with the underlying written medical records so that we

have experience in that.

The program will involve what you might think of as three filters which help clear the huge mass of data that comes out of very large insurance claims databases. We will be looking at treatment emergent diagnoses, meaning diagnoses that have occurred for the first time after a drug has been started. We will be looking at comparator groups. We will be using statistical methods that we have been using over the past several years to create comparison groups that are highly similar to the users of the new molecular entity with respect to demographics, medical history, healthcare utilization, concomitant drugs, and the like. We will be using data mining techniques in this, sort of more refined and symmetrical kind of data universe that really amounts to looking for interactions between potential adverse effects and concomitant drugs, diagnoses, and the like

All of these filters have tradeoffs and I will describe them in a little bit more detail, but the advantage in applying filters to insurance data

is that there is a tremendous amount of noise and you have to somehow systematically remove it or you spend your life setting noise and that gives the differences that remain some greater impact to interpretability. The disadvantage is that there are clearly adverse effects of drugs that will get screened out by these filtering techniques. So, I am not proposing that this is going to be a universal solution to drug safety monitoring, but simply an addition to the process.

Let me just go quickly through each of those filters. We will be looking at hospitalizations, doctor visits and other services that occur after the dispensing of the new drug or the comparator, not sharing the first three digits of the ICD coding scheme with any service in the six preceding months. The idea here is to take advantage of the ICD scheme. That is how diseases are identified in our data that is not MedDRA, and to use that to filter out progression of the disease and concomitant illnesses that are simply continuing. Clearly, when the adverse effect

relates to progression of disease or progression of concomitant illness you don't get it.

The identification of comparator drugs goes to some of the issues that new things that happen when people are taking drugs aren't all due to the drug that is being taken. What we are doing here is to choose with each new drug that comes in a drug used for a similar indication, typically the most prescribed drug for that indication, and then choosing users who have similar diagnoses, healthcare utilization, claims history and a whole variety of aspects. The purpose of the comparator group then, just as in a clinical trial, is to give a standard by which to judge things that occur in the general population.

The general pattern then of surveillance within this is to, on a quarterly basis, identify all people receiving the drug under surveillance and to select the control group or the comparison group with these statistical techniques, initiate follow-up for that cohort and then again repeatedly look at each of these cohorts.

The data mining that we are proposing is actually considerably less complex than you heard from Dr. Gould. It really amounts to a group for

subgroup review of every possible subgroup data defined by concomitant medications, diagnoses and demographics to look for subgroups within which differences between the drug and the comparator are more striking than they are on the marginal tables.

The plan that we have is to begin to generate quarterly reports on the new molecular entities that have come in, which will include mostly what you might think of as sort of routine reports, by comparing treatment patterns, treatment emergent diagnoses. There will be ability to do some querying of the data and data feed. We plan to produce annual print and web-based summaries which will be widely available to summarize the information.

We will have a beta version of this available by the end of August. We have picked our four NMEs from the drugs that within the United Health Care data have the largest number of

individuals receiving the drug among the NMEs introduced in fiscal 2004. Those are Cialis, Cymbalta, Sprivia and Ketek, and the comparators that we have chosen for each of these are Viagra, Effexor, Atrovent and Biaxin.

Just to give you a sense of numbers, for the Cialis and Ketek we have well in excess of 30,000 people in the drug group and similarly in the comparison group; smaller numbers in the Cymbalta and Sprivia. Don't let that threshold of 1,000 to get in lead you to think that we are dealing typically with a small number sort of situation. Thank you very much.

DR. GROSS: Thank you, Dr. Walker. We will now hear from our last registered speaker, speaker number five.

Boston Collaborative Drug Safety Program

DR. JICK: Good morning. My name is Susan Jick. I work with the Boston Collaborative Drug Safety Program. We are part of Boston University Medical Center.

I would just like to start by mentioning

that we have submitted a background document to the committee and this provides a reasonably detailed review of the history and accomplishments in the field of drug safety epidemiology over the past 40 years. It is our view that a great deal is known and published in this area. This body of information is available in publications to anyone who is interested and wants to take the time to review the evidence. We hope the members of the committee will carefully review this background document that we have provided.

In the short time that I have, I would like to mention our views of the most crucial elements that could lead to an informed judgment of the current and future status of drug safety epidemiology.

First, I would like to say that, as I am sure people are well aware and as has been mentioned before, drug safety is a complex and subtle area of medical research, as complex as, say, molecular biology and it logically follows that in order to become a professional level

scientist in this field, one requires years of intense training and a long time spent in direct experience conducting research in this area.

Secondly, drug safety epidemiology is a unique area of epidemiology whose principles and methods are as different from, say, chronic disease or infectious disease epidemiology as surgery is from psychiatry or pediatrics. Therefore, a postgraduate degree in traditional epidemiology does not itself provide a direct basis for being informed about the substantive issues related to drug safety epidemiology.

The ability to produce high level, quality research in this area requires, as noted, appropriate training and research experience, and the necessary tools to conduct this kind of research, namely high quality, reasonably complete information on relevant clinical medical history in a standardized nature and in a very large number of people.

Having made those comments, let me say that, again, there is a large body of information

out there in the drug safety area, going back about 40 years. Prior to 1966 there were no formal epidemiological studies. There were none in the area of clinical drug safety. The information and the knowledge in this area was derived really from animal research data and from personal opinions based on clinical experience.

The first continuous large-scale multipurpose formal study on drug adverse effects was begun in 1966 by the Boston Collaborative Drug Safety Program, and this was funded at that time by NIH. The study was based on observational information. The original objective was to document and quantify the acute, that is, short-term toxicity of osteomyelitis marketed drugs at that time. The design used was restricted to the study of hospitalization patients where drug exposure at that time could be fully recorded and carefully followed up.

Over the years, the study design was introduced into some 40 hospitals in seven countries, and by 1982 the available information

encompassed about 70,000 patients and 700,000 drug exposure episodes. This design provided short-term follow-up safety information for large cohorts of users for all drugs that were used in those study populations.

During the time that the study was going on, the BCDSP introduced an additional data collection element that specifically allowed for evaluating the risk of hospitalization for certain illnesses caused or prevented by drugs that were taken on an outpatient basis prior to hospitalization. The evaluation utilized the so-called case control design. Numerous drug-disease relations were identified using this technique including, among many others, the negative association between aspirin and MI. This was published in 1974 in the BMJ, British Medical Journal. Also, the positive association between estrogen use and gallbladder disease was discovered and this was published in 1974 in the New England Journal of Medicine. I would also like to mention that both of these findings have been confirmed by

clinical trials since that time.

For those of us who were engaged in full-time activity in the drug safety field, starting in the mid-'60s, it had become evident that far greater efficiency was required to conduct the necessary research for the hundreds of marketed drugs that had become available. The increasing availability of computer-recorded medical information offered, at least in principle, the opportunity to achieve a major advance in the efficiency of obtaining the required information to conduct these safety studies.

To this end, in 1978 the BCDSP developed a cooperative agreement with Group Health Cooperative of Puget Sound, a health maintenance organization out in Seattle, Washington. The HMO, which at that time had a membership of about 300,000 people, had begun to put diagnosis information on all member hospitalizations on their computers in 1972. In addition, they computerized all of their local pharmacies and that was completed in 1976. Finally, they had centralized record rooms. It was

possible to get the actual clinical records on members of the cooperative, and this was very important for validating the data resource.

Subsequently, early studies done by the BCDSP using this data resource documented the high quality and completeness of the computer-recorded information on drugs dispensed in the cooperative and hospitalizations that occurred.

In view of the expense and administrative tediousness of the previous means of conducting drug safety studies, it was immediately evident that the availability of this high quality computerized data represented a major advance in the ability to conduct drug safety studies. The research which, by 1995, encompassed 20 years of follow-up information, proved to be highly useful. Over 50 drug safety papers based on Group Health Cooperative were published in peer reviewed journals.

By 1991, the BCDSP had published more than 35 articles related to drug safety in JAMA, more than 20 in The Lancet and more than a dozen

articles in the New England Journal of Medicine. One could argue that these papers formed the foundation of a large part of our quantitative knowledge of drug safety for many of the drugs that were on the market in 1991.

While the initial use of computerized medical information, which began in the late 1970s, had provided new, highly efficient research output in the area of drug safety, the available resources did have limitations based on certain characteristics of the resources and certain types of studies could not be conducted.

In considering where one might go to generate a larger and more efficient source of data n directly accessible medical information, it has become apparent to us for many years that the U.K. provided a unique medical environment to create an optimum computerized medical data resource. The characteristics of the U.K. medical system are such that the GP is the gatekeeper for all patient care so that the characteristics of all of the patients are known or are kept with the GP's office. A

comprehensive record of prescriptions written, outpatient diagnoses and referral letters to hospital are all available in GP files.

In late 1980s, VAMP Health, a commercial company, designed and marketed a GP computing system which recorded comprehensively the medical records for individual patients. This system enabled computer recording of patient demographics, all prescriptions, all important medical advances, along with a considerable amount of other important medical information. Numerous studies have repeatedly confirmed the quality and completeness of the computer-recorded diagnoses in this system and found the quality to be very high.

After completion of the many validation studies using this data resource, it was possible to conclude that the database could be relied upon to provide efficient access to clinical information suitable for drug safety studies, and that GPs were cooperating and providing photocopied referral letters which were necessary to validate the diagnoses in the database. It was, thus, possible

to document the size of large cohorts of drug users, and these unprecedented cohort sizes were then provided and made available for drug safety studies for quantification of many events, including rare events.

The United Kingdom's Office for National Statistics assume the responsibility of collecting and maintaining the VAMP database in 1995. The database is now known as the General Practice Research Database, or the GPRD. This database, in 2000, was taken over by what was then called the Medicines Control Agency and has been in there since that time.

The BCDSP has now published over 150 papers based on research conducted on the GPRD. The research output has fully demonstrated the unique utility of the resource to provide comprehensive, well-documented study results on a large number of safety issues, and relatively quickly, at relatively low cost. Our references are available on our web site.

Our studies published to date demonstrate

that the GPRD is capable of documenting and quantifying important suspected toxicity population-based database, and we can calculate rates. The database is capable of providing substantial evidence of drug safety for commonly used drugs, and providing reasonably precise drug-specific quantification of recognized toxicity.

Additional evidence supporting the validity of the GPRD comes from published studies that provided results highly concordant with those of studies published by other investigators. Examples of some of these would be studies looking at the association of oral contraceptives and venous thromboembolism, NSAIDs and GI bleed, HRT and venous thromboembolism, MMI vaccine and autism, to name just a few of those.

This brief history of the development of the research in the drug safety area, conducting formal studies in postmarketing, I hope gives a sense of the availability of the information that is out there. There are also many other resources

available out there at this time and other studies being conducted.

Let me just add some comments about the interpretation of observation of drug safety studies, now that I have talked about the available resource, all drug studies based on observational data are subject to biases that are inherent in non-experimental research. The ubiquitous presence of such biases are difficult, if not impossible, to control require a research design meticulously adhering to fundamental epidemiological principles just to minimize the effects of bias on the interpretation of the results. At the same time, where possible selection or information biases are likely to be unimportant to the interpretation of a particular study they should not be invoked indiscriminately to question the conclusions drawn from well-designed observational studies, which provide convincing evidence of causal relations of a drug to a particular illness.

Unfortunately, it is still possible to find comments such observational studies are

weakened by biases and confounding factors and cannot be relied on. While this may apply to some studies, certainly it does not apply to all observational studies in this area.

There are many examples where results of observational studies have been confirmed by randomized trials. The BCDSP in 1974, as I mentioned, reported a strong protective effect of aspirin on MI and this has since then been demonstrated in trials. Dr. Molensy et al. reported a strong negative association between folic acid intake in the first trimester of pregnancy and risk of neural tube defect. There have been observational studies reporting an increased risk for VTE among estrogen users. All of these findings have been confirmed by subsequent randomized trials, as have many others.

In summary, it is our view that the drug safety epidemiology is a complex and subtle science, and as complex as many other areas of medical research. It is therefore, necessary for persons who conduct such research to have training

and experience to develop the expertise at a high professional level.

Few people who publish in this area have this training and experience and publication of unsatisfactory research is ubiquitous and contributes to the sometimes ugly public controversy in the drug safety area. Unfortunately, those who have responsibility in this area, in industry, in regulatory agencies, in journalism both medical and public...

[Microphone is turned off]

DR. GROSS: On that note, Dr. Jick, thank you very much. We have a few minutes. Is there anyone in the audience who would like to add a comment at this open public hearing? Seeing no hands raised, I would ask those around the table if they have any questions of any of the speakers who presented in the open public hearing. Ruth Day?

DR. DAY: I have a question for Dr. Bennett and the RADAR system. He presented a very nice investigational flow diagram for ADRs and it makes a lot of sense. It all starts with "get a

signal." Can you comment on the types of things you consider a signal? Because we did hear of an interesting case of one needle stick that then led to eight within a city, and so on. So, what signals that you have a signal you ought to act on?

DR. BENNETT: Thank you, a very good question. The signals we look for would be very serious side effects, that is, severe hepatotoxicity. The Lovenox bleed was big inguinal bleed after a cardiac cath. procedure where a person ends up on a ventilator after taking the drug, a person who ends up on plasmapheresis in ICU after taking Plavix. Those kinds of signals are very, very strong, very apparent. So it is because we are so clinically attuned, and we are also in the middle of a hospital setting where we could find them. So, if we were out doing data mining this might not show up but it is really because we are in the middle of a clinical setting that we see some severe side effects.

DR. GROSS: Thank you, Dr. Bennett. Jackie Gardner?

DR. GARDNER: I have a question for Dr. Walker. Dr. Walker, beyond the screening, can you tell me whether you are succeeding in making some

headway in access to medical efforts for follow-up in that system that you are working in now?

DR. WALKER: Our research projects involve medical record review. That is still under IRB approval and under all the usual constraints to assure privacy. I don't see it as a routine procedure, although I suppose if FDA were wanting it, it could be figured out.

DR. GARDNER: So, it is available--

DR. WALKER: Absolutely, yes.

DR. GROSS: Thank you all very much. We will now break for lunch and we will begin promptly at 1:00 p.m. Thank you.

[Whereupon, at 11:55 a.m., the proceedings were adjourned for lunch, to reconvene at 1:10 p.m.]

- - -

A F T E R N O O N P R O C E E D I N G S

DR. GROSS: I would like to welcome Dr. Mary Willy, epidemiology team leader in the Office of Drug Safety, who will talk about active surveillance for drug safety signals: past, present and future.

Active Surveillance for Drug Safety Signals: Past, Present and Future

DR. WILLY: This afternoon I am going to provide you with a brief overview of the use of active surveillance for drug safety.

[Slide]

My talk will include a brief background and history of active surveillance, and then provide some examples of different surveillance programs here, in the U.S. and elsewhere. I will discuss the challenges associated with trying to do active surveillance, and conclude with a discussion of possible U.S. applications.

Following my talk, there will be two examples of active surveillance systems being piloted by FDA. Dr. Mendelsohn will discuss the

National Electronic Injury Surveillance System: Cooperative Drug Event Surveillance Symptom. Dr. Graham will discuss as second pilot program that involves active surveillance of longitudinal data.

[Slide]

This is our postmarketing surveillance program which you have seen before today. As you see, it includes active surveillance in equal parts of the picture.

[Slide]

The Office of Drug Safety has a group of epidemiologists that work to study drug safety signals using case report series, recording rates, background rates, drug utilization data and literature. Because there are recognized problems with underreporting to AERS, other databases are used at times to try and explore possible drug safety signals, including claims databases, electronic medical record databases and national surveys.

[Slide]

In the future, the office would like to

collect additional drug safety information using active surveillance systems. Active surveillance might be defined in a number of ways but for the purposes of this presentation I will define active surveillance as the regular collection of case reports from healthcare providers or facilities. The focus for surveillance may be on outcomes, settings or drugs. An active surveillance system has the potential of collecting more complete information, although one system is unlikely to address all drug safety problems.

[Slide]

Active surveillance programs might be used in two ways. First, the program might be used to help identify drug safety signals, although the method used for how a signal might be defined is not clear. Second, a surveillance program might be used to collect additional cases to help validate drug safety signals that were identified through passive surveillance.

[Slide]

There have been different active

surveillance initiatives in the past that the FDA has participated in. In the 1960s there was a joint project with NIH and a large HMO to develop medical record linkage. This project attempted to link scanned forms from different parts of the health system, but failed because computer technology was still in its infancy.

[Slide]

In the 1970s there were efforts to collect drug exposure information from patients hospitalized with specified outcomes. But these programs did not meet the FDA's needs due to funding challenges, underdeveloped technology and the low yield of new information.

[Slide]

One way to think about how active surveillance systems might be described is according to the strategy used for doing surveillance. A drug-based system would follow large numbers of patients exposed to new molecular entities after their launch for all or for specified adverse events.

A setting-based system would be implemented in a relevant setting, such as a hospital or emergency department, and would work to

detect drug-related events likely to present there, for example, anaphylaxis in an emergency department.

A disease-based system would be some type of program that would be developed to collect comprehensive disease-specific information for selected drug-induced diseases.

[Slide]

I will spend the next few minutes talking about examples of surveillance systems that can provide drug-related information. The examples I am providing do not represent every system currently collecting drug safety data, but only a series of systems that use different strategies that might be relevant when thinking about developing some type of active surveillance system. In particular, I will mention two non-U.S. systems as examples of some of the foreign programs that have been initiated although their relevance to the

U.S. may be limited.

The Drug Abuse Warning Network, or DAWN, is funded by the Substance Abuse and Mental Health Services Administration. The program collects information from a nationally representative sample in emergency departments so it is an example of a setting-based surveillance system. A sample of medical examiners that are not nationally representative also provide information. The survey was recently revised and now includes 22 metropolitan areas. The population studied includes ages 6-97 years; collects information on any kind of drug-related event' and includes chart review of cases.

[Slide]

While DAWN provides nationally representative data that can be useful to monitor drug safety, it has recently been revised so it is difficult to study trends in drug safety.

[Slide]

Another setting-based example is the Toxic Exposure Surveillance System, or TESS. TESS is a

database established in 1983 and maintained by the American Association of Poison Control Centers. There are over 64 poison control centers in the U.S. which serve nearly the entire U.S. population. The centers provide information to callers, but they also collect data from collars about poisonings.

[Slide]

TESS has the advantage of collecting data from almost all the country on any drug, including over-the-counter drugs. Unfortunately, only limited data is available to the FDA and the information that is collected by the centers is not validated. Additionally, data on events may be missed if a call is not made about the event to the center.

[Slide]

The Acute Liver Failure Study Group is an example of a disease-based surveillance system. The program is funded by NIH and collects data from 25 adult and 25 pediatric sites. Patients hospitalized with severe hepatotoxicity are

enrolled in the study. A subset of patients with drug-related events can be identified from the study.

[Slide]

The Acute Liver Failure Study Group is an example of a system that provides detailed, validated information for patients. This program is not nationally representative and may miss cases that die before reaching an expert or are less severe and don't need an expert's attention.

[Slide]

An example of a program that uses drug-based surveillance is the United Kingdom Prescription Event Monitoring System, or PEM. The system started in 1980 and is funded mainly by unconditional grants from pharmaceuticals. In this program newly approved drugs identified as important are chosen for monitoring. Prescribers of the study drug are identified by the national prescription system. Six months after the first prescription for an indication is written, the prescribers are sent a questionnaire, which is

called a green card, and are asked to identify every significant event. An average cohort of patients studied in PEM is 10,000.

[Slide]

Although this system involves the majority of physicians in the country, the typical cohort size may not be large enough to capture rare events. They do not monitor hospitals or over-the-counter drugs, and they have a response rate of 58 percent. Most importantly for the U.S., this type of surveillance system relies on the identification of all prescriptions and prescribers, something that the U.S. does not have in place.

[Slide]

Another foreign program to mention is the French Pharmacovigilance System. The system was started in 1973 and was decentralized in 1979. It includes a network of 31 regional centers that are located in the departments of clinical pharmacology. The centers collect adverse event information and provide feedback to professionals,

and they also conduct research. The 31 centers are connected together by a national database, and funding is provided by the French Medicines Agency based on performance and scientific publications.

[Slide]

This system is an example of a program that collects information from centers that are distributed throughout the country, but it may not be representative of the general population since the centers are located in academic centers.

[Slide]

Well, there are challenges to every system that I have described that might be used for collecting active surveillance infection. First, it can be difficult to obtain timely information since data often needs to be cleaned up and then transferred to a central database. Validated information may be difficult to obtain. The ideal system would cover both inpatient and outpatient settings but such a system would be very costly. Identifying rare signals also requires access to information from a large population. Having a

system that is efficient at identifying true cases may be difficult to develop since programs may identify possible cases that, upon further investigation, are found to be non-cases. Finally, obtaining a broad enough scope across the U.S. is difficult.

[Slide]

I would like to spend a little time talking about how we might apply active surveillance in different situations. First, many of the adverse events that are identified by AERS are rare events, such as acute liver failure that has a background rate of 1/million person-years. Could active surveillance help identify these kinds of cases? Perhaps a disease-based program might help identify cases, But the challenge would be getting the providers to attribute the disease to the drug, otherwise the surveillance system would miss the cases.

[Slide]

What about an event with a high background rate, like acute myocardial infarction? The

background rate for acute myocardial infarction depends on the population you choose but has been reported to be 4/1,000 adults. Could active surveillance help identify a signal in this situation? A drug-based surveillance program might help collect information to help quantify and describe the association, but it would be important to have a comparator. [Slide]

Hospital-related events be difficult study, particularly those that are anesthesia related. Could active surveillance help identify these cases? A setting-based system might identify cases if an effort is made to monitor for the event of interest and prior drug exposure.

[Slide]

In conclusion, active surveillance is a complex process that might require multiple strategies. The current surveillance systems outside the FDA may provide useful information but are limited. Progress in computerized medicine will make the development of a timely active surveillance system more likely.

[Slide]

As we consider options for active surveillance, there are several questions. How

would active surveillance complement the passive surveillance system that is in place? Would active surveillance be any faster at finding signals? And, how would a signal be identified?

[Slide]

The agency remains very interested in developing some type of system. A request for information was announced in April to collect information about U.S. active surveillance programs. The Office of Drug Safety will continue to explore opportunities for active surveillance and will participate in any initiative to link health information that might prove helpful for active surveillance. Thank you.

DR. GROSS: Thank you, Dr. Willy. Dr. Aaron Mendelsohn, epidemiologist in the Office of Drug Safety, will talk about the National Electronic Injury Surveillance System.

NEISS-CADES--National Electronic Injury Surveillance System:
Cooperative Adverse Drug Events Surveillance System

DR. MENDELSOHN: Good afternoon.

[Slide]

My name is Aaron Mendelsohn, and I am an epidemiologist in the Office of Drug Safety at the

Food and Drug Administration. Today I am going to present to you a description of the National Electronic Injury Surveillance System: Cooperative Adverse Drug Events Surveillance Program, or NEISS-CADES. NEISS-CADES is an active surveillance system the FDA recently acquired for assessing adverse drug events in the outpatient setting.

[Slide]

NEISS-CADES has its roots in a system created over 30 years to. In 1971 the Consumer Product Safety Commission implemented the National Electronic Injury Surveillance System, or NEISS, for detecting injuries related to consumer products and presenting to a random sample of hospital emergency departments in the United States.

Over the years the NEISS system was

continually adapted. A significant milestone happened in the year 2000 when NEISS was expanded to collect data on all injuries, including but not limited to occupational and violence-related injuries and adverse drug events.

Just recently, in 2002 the FDA and the Centers for Disease Control and Prevention formed a collaboration with the Consumer Product Safety Commission to collect additional details on adverse drug events detected through NEISS, such as the route of administration and dose. I will describe all the data elements that are collected shortly. This new effort became known as the Cooperative Adverse Drug Events Surveillance System, or NEISS-CADES.

[Slide]

NEISS-CADES is ongoing survey of 64 U.S. hospitals. These hospitals constitute a stratified probability sample of U.S. healthcare facilities with 24-hour emergency departments and a minimum of six inpatient beds. The sites were selected based upon the geographic region, their size and whether

they primarily catered to a pediatric or adult patient population. As the sites were chosen to be representative of the entire U.S. and its territories, it is possible to make national projects with NEISS-CADES data.

[Slide]

This slide shows the participating hospitals in NEISS-CADES. Notice how the sites are distributed throughout the United States, with increased concentration in the more populous areas. Alaska and Hawaii are not shown on this map as there are currently no participating sites from either of these states. Hospitals from Alaska and Hawaii would certainly be eligible for inclusion in this surveillance system however.

The NEISS-CADES sample is adjusted periodically to reflect changes in the healthcare environment. Additionally, the analytic weights for making national projections are updated on a yearly basis.

[Slide]

NEISS-CADES captures adverse events for both prescription and over-the-counter drugs and includes topical preparations. The system also

detects safety concerns with vaccinations, along with the adverse events associated with alternative therapies, namely, vitamins and minerals, dietary supplements and herbal products.

[Slide]

NEISS-CADES defines an adverse drug event as an injury related to the outpatient use of a drug and resulting from one of the following mechanisms of injury: allergic reactions, in other words immunologically-mediated effects: side effects, defined as undesirable pharmacologic effects at recommended doses; unintentional overdoses, defined as toxic effects linked to excess dose or impaired excretion; and, finally, secondary effects such as falls or choking associated with the use of a drug.

Note that this definition excludes intentional cases of self-harm such as suicide attempts. In addition, injuries resulting from

alcohol use, tobacco products and illicit street drugs are not included in NEISS-CADES.

[Slide]

The key data elements from the FDA's AERS system were incorporated into data collection for NEISS-CADES, as shown on this slide. In addition to patient demographic data such as age and sex, NEISS-CADES obtains details on suspect drugs including the name of the drug, the drug's dose, frequency, duration of use and route of administration, and information on any concomitant medications that the patient was taking.

NEISS-CADES collects diagnosis information, along with data on diagnostic tests performed and treatments received in the emergency department. The patient's disposition following care in the ED is also collected. Finally, any other relevant details can be recorded in a brief narrative field.

[Slide]

Here we have a schematic showing the flow of data NEISS-CADES. A patient visits an emergency

department having experienced an adverse drug event. A physician diagnoses the event and documents the incident in the patient's medical chart. A coder, physically present at each hospital, performs daily reviews of all emergency department records. Upon detecting an adverse drug event, the coder abstracts the pertinent information from the patient chart. The coder then sends these data electronically to the Consumer Product Safety Commission. At the CPSC quality assurance personnel review the data. They check the information for completeness and they perform checks. Should there be any problem with the data, the QA staff go back to the coder for resolution. A second, and more thorough level, of quality control is then performed by the CDC. Once the data are cleaned, they are assigned codes using MedDRA, an internationally recognized system for classifying adverse drug events.

Data that have been MedDRA coded are available for analysis which is conducted jointly by the FDA and the CDC. The ultimate goal is to

use the findings to develop targeted interventions for preventing future adverse drug events.

[Slide]

Recently a team of researchers from the FDA, the CDC and the Consumer Products Safety Commission conducted a pilot study to describe the adverse drug event data obtains from a stratified convenience sample of 9 of the 64 NEISS-CADES sites. This pilot study was based upon cases detected in the first quarter data collection with NEISS-CADES, that is, between July, 2002 through September of that year. I will highlight the key findings from this study in the next several slides. Members of the advisory committee will find a copy of the final study report, which was published this year in Annals of Emergency Medicine, included in the meeting's background packet.

[Slide]

A total of 598 patients experiencing adverse drug events were captured by the 9 pilot study sites during the 3-month study period.

Adverse drug events among patients of all age groups were detected. The median age was 41 years, with a range from 0-101 years. One half of the cases were either pediatric patients or persons above the age of 62 years. Nearly two-thirds of the adverse drug events, 64 percent, were among female patients. Most of the patients, 90 percent, were treated in the emergency department and subsequently released. Roughly 9 percent of the adverse drug events resulted in inpatient hospitalization. None of the patients was reported to have died in the emergency department.

[Slide]

This slide shows the most common drug classes associated with the adverse drug events in the pilot study. Approximately 16 percent of the ADEs were related to antimicrobials, followed by diabetic agents at 13 percent, and cardiovascular agents at 9 percent. Though anticoagulants were associated with only 5 percent of the ADEs, they were responsible for 15 percent of the adverse drug event associated hospitalizations. Cardiovascular

and diabetic agents were responsible for 23 percent and 17 percent respectively of the hospitalizations.

[Slide]

Forty-four percent of the adverse drug events in the pilot study were due to unintentional overdoses. Side effects and allergic reactions were associated with 31 percent and 26 percent of adverse drug events respectively. Unintentional overdoses were responsible for nearly three-quarters of the hospitalizations, followed by drug side effects at only 15 percent and allergic reactions at 8 percent.

[Slide]

Permit me now to illustrate a few specific examples of the types of adverse drug events that have been detected through NEISS-CADES. A 68-year old male with gastrointestinal bleeding following warfarin use was held for observation. This would be classified as an unintentional overdose.

A 54-year old female became hypoglycemic following an overdose of insulin. The patient was

treated and released. This is also an unintentional overdose.

A 7-year old female had a rash following the use of an antibiotic, designated antibiotic A for the purposes of this presentation. The patient was treated and released. This is an example of an allergic reaction.

Finally, a 2-year old male with tremors following albuterol was treated and released. This is an example of a side effect.

[Slide]

The FDA and CDC are currently in the process of analyzing the data from the first 12 months of data collections with NEISS-CADES. The specific goals of this one-year analysis are, first, to obtain national estimates of various types of adverse drug events; second, to identify the drugs and drug classes that are most often cited as being associated with adverse events. To the extent possible, we will try to incorporate the nominator data in this analysis. For example, we will consider the number of prescriptions dispensed

for a given drug or drug class. Third, to identify the most common mechanisms of injury such as unintentional overdoses. Such information would be useful in targeted interventions for reducing the incidence of adverse drug events. Fourth, conduct multivariable analyses to determine which factors, including patient and drug variables, are independently associated with adverse drug event-related hospitalizations.

In addition to the one-year analyses, the CDC and FDA team also plan to study specific subgroups which have received limited attention in the past. This would include such persons as the elderly and the pediatric patients. Both of these groups seem to be adequately represented in NEISS-CADES.

[Slide]

NEISS-CADES is not without its limitations. It only captures certain types of adverse drug events. When NEISS-CADES is able to detect acute events, it will likely miss long-term negative consequences of a drug. For example,

cardiovascular events associated with the use of COX-2 products are probably underestimated in NEISS-CADES. In addition, the NEISS-CADES system only detects outpatient adverse drug events that are seen in the emergency department setting. Thus, inpatient events and those not necessitating an emergency department visit, such as less severe events, will not be captured. Finally, adverse drug events must be recognized by the emergency department physicians and be documented in the patient chart.

As the data come solely from the emergency department records, NEISS-CADES is dependent upon their quality and completeness. Another limitation of NEISS-CADES is that the data coders need to be well trained to adequately perform chart reviews and to detect adverse drug events. The principal stakeholders, therefore, conduct extensive training programs for the coders, including administrative practice cases and periodic continuing education sessions.

[Slide]

There are a number of features about NEISS-CADES that make this system attractive to the FDA. NEISS-CADES is one of the only nationally

representative active surveillance systems for detecting adverse drug events in the outpatient setting. As noted, national projections are possible and are currently being calculated for the first year of data collection. Using these nationally projected estimates, we can quantify the magnitude of a drug safety concern. This is something that is not easily possible with the passive AERS system.

Additionally, unlike AERS, NEISS-CADES does not differentially obtain data for newly recognized versus well established adverse drug events. To clarify, NEISS-CADES look like data for expected events such as bleeding related to warfarin and hypoglycemia associated with insulin, as well as being theoretically able to collect data regarding the unexpected serious adverse events that require expedited reporting to AERS.

Another strength of NEISS-CADES is that

the system is stable. As mentioned earlier, its parent system was created over three decades ago and has a well-developed infrastructure. Perhaps one of the reasons for the stability of the NEISS family of surveillance systems relates to the fact that these systems were designed to be easily adaptable. It is very simple to modify NEISS-CADES to address the ever-changing needs and increasing demands of this stakeholders.

[Slide]

As mentioned previously, NEISS-CADES collects detailed information on adverse drug events. Specifically, the system seeks to record the primary data elements from the MedWatch form. As most of these data are collected for routine medical purposes and are, therefore, readily available in the patient medical chart, few data are missing in NEISS-CADES.

NEISS-CADES is a timely system. Approximately 70 percent of the adverse drug events are available within one week of the emergency department visit. In the case of an urgent,

time-sensitive issue it would be possible to have adverse drug event cases available in real time, although these cases would not have been evaluated for quality purposes, nor would they have been MedDRA coded. Finally, NEISS-CADES represents a successful, cost-efficient collaboration between multiple federal agencies.

[Slide]

I would like to acknowledge the primary contributors to the surveillance system. I specially would like to recognize my colleague, Dr. Daniel Budnitz from the CDC, who has been primarily responsible for the success thus far of NEISS-CADES. Thank you.

DR. GROSS: Thank you, Dr. Mendelsohn. Before going on, I would like to acknowledge Dr. Curt Furberg. Curt, if you would go through the routine that we have done before and introduce yourself and tell us your area of interest.

DR. FURBERG: I am Curt Furberg. I am from Wake Forest University, Professor of Public Health Sciences. My interest is primarily in the

area of cardiovascular drugs.

DR. GROSS: Thank you, Curt. I appreciate you taking a "red eye" to get here. The next speaker is Dr. David Graham, a medical officer. He will talk about active surveillance using longitudinal data: a pilot project. David is with the Office of Drug Safety.

Active Surveillance Using Longitudinal Data: A Pilot Project

DR. GRAHAM: Good afternoon.

[Slide]

Today I will talk about an active surveillance pilot project using longitudinal data. What I will be discussing is really based upon the work of a number of other people, and I have only made a small contribution. So, those people should be acknowledged--Dr. Richard Platt, Dr. Arnold Chan, Dr. Martin Coldorf and Dr. Robert Davis. They have been the driving force behind this.

[Slide]

As a background, there are multiple potential approaches to active surveillance, and

Dr. Willy has talked about these previously. There are a variety of statistical methods that one could apply to active surveillance. Dr. Gould talked about some of these. Each of the methods that has been discussed in a formal manner has been dealing with data that is not longitudinal in nature. So, what I will be talking about today is data following people as they progress through time.

[Slide]

Most pharmacovigilance relies on spontaneous reports, and we have heard about that before. The problems with spontaneous reports are underreporting, the absence of a reliable denominator, the possibility of data mining for signals.

The question was could we develop a population-based approach to adverse drug reaction screening that would be longitudinal and potentially, once perfected, prospective in nature so that it could occur in real time?

Some of the questions that would come up in such a system, however, would be what is the

value of a positive signal? In other words, can you believe a positive signal when it says it is positive, and can you believe a negative signal when it says it is negative? So, what I will be describing are two complementary statistical techniques that are being used in this project.

[Slide]

As background to the project, it is being performed within CERT, which is one of the Centers for Excellence in Research and Therapeutics that are funded by the Agency for Health Quality and Research. It is a consortium of 10 HMOs, with about 11 million enrollees. These systems are record-linked and have traditionally been used for hypothesis testing purposes. Three of the HMOs within the CERT are also funded by FDA within our cooperative agreement program, and it was that linkage that enabled us to participate in this project.

[Slide]

The first approach that is being explored is something called sequential probability ratio

testing. What this basically amounts to is a periodic analysis of data as it accumulates, the sequential measure of signal strength. It is statistically based and it can adjust for a variety of covariates that are present in the data that you are using.

[Slide]

The way the system works is that it takes some period of time that is predetermined by the research question and looks at patients who begin the drug during each segment of time, and then each of those groups then represents an independent cohort of patients that will be followed for a certain period of time for the occurrence of the event.

So, in this slide what I am trying to show--this isn't my slide; this is Dr. Davis' slide--basically, each week you got a new cohort of people who are starting a particular drug, in this case maybe a child being vaccinated. Then, they are followed for 30 days until the occurrence of an event. So, each of these cohorts in the analysis

will represent an individual data point in the sequential probability ratio testing.

[Slide]

This slide is by courtesy of Dr. Platt. It is to show what the system, if it works in its optimal fashion, might be capable of doing. To orient you to the slide, along the X axis we have calendar time over which the surveillance was occurring. What we have with the green lines are sort of two statistically set points, the null hypothesis which would basically be if your measures come close to that--and I will describe what those measures are in a minute--it suggests that you suggests that you don't have a signal. If you were to cross some threshold, which we have labeled the alternative hypothesis and here this alternative hypothesis is an event that would be occurring ten times greater than expected, you would have what you would call a signal.

What happens is you can see along the Y axis we have the likelihood ratio, the log likelihood ratio, a statistical measure of

association. So, basically, the higher that number, the greater the difference is from expected to what you have. You can see a series of points, and each one of those points represents a separate weekly cohort of children vaccinated, in this case, with rotovirus vaccine. The event that we are looking for here is intussusception which is a relatively uncommon event.

So, here we can see that in each of these different cohorts there were no events. So, the log likelihood ratio which compares basically what is observed with what is expected is calculated out. Since the expectation is very low, you end up with a low ratio. Then, one case is determined in this cohort at this calendar time of February of 1999 and that raises the log likelihood ratio. Then, subsequently with each accruing case in sporadic cohorts we get the accrual of evidence of a signal. By this time point--I guess it is sort of in April or May--the signal cross the threshold of the alternative hypothesis and remains elevated.

Now, it turns out that the rotovirus

vaccine was licensed in August of 1998, and through the period of July, 1999 when the immunization was halted and the drug was then subsequently withdrawn, there were 15 VAERS reports, that is the Vaccine Adverse Event Reporting System. It is an AERS system for vaccine-related adverse events. So, at a similar time point that reports were accruing in the VAERS system this sequential random probability testing algorithm would have detected a signal of intussusception.

Now, this experiment was done retrospectively. The data were collected in a longitudinal database within the HMO and we are looking retrospectively to see what would have happened if we had been following these data prospectively in real time. It is sort of a demonstrated proof of principle. Subsequent work has shown that there needs to be a modification of what gets measured and now, rather than measuring the log likelihood ratio, we are measuring the log of the relative risk, and there are statistical reasons why that is a better measure.

[Slide]

The other statistical approach that we are using in this pilot project is something called the

tree-based san statistic. This is an approach that can be used with hierarchical data. That is data that is basically embedded within increasing granularity of data. ICD-9 and MedDRA coding, for example, drug classification systems are examples of hierarchical data structures. In the tree-based scan statistic we make no a priori assumptions regarding associations. What happens is that, if you can sort of think of a tree and you have your trunk of a tree and then it has different branch points, what the tree-based scan statistic does is it moves up the tree and at each branch point conducts a statistical test comparing a particular drug with a variety of reactions or a particular reaction with a variety of drugs, depending on which hierarchy you want to move up. At each cut it evaluates what the maximum likelihood is and then pursues that likelihood to the next cut point.

[Slide]

This is just to illustrate what we mean by hierarchical data structure. You start out with all drugs, and within that drug there may be a particular subclass of analgesics. Within the subclass of analgesics we have non-narcotics. Within non-narcotic analgesics we have NSAIDs.

Within NSAIDs you have specific classes of NSAIDs and you can get right down to actual specific drugs. So, you can kind of see how the data are embedded and you can imagine a tree with branches. For adverse drug reaction events we have the same sort of hierarchical structure.

[Slide]

This is just my crude attempt to illustrate what that might involve. We get up to the first branch point and we are looking maybe at a particular drug and we are looking at a host of different reactions, and we want to see what the maximum likelihood is and we see that it goes here. So, we follow that branch of the tree to the next branch point and then conduct another test, and it sort of guides us to the place where we have the

maximum likelihood for a particular association.

[Slide]

In any event, that is sort of a high altitude level of the pilot project that we are involved in. I probably got some of the details wrong and, Rich, you can correct this and disabuse the committee of the mistakes that I have made.

What we are doing is this pilot testing of candidate drugs and candidate adverse reaction events. We have specific sort of known associations, ones that we believe are real and are pilot testing those, as well as a couple that we don't know what they would show, and then ones where we believe there is no association to basically test the sensitivity and specificity of the model. If we find that it appears to be successful and that it is worth pursuing, the goal would be to modify it so that one could do prospective surveillance in real time. So, thank you for your attention.

Question and Answer Period

DR. GROSS: Thank you very much, David.

We now have time for questions. We actually have more time than was allotted to make up for what we missed this morning. Anyone from the panel have a comment? Jackie?

DR. GARDNER: I have a question for Dr. Graham and probably also Dr. Willy. David, how do you see implementing the prospective active surveillance? Would you take all new chemical entities, new molecular entities and begin enrolling cohorts and then watch them over time? Would you need to have a hypothesis? How would you operationalize what you have just shown us?

DR. GRAHAM: I think at this point we are working with specific hypotheses and looking for those. I think it might be possible to generalize sort of to a data mining perspective. In a sense, what we are doing is data mining. It is data mining that is closer to hypothesis testing because we have prespecified what the combinations are that we are looking for. The hope would be, I think, to generalize that so that it could be used as a more global screening mechanism. The question then

comes, you know, which drugs do you do it for and I think that would be dependent on resources, extent of use of the drug product, and then being able to sort of marshall, in this case, the HMOs to provide the data on an ongoing basis to accomplish that.

DR. GARDNER: I would expect that. I guess my question is do you have a professional feeling for what might be the best way to use this?

DR. GRAHAM: I think at this point, and in your discussion I would refer you to Richard Platt because he can speak more to this probably than I can--I think at this point we would probably be interested in looking at associations that we suspect might be real until we sort of have confidence in what the value of the method is in terms of its positive predictive value and its negative predictive value. So, I see it as an adjunct. It is possible that we could come to a point where for some things it could compete with passive surveillance, for example, in terms of signaling an event but I think we are pretty far away in time from when that would happen.

DR. GROSS: Dr. Platt, any comments?

DR. PLATT: Well, David did a great job of summarizing that. As you could tell from his

description, this is being built on work that CDC-sponsored vaccine safety data link has been doing, and it is being moved into production mode and is intending, over the next couple of years, to use this approach to monitor all newly released vaccines that are used in this pediatric setting. The intent there is to focus for each vaccine on a small number of outcomes that are of interest because of prior knowledge or because of information that comes from the pre-licensure testing.

A second piece that is really not sequential in nature is to look at all events that result in hospitalization. It is fairly computationally intensive to deal with each of the potential outcomes so I think at present it is not a great way to be very efficient, to look at all potential outcomes.

DR. GROSS: I will remind the group that

Dr. McCloskey's presentation is also available if anyone has any questions. Dr. Crawford?

DR. CRAWFORD: Thank you. My question is for Dr. Mendelsohn. Dr. Mendelsohn, I am wondering if the NEISS-CADES database also captures and is able to differentiate certain sources, such as if someone presents to the emergency department where an amphetamine prescription drug versus crystal-meth. or illicit street drug or some other sources? How does the database handle these, especially opioids and amphetamines?

DR. GRAHAM: Right. To make sure I understand your question I will just reiterate the point about the street drugs or illegal drugs, are they excluded from NEISS? Sometimes those will slip through but in the quality assurance that the CPSC does and that the CDC does we will find those cases and we will get rid of them. The related DAWN system that Dr. Willy had talked about, that will collect cases of the illegal products.

DR. GROSS: Dr. Ruth Day?

DR. DAY: Dr. Graham, with the tree-based

scan statistic--I was sitting on the edge of my seat, waiting to hear what this might buy us and I have some thoughts but I am not an expert in this area. Could you tell us--I know this is just a pilot project but, say, after X number of years what we might see, and how that might then inform the way data are collected, coded, analyzed, and so on in other systems? Do you think there will be a ripple effect to other systems already in place?

DR. GRAHAM: I suppose it is possible to think that other data systems might want to apply this technique. Certainly, large healthcare organizations that have collected their own data, I could very easily see them wanting to adopt techniques like this or other ones. You know, right now it is sort of experimental and we are trying a bunch of different approaches to see which ones have higher yield and are most efficient, and the like. But I think once methods that have relatively good predictive value and are efficient are sort of identified, it wouldn't surprise me at all if most healthcare organizations that have

large databases would implement systems like these for their own quality control and for their own risk management, if not for more public health oriented things where they are going outside of their health plan. So, if there are particular data requirements that aren't being collected that would be useful for this, then I can imagine that health plans might want to introduce those but, based at least on my experience with healthcare databases, I think the types of data we are working with now--all of those databases have that type of data. The real problem would be how accessible the data is; how linkable these things are; and what their computing capacity is to deal with this.

Now, it may turn out at some point that there is additional data that people identify that actually might make approaches like this, or other approaches, even more useful and more informative, data that is not currently being collected, and I could very easily see a health plan implementing the collection of that type of data if it wasn't too intrusive and if it wasn't too difficult to do.

I suppose ultimately--you know, we had the press announcement from Dr. McClelland, the head of CMS, the other day talking about the broad vision of being able to link Medicare data, Medicaid data and data from everywhere in the world, and data mine it and do everything else to it. Maybe down the line these things will happen. But I think in the short run we are in the experimental mode and basically people are trying a host of different approaches. Listening today to the different approaches, the ones that were presented in the public session and the ones that were presented by FDA and Dr. Gould's talk, to me, I found the diversity of approaches to be delightful. You know, it is great to see that.

DR. DAY: It just might be that the level of the nodes that yield payoff might emerge.

DR. GRAHAM: Yes.

DR. DAY: So, in any hierarchical structure there may be some break points that are really important for other systems to implement.

DR. GRAHAM: That is true. Are you speaking from sort of a personal experience or from

your domain of expertise?

DR. DAY: I am speaking from the domain of mental representation and hierarchical structures and all the research on how memory is organized in hierarchical structures, and how easy it is to scan and understand information in tree diagrams, and that levels of nodes can be very important.

DR. GRAHAM: That is great! Here is an example where the person who is being questioned asks the questioner!

DR. GROSS: It sounds like you are going to have a good conversation afterwards. Allen Mitchell, please?

DR. MITCHELL: Yes, I haven't had my second cup of coffee so forgive me if I am confused, but I thought that today we were talking about active surveillance approaches and yet, David, as you have just described, you are really referencing a very interesting technique for mining available data and I am not sure if I missed the transition or whether we are talking about still active surveillance in the sense that it was

defined earlier as active surveillance by healthcare providers or facilities.

DR. GRAHAM: The idea is active surveillance broadly defined. Mary Willy sort of tried to give a definition of what active surveillance might be in various sort of dynamics or sort of approaches that one might take and then Aaron's and my talks were designed to illustrate basically I guess a couple of things. One, FDA's interest and active involvement in the notion of active surveillance and, two, as examples of what active surveillance approaches might look like. Our attention wasn't to narrow your thinking at all. It was basically just to sort of help to get the juices flowing so that the synapses will connect and you will come up with some really good ideas--not you, personally, Allen, but collectively. Then we could take that and benefit from that collective wisdom.

DR. MITCHELL: So, if I can rephrase it, the active surveillance sort of concept that is being used here is more of a process than it is a

focus on particular data sets.

DR. GRAHAM: The way we presented things, if we have led anyone on the committee to think we are wedded to a particular approach or a particular idea, please--

DR. MITCHELL: No, no, David. I am sorry to interrupt, I didn't mean that. I didn't mean that you were leading us towards a particular approach but, rather, active surveillance is an activity as opposed to a definition of data sources.

DR. GRAHAM: yes, it is an activity but it will frequently involve data sources of one sort or another.

DR. MITCHELL: Clearly. Thank you.

DR. GROSS: Elizabeth Andrews?

DR. ANDREWS: Actually along the same lines, a question for Dr. Graham, I am a big fan of using healthcare databases but one of the historical problems in using them for more active surveillance has been a lag time--

DR. GRAHAM: Yes.

DR. ANDREWS: --before data are available, and I wonder if you could comment on where we are with that now, especially with the system you are

describing, as well as what do you see in the future?

DR. GRAHAM: Well, the system that we have was, as I said before, retrospective, and it was retrospective for a reason. It is because, one, the technique is not developed and, two, there is this inherent lag in data and to some extent I think that that will be probably insurmountable, depending upon how certain you have to be and the seriousness of what it is that you are looking for. If you are talking about a reaction that sort of results in hospitalization or that is indicated by a particular procedure, it may be that within three months the lag is short enough, and traditionally we talk about a six-month lag or something like that. Prescription data comes in and maybe within a month we have all of that. Sometimes we have that almost instantaneously but the procedures and the diagnoses, and the like, have sort of a broader

distribution. So, that basically then locks you into a built-in lag.

I suppose one could look at systems to see what is the distribution, and I am sure each health plan knows what is the distribution of lag within it, and at what point do I have 75 percent of the data. I think that would probably be a reasonable place to initiate an active surveillance of one sort of another, and it may be that for expensive procedures, hospitalizations, maybe those data come in more quickly than data that come from physician offices, and the like. So, that may affect lag time as well.

But I think inescapably we are probably never going to be in true real time. I don't imagine that we will probably ever be in a situation where it is less than three months lag. At least for the systems that we are dealing with here, as Rich mentioned before, they are resource intensive in terms of what is required to get the data, organize the data and then to process the data. So, what that means is that probably the

types of active surveillance that I have described would not be done on a continuous basis like RADAR that is going continuously, and I see the little blip and I can watch it get brighter, brighter and brighter. It is probably something that would be done on some periodic basis, whether that basis is monthly, quarterly, semiannually will remain to be determined. But I think that, at least the systems that we are talking about here, unless there is a huge advance in computer technology and the way files are organized, I think that we are probably stuck. At least, that is my own impression.

DR. GROSS: Thank you. Lou Morris?

DR. MORRIS: I have some questions for Dr. Mendelsohn. I was very struck that this is the only database that has been discussed as nationally representative. I was wondering how representative it is and if you can give us some details. For example, what percentage of emergency rooms are representative? What is the sampling error? What is the participation rate? How often is a frame re-sampled?

DR. MENDELSOHN: I am happy to address that. First let me clarify that it is not the only nationally representative active surveillance

system; it is one of the only ones. There are other systems that are out there, of course, the DAWN system that Dr. Willy had mentioned. But, you know, this is the one that FDA has been exploring thus far so it is the one we know the most about.

There are only 64 participating hospitals at this point so it is a very small sample. In terms of the projection factor is, that I do not know but I can tell you that statisticians from the Consumer Product Safety Commission update the analytic weights and the sampling frame every single year based upon changes in the healthcare environment, hospital mergers, that sort of thing. So, it is updated periodically.

DR. MORRIS: Do you know the sampling error? How wide is it?

DR. MENDELSOHN: That I don't know. I can tell you that, having spoken at length with some of the statisticians from the Consumer Product Safety

Commission, generally they like to see about 20 adverse events for a particular drug or drug class to feel comfortable making projections for that drug.

DR. MORRIS: Do you know how many emergency rooms there are nationally?

DR. MENDELSOHN: That I do not know. Sorry.

DR. GROSS: Annette Stemhagen, please?

DR. MENDELSOHN: Judy, about how many hospitals are there in the U.S.? About 5,000? There are about 5000 hospitals in the United States. Not all of them have emergency departments but if we were to say, you know, two-thirds of them have emergency departments we are talking about probably 3,000, 4,000, somewhere in that neighborhood of emergency departments would be my guess.

DR. STEMHAGEN: My questions are also about the NEISS-CADES database. It seems like it is a very unique database for finding adverse events for people going to emergency departments.

I wanted to clarify first though that in order to identify that adverse event the physician must indicate in the chart that it is an adverse event.

DR. MENDELSOHN: That is correct and one of the limitations of the system is that if the physician does not indicate that in the chart then, of course, it will likely be missed.

DR. STEMHAGEN: It seemed like many of the examples that are very well known, insulin and hypoglycemia, for instance.

DR. MENDELSOHN: Exactly, and that is why I said that in theory NEISS could capture the unexpected serious adverse events. Again, in theory.

DR. STEMHAGEN: I guess the follow-up to that, and it is something that I think people have been talking about for a long time, is the education that goes along with it, not for the coders but for the physicians, and are there processes in those 64 hospitals to actually educate the emergency room docs on adverse event identification? Because otherwise I think we are

really only going to get the very common things that are pretty well known.

DR. MENDELSOHN: I would certainly agree with that, and that is why we have not done any training, any education for the physicians. We try to make it as simple, as non-invasive for the hospitals as possible.

DR. STEMHAGEN: And having spent time doing research in emergency departments, people are very busy and the likelihood of things getting documented if there is not reinforcement is probably--I think it has a lot of potential here but I think there is a lot of difficulty unless there are some other things that are sort of overlaid on top of it.

DR. MENDELSOHN: Right, and I think the primary selling point for NEISS versus the AERS system--it is not going to replace AERS in terms of collecting the unexpected events, but in terms of being able to quantify events that we do know about, which would certainly be useful information, that is where a system like NEISS is extremely

helpful.

DR. STEMHAGEN: I mean, I guess in the risk management the interventions on even known events might be able to suggest some things, but in terms of using it for other signaling I guess it is difficult.

DR. MENDELSOHN: Right, I would agree with you, it is not there at this point.

DR. GROSS: Anne Trontell?

DR. TRONTELL: I only wanted to make some suggestions. We will certainly see shortly national projects for NEISS but it clearly does rely on the recognition of adverse events, and most commonly those will be things physicians already know about. It actually shows some promise for some of the materials that Dr. Holquist talked about this morning, medication errors, people who overdose or inadvertently overdose not simply related to insulin or warfarin. So, we have some provocative examples in an area where we have perhaps even greater underreporting than we do with adverse events.

DR. STEMHAGEN: That is why I think in terms of risk management for the misuse kinds of things it would probably be very useful. That is

who is going to present to the ERs most likely anyway.

DR. GROSS: Henri Manasse?

DR. MANASSE: One of the things that impressed me from Dr. Bennett's presentation before lunch was the sort of rapid cycle improvement concept that was used from the data itself because it was collected locally; there was a follow-up locally; and there were certain improvements that were implemented fairly quickly. If we jump over to, say, a NEISS program and some of the other programs that Dr. Willy addressed, how is that data being used to improve patient care? And, I think maybe we will get into this later, I think there is a significant difference between using these data to improve care and improve safer use from the bigger policy questions about should this drug remain on the market.

DR. WILLY: Well, the programs that I

described, they are all surveillance systems but we are not necessarily using them for active surveillance. So, at times we may use data from those sources to help us but I can't speak to how we are using them currently. It is a very good question though.

DR. GRAHAM: Active surveillance at FDA is in its infancy so we are basically experimenting with it and learning about it. That is one of the reasons why we sent out this RFI, this request for information. It is a relatively new field and there is very little that has been written about it. The Centers for Disease Control have done active surveillance for a long time and they have infrastructure looking for nosocomial infections, for injury and for a variety of other things, sexually transmitted diseases, HIV. But for classic sort of pharmacovigilance we don't have sort of an infrastructure or a system in place. So, that is what we are trying to develop now.

The question that you asked about can we use it for continuous improvement to improvement

patient care, and the like, at this point nothing that we do can accomplish that. What Dr. Bennett talked about this morning I think would fit very nicely as a form of active surveillance. You are kind of beating the bushes for cases and you are doing it based on an observation, intuition, clinical experience that sort of tells you that there could be something going on and then having the moxie to go out and look.

The French pharmacovigilance system is probably the closest system in the world to what Dr. Bennett has described, and there it is a national system and they have regional clinical pharmacology departments located within university hospitals, and the physicians who live in that geographic area, around that research unit, receive periodic lectures from the staff of the pharmacovigilance unit. They understand that it is important to report things and so there is almost a personal relationship, if you will, between the unit and the physicians in their area. They are also able, because of that relationship, to feed

information back to those people if they identify that there is a particular pattern of misuse, let's say, or off-label use or overprescribing or using a higher dose than maybe should be used. They are able to feed back, with the idea that maybe you can improve patient care.

Now, there are systems, not that the FDA uses them, but health plans with automated data sources, Medicaid does this and they have something called drug utilization review. There, they are looking sort of basically for pattern recognition of inappropriate prescribing. So, there are algorithms for what would constitute inappropriate prescribing. The idea is that you can feed back to the physicians saying, you know, you are prescribing too many narcotics or you are prescribing too much of this or too much of that, or did you realize that this patient is over age 65 and shouldn't get this drug, with the goal of trying to modify prescribing behavior. FDA doesn't do those things but there are other models out there.

DR. GROSS: Can any of you help me with making a connection between passive and active surveillance where you detect a safety signal

through passive surveillance and then try to document it further in large databases, and perhaps initiate an active surveillance project to confirm it?

DR. GRAHAM: I have been perplexed by that question as well and I will give you sort of my view. I don't know if anyone else shares it. With the passive surveillance things are just basically coming to us and there is no rhyme or reason to it and we can't predict what we are going to get. But it does provide information that can signal things that are important. Active surveillance, at least the way I think about it, is a parallel mechanism that might be used for a similar purpose to identify new signals. You might use it to try to strengthen a signal that you see in a passive system. But when you are doing that what you are almost doing, in a sense, is a hypothesis testing study. So, the lines become blurred between what

is active surveillance to validate a signal and what is actual hypothesis testing.

I think that the attraction of active surveillance is that you are doing it maybe prospectively and that you can rely just on automated data to do it so that you don't have to go to validate diagnoses, and the like. That would require sort of a third step which would be an actual formal study. But then, in my own mind, that raises the question if I have done active surveillance in a particular data resource and I have identified a signal in that database, is it legitimate for me now to do a hypothesis testing study in that database upon which the signal was generated? Because now what I am trying to do is I already have a Bayesian prior that I am going to find something there. So, statistically--you know, I think eventually we are going to have to come up with techniques where maybe we do our active surveillance on part of the database so that we are saving the rest of it so we can actually do confirmatory studies if that is necessary.

That is my own idea. So, I don't know if that helps in any way. I think what you are struggling with is what we are struggling with and

why it is on the agenda here, which is sort of what is active surveillance? What should active surveillance be? What might it look like? And, you know, none of us can really describe it very well. So, we have done what we can to try to describe some ideas and maybe the committee can provide us with greater clarity or sort of a construct to think about it, an organizing principle. I think that is what we are looking for, an organizing principle.

DR. GROSS: Well, if there are not questions, why don't we take a break early? Then, we will reconvene and the committee will address the questions that it has before it. Why don't we take a 15-minute break and reconvene at about 2:30?

[Brief recess]

Questions to the Committee

DR. GROSS: We do have to proceed with the meeting. While you are sitting down, I am going to

start by reading the questions. You all have them. We are going to discuss and compare passive and active surveillance. The passive surveillance will focus on the Adverse Event Reporting System that the FDA discussed earlier today. They would like us to comment on the question what types of safety problems are most effectively addressed by using a passive surveillance system such as AERS that depends on voluntary reporting.

Secondly, are there safety problems where use of this system is less effective? If so, please specify the type or nature of these safety issues where passive surveillance is ineffective.

Number three, how can the FDA passive surveillance system be improved?

Under active surveillance, how can active surveillance systems be used--I don't think I am going to read the rest of this. You all know I can read! Why don't we start and doing this one at a time? I am sure there will be some overlap but that is okay. So, let's start off with passive surveillance, the first question, what types of

safety problems are most effectively addressed by use of a passive surveillance system such as AERS that depends on voluntary reporting? Some of this, of course, was already answered this morning. Would anyone like to reiterate? Curt?

DR. FURBERG: Serious and rare events, as I see it that is probably the primary use of the passive system.

DR. GROSS: You weren't even here this morning so you pass the test!

[Laughter]

Very good. Anyone else want to comment? Yes, Sean?

DR. HENNESSY: First, I just want to get on the table that in the paradigm of risk identification, risk measurement and risk management I think we are still talking about risk identification so that none of these systems is, in and of itself, going to impact on patient safety unless the downstream things are done as well.

DR. GROSS: Okay, that is a good point. Anyone want to add anything other than that serious

and rare events are the thing that we can count on picking up by the AERS system? Henri?

DR. MANASSE: I would like to perhaps not answer that question directly but take a little different view. What we are seeing across the country is increased utilization of pharmaceutical agents. We are seeing a broader range of medications being available in the marketplace and consequently many drugs taken concomitantly, and relying fully on an AERS system I think voids our capacity to be able to pick up all the signals in the marketplace.

From a prospective point of view, I think we have to critically examine whether we want to continue in this direction. Are we really serving the public well from a safety perspective by continuing on relying largely on an AERS system in the context of this ever-growing complex drug environment?

DR. GROSS: Henri, I am going to take the chair's prerogative and ask a question myself. Of the major safety problems that have been

encountered in recent years, could someone from the FDA comment for us on by what systems have these problems been initially identified? Were they the AERS system?

DR. TRONTELL: I don't have, you know, in the top of my mind the list of products withdrawn or with significant safety problems. I am going to say broadly that the majority have come from the spontaneous reporting system, it has at least for these rare adverse events--liver failure with troglitazone, certainly, continued pregnancy exposures with isotretinoin. There is actually active surveillance for that.

The number of instances where other mechanisms have identified the safety problems are quite small. As this committee knows from its meeting in February, we had clinical trial data inform us about the safety risks associated with the COX-2 selective NSAIDs. But that really represents the minority. But I will appreciate any others from FDA to add to that.

DR. AVIGAN: I would just add to that that

in many cases a safety signal was appreciated early on in the AERS system as, in fact, the presence of a signal, the presence of a series or a number of reports. What it didn't offer was a quantitation of risk or a quantitative sense of the burden or the range of severity. So, in some cases information about the signal was then complemented with other kinds of studies that were done either in parallel or sequentially after the signal was initially detected. There are examples where safety concerns were raised at the time of approval of a drug but the range of severity or the implications with regards to safety in a larger population exposure weren't really known until the drug was put out into the marketplace.

So, I would say the answer is that AERS had played an important role to determine that there has been the presence of a safety problem during marketing, and this information has been usually or often complemented by other kinds of studies, other kinds of tools for quantitative risk analysis.

DR. GROSS: Mark, I think that is probably the key. Until we have something better, we probably shouldn't throw out the AERS system at

this point. But maybe what our advisory committee can do is offer some suggestions as to how the safety signals from the AERS system can be followed up better, quicker, in order to try to come up with the answer a little sooner. So, that may be another question that would need to supplement the questions that are here, what should be combined with AERS. As Henri said, we need other kinds of surveillance systems, but we are probably going to need to have both at our fingertips and be able to access both. David Graham has sat down but he was up before. I think he wanted to say something.

DR. GRAHAM: It was just to say that if you look at major drug safety problems in the last 20 years, probably 90 percent of those came as a result of the AERS system or its predecessors. Encainide and flecainide were discovered in a randomized clinical trial. There are a few other examples where they came from clinical trials.

What Dr. Trontell said is basically that that is the bottom line.

I just have a few remarks about subsequent remarks that Mark Avigan made, which is that if you have a rare serious event, acute liver failure that has a because rate that is impossibly low, you don't necessarily need confirmatory epidemiology to tell you that you have a causal association. So, I guess in that position I have a different opinion than Dr. Gould did this morning.

So, the notion that you need to wait months or years to confirm something before you do something I think would be a misapplication of the primary purpose of spontaneous reporting, which is to identify disasters and deal with them.

DR. GROSS: Thank you. We have a couple of other questions. Allen Mitchell?

DR. MITCHELL: I think, Peter, that in a way we are heading to the third part of the question--

DR. GROSS: Yes.

DR. MITCHELL: --which is how can the FDA

passive surveillance symptomatic be improved. Two comments I guess. One is that by definition there is competition for resources so any resources that would be devoted to the improvement to AERS would, I would think, come at the compromise of other opportunities, number one. Number two, I am not sure that it is at all clear--I think there is a great role for alerters or astute clinicians, as evidenced by some of the information just provided. The question is whether "improvements" in AERS is going to make any difference in that outcome. To me, that is really unknown and my own view is that it wouldn't because, as David is describing, the risks we are talking about are huge and fine-tuning those with a slightly better this or a slightly better that imposed on a system that is inherently flawed--I mean, we all have discussed the flaws of AERS--I think would be to throw a lot of resources against something that is functioning reasonably well for what it is and maybe ought not to be a candidate for further improvement.

DR. GROSS: I would think then, let's say

if that is a conclusion, it would be important to document, based on past experiences, whether or not safety signals, once identified--was additional information accumulated fast enough to deal with the problem.

DR. MITCHELL: You mean within AERS?

DR. GROSS: Yes, within AERS or other surveillance systems.

DR. MITCHELL: But my understanding was that the purpose of AERS was to identify, if you will, a hint of a signal as early as possible, and that signal would then be taken to any number of potential data sets, both within FDA and beyond FDA, for testing. If that is the case, then AERS is designed only to identify signals, not to confirm them by further data collection because one would hope, once there is a signal, you have gone off to a more reliable data set.

DR. GROSS: Right, but that is in theory. The question is, is that what has been done and is that what we should do?

DR. AVIGAN: Can I add just one more

thing? It is important to realize that besides bean counting, besides actually counting the numbers of cases of a certain adverse event and determining the numerator burden with all the underreporting, etc., the analysts who are looking at these cases are looking at other dimensions as well which are very important in understanding that there is a signal, particularly the range of severity of the adverse event over time as more of these adverse events accrue in a large exposure population, to see what is the distribution, and also to do a causality analysis given that in some cases the information that is provided is insufficient to allow for a very sort of crisp likelihood analysis. But in other cases that come over time this causality analysis is possible. So, over time you develop a collection of cases, a so-called case series, that allows you to get some handle on the range of risk not only from the point of view of are you collecting cases, but what actually is the nature of these cases and what is the linkage to the drug. The analysis of AERS

cases is useful for all those.

I think Joyce Weaver, this morning, really emphasized this idea of case review. This is an extremely important element in the analysis because it allows you to look at the collection of cases that you have as a totality and look at the distribution with regards to range of severity, number one and, two, to look for those sentinel cases where a likelihood analysis allows you to link the drug to the event.

DR. GROSS: Robyn Shapiro?

DR. SHAPIRO: This is really not a totally voluntary system. The sponsors have to report. I can't understand why we wouldn't make a recommendation for mandatory reporting by providers. This happens in other public health arenas. JCAHO--you have to report a sentinel event; public health, you have to report a death in a nursing home. You have to report a gunshot wound. This makes no sense. By just changing that with an amendment in the law we could deal with so many of these problems. So, this is probably

something naive but I just don't understand it.

DR. GROSS: Good question. Ruth Day?

DR. DAY:; The movement seems to be towards active surveillance systems and I would like to speak up in favor of the so-called passive ones. There are some small local fixes that maybe have already been done and can be done very cheaply. We heard a number of times this morning that one of the problems with AERS is that there are duplicate reports. Well, you can write a template-matching little sub-routine in the programs to match cases and if X number of elements match, then a case reviewer looks at them and you can then reduce the duplicates, and so on. That is I think a pretty small, easy thing to take care of. So, before movement away from this system gets too far, looking at some quick and easy fixes like that I think would be important.

Also, as the system exists now it can help us understand more about what we are not supposed to be looking at, namely, expected events. We are supposed to be looking at rare, serious unexpected

events. Well, in clinical trials there are certain things that we are supposed to expect a little bit. What if the rate is higher than we thought or if it is in some subpopulation? So, the possibility of having age and gender and ethnicity popping for us is great. I mean later, of course, you can go and look for them specifically but it is a little bit of a circularity. In active systems you have to know what you want to look for and you can then go and look for certain kinds of patients, settings or diseases, and so on, and you may miss where things are going on.

So, whatever happens on the active surveillance front, I think small fixes to help AERS and such systems now can really bring us things that we can't even think of. So, there are more things in heaven and earth than are dreamt of in our philosophies, as was said in Shakespeare, and the same in public health, we don't know until things pop up and this is the way to see them.

DR. GROSS: Curt Furberg?

DR. FURBERG: Two comments, one is on the

effectiveness of the AERS system. I just want to remind you about the paper by Friedman and co-workers, from the FDA, published in JAMA around 2001, where they talked about five drugs taken off the market and they had a table there showing the number of people exposed to the drug before it was taken off the market. When you added up the number of people exposed to the five drugs, it was close to 20 million. So, about four million people are exposed to drug before it is taken off the market and that represents more than 10 percent of the adult population. So, the system is not effective; it is a failure.

In terms of how do you detect side effects, there is a paper from quite a while ago in the British Medical Journal, by Ware, who addressed this specific question, how do we discover adverse effects? And most of them were through case reports and letters to the editor. That is the first time something came up and very, very few came through clinical trials and large scientific studies.

DR. GROSS: I guess one of the questions is, is the AERS system like democracy? It is a terrible system but we haven't found one that is

any better yet. I guess when we do it can be replaced. Art Levin?

MR. LEVIN: Following up on that, Peter, I guess I think what we are talking about--and I am just looking for clarification--is really should the passive surveillance system be improved, and it goes back to the resource question. So, I guess I need help in figuring out, because I think I am hearing that more wouldn't make the system perform any better. In other words, our concern about underreporting is legitimate but going from 5 percent to 10 percent of what is out there, 20 percent or 25 percent isn't really going to give us any benefit. And better might not give us any benefit. In other words, more detail in the reporting might help a little bit; might speed up the process of analysis and case review if the data in a case was more sufficient to the task.

So, that is why I would almost reword (c)

because it is a resource allocation question so it is should we make it better or should we say it is what it is; it does about as much for us as we can hope for and we really need to be looking elsewhere for real improvement.

DR. AVIGAN: Could I just quickly comment? I think there was a comment made before that if better reporting could be solicited for those kinds of adverse events that we are interested in specifically from physicians where differential diagnosis and the important background information to allow a good likelihood analysis was given, from my vantage point as an observer to this process, would be a great reform. So, the idea of soliciting the healthcare system, the reporter population, physicians etc. to do a better job; to be better citizens about reporting, if that could be achieved and could be implemented in a way where we communicated this to the health provider world, that would be a great advance I think. So, I just want to take exception with that point. Quality of reporting is very important.

DR. GROSS: Anne Trontell?

DR. TRONTELL: Actually, these questions and comments all converge on what Robyn Shapiro

asked about mandatory reporting and how that might improve volume and quality. There is some experience in other FDA mandatory reporting systems. The Vaccine Adverse Event Reporting System is technically mandatory but, to my knowledge, has no better reporting of information than the Adverse Event Reporting System. Similarly, for medication devices there is a requirement and in that instance they have actually gone to develop an active surveillance system because the volume of reports received was quite low.

In the area of requirements, FDA at least has proposed a rule, the adverse drug reaction reporting rule, that actually speaks to activities on the part of pharmaceutical companies to do what is called active query of these cases, the feeling being that at such time as the clinician has contacted you with an adverse event is really your

opportune moment to collect the most information. In instances where there is a delay or paper feedback back and forth might impede the ready collection of information so you can have targeted inquiry to the specific adverse event reporting. That might give us more complete information. That is the proposed rule at this time.

DR. GROSS: Before we get to the third item, I am going to ask you to comment on the second item. Are there safety problems where use of this system is less effective? If so, please specify the type or nature of these safety issues where passive surveillance is ineffective. Anybody want to comment on that? Curt? Jackie?

DR. GARDNER: It is not clear to me the committee can improve upon what the FDA has already presented to us this morning. I think they have a very comprehensive catalog of limitations for all of these systems, and the answer to that, of course, is yes and it is effects of long latency and effects that have high background rate. I don't know that we are going to add to that. It

seems to be a question that we could spend a lot of time talking about and shed no light on.

DR. GROSS: There are a couple of other comments. Allen Mitchell?

DR. MITCHELL: Yes, I want to be clear that I wasn't suggesting that AERS was of no value. Quite to the contrary, I think it is of considerable value. My question comes back to what I posed earlier, is improvement going to be cost effective, if it can be achieved at all? I will go down on record as saying to ask doctors on a mandatory basis to report adverse reactions is tilting at windmills, at best. I think the information you would get would be ampicillin and amoxicillin rashes but not things where they are likely to get a call from FDA for follow-up; not likely to be anything where there could be a prospect of litigation. They run very busy and I can't see this as something that they will do. I would urge that the resources be put into alternatives that can buffer or buff up the signals that are identified in AERS. It is just a question

to me of additional improvements, not whether the system is useful as it is. It is quite useful I think.

DR. GROSS: Stephanie Crawford?

DR. CRAWFORD: Thank you. I wanted to tough on something that Curt and Robyn mentioned earlier. I am looking at question 1(c), by the way, how can we improve the system? I do think we need both the spontaneous reports as well as active surveillance systems.

In terms of getting reporting, I don't know the latest figures on where the reports come from. In the past I have seen that of healthcare practitioners the pharmacists submitted the most reports. I am not sure if that is still the case. Then perhaps nurses and physicians. I am not sure where consumers come in there. But I do know patients do submit some. In terms of health practitioners, what we hear is that quite a few come from published case reports. Perhaps there could be some incentive to editors of the major journals--just as we do studies, part of the check

list asks is this approved by the IRB. If they give an adverse event case report, perhaps one of the check lists questions could be has this been reported to FDA, MedWatch, whatever the appropriate agencies are. It might be the hospital system; it might be a state board, but just as a check list question.

I think in order to incentivize practitioners, which is necessary to make the passive surveillance system grow, there needs to be some discussion with the practitioner-based organizations. Several of them run the MedWatch program. Maybe there could be occasional, periodic sections for some of the journals for the different professions to say "lessons from the MedWatch" or something that shows the practitioners that this is used and maybe they should be more sensitive to this, and it may incentivize others to report by simply seeing how it is being used, not that it is just a data collection method, as well as ongoing continuing education.

For consumers, I am not sure the typical

consumer knows what to do if they experience what they may or may not recognize is a potential adverse reaction. If my mother experienced one she would want to talk to areal pharmacist so she might not ask her daughter--

[laughter]

--but in some manner I think there have to be ways that they will either directly--I would actually prefer if they felt more comfortable at least to mention it to their pharmacist and/or physician and/or other primary care providers, but in some manner that consumers actually know because I think the ones who actually are reporting have to be very savvy to know to report to FDA.

DR. GROSS: With familial issues you have a lot of company, Stephanie!

[Laughter]

Richard Platt?

DR. PLATT: I think it is important to recognize how rapidly and dramatically the range of options available to FDA is changing. We are really living in a different world now than we were

two or three years ago and in two or three more years it will be dramatically different. At least a quarter of the U.S. population will be available for study as defined populations to FDA. In that environment I think the question is where should the FDA put its resources.

I agree with Curt that the passive system will be useful for the foreseeable future--foreseeable is probably five years--for very serious rare events. Otherwise, I really think FDA could do better to put some of the resources that are currently going into AERS into accelerating its ability to work in the very rick defined population data resources that will increasingly be available to it. So, I would say the question isn't just should more resources be put into making AERS better but should some of the current AERS resources be invested in other opportunities. Personally, I think that that would be a wise choice of the much too limited resources that FDA has. FDA is working with less than ten cents on the dollar that it could profitably use

but since that is the world you are living in, I think that making the wisest choice in spending is terribly important.

DR. GROSS: Lou Morris?

DR. MORRIS: I guess before thinking about how we can improve the system it would be helpful to know in what way we are trying to improve it. Curt had a really good point. My problem is I don't understand the problem because the problem that there is too much of a time period between when a drug is approved and when a drug that shouldn't have been on the market is taken off the market; that we are not getting enough signals and we think that there are a lot of things out there that we should be aware of; and based on how we define the problem, there is a very different solution.

I would just feel it would be very helpful to get some discussion on what do we think should be the best way for protecting the public health to figure out what the problem is before we start discussing solutions.

DR. GROSS: Do you have any proposals?

DR. MORRIS: Well, you know, I tend to think, like Allen, that it is not a matter of lack

of signals and Curt's idea that it is more a matter of the rapidity or the speed with which these signals are worked up and some decision is made. Maybe the most direct way--I mean, I always think that it is easier to fix it on the back end than the front end because you never know what is going to happen at the front end if you do a systems analysis, and the problem may not be with the passive system. The passive system may be working fine, but the problem may be how do we work it up and how do we coordinate it so that the whole system works better, and look at this more as a systems problem than it is looking at the individual tools.

DR. GROSS: So, you are really getting to item number three, how can the FDA passive surveillance system be improved, or the issue is should it be improved. Why don't we try help the FDA with this particular part of the question

because we are not saying that AERS should be the only system for identifying signals and confirming them, but it is a beginning. It is one that has most commonly identified problems to date. How can it be improved? Henri?

DR. MANASSE: As I listened this morning to some of the barriers to reporting, it ranged all the way from, you know, it just takes too much time to I don't want to be bothered with the follow-up work that is going to go on; I see myself as potentially offering myself up to liability. And, all of those reasons obviously need to be addressed.

But I wonder if we need to perhaps have a different perspective on this as well. I think the notion of reporting to some big federal agency on Rockville Pike probably doesn't find a lot of attraction in community practitioners who are, you know, in South Dakota. My point being that perhaps if we begin to regionalize the system with peer review that can go on, on a regional basis, where you are really reporting to your colleagues for

screening and follow-up rather than just throwing a piece of paper into a big bin in Rockville, whether that might be a better issue.

This is an awfully big country, as we all know. We know that there are regional variations not only in practice but behavior, and it may very well be, if we began to approach this from a regional perspective, that we would be getting maybe expanded reports, maybe different reports and begin to pick up things that this system presently isn't picking up.

DR. GROSS: So, would you like to propose a pilot system or a pilot study where that is done in maybe a couple of areas of the country?

DR. MANASSE: As I think further about this, Peter, we have about 125 schools of medicine. There are about 95 schools of pharmacy. My guess, and I think I am pretty accurate here, is that they are just about in all states in this country, and is there a way, a bit like the French system, where you use these academic centers and not only bring a level of local credibility to the reporting and the

importance of it, but also bring a dimension to research and, to my point earlier, to process and care improvement to the system.

DR. GROSS: Would anyone else like to comment on Henry Manasse's suggestion? Yes, Sean?

DR. HENNESSY: I am wondering if anybody in the room knows with there are adverse events that were identified by the French system earlier than we did, or adverse events that were identified by the PEM system in the U.K. before we identified them. They sound like great ideas on the face of it, and it seems like maybe there is some experience to be able to evaluate how they perform in relation to a spontaneous reporting system that we have.

DR. GROSS: Anybody here from France?

DR. GRAHAM: As far as PEM is concerned, it has never identified an adverse reaction signal that has been acted upon. Culturally, British physicians are sort of educated to report adverse drug reactions so they are reporting rates that are higher than in the United States, and they are also

trained to identify things that are more serious, not necessarily rare but more serious, to report those. And, reporting is only limited to physicians although they have recently widened it so that pharmacists can report as well.

Regarding the French system and comparisons with the U.S., there are no published studies looking at the two systems, when was a signal identified, and what regulatory actions may or may not have resulted.

I think that when you look at the literature, say, in the last 20 years when drugs came off the market, the places where the signals originated were either in the United States or the United Kingdom. They did not originate with the French pharmacovigilance system. Now, in defense of the French, it is a much smaller country than the United Sates in terms of population so I think that we have to take that into account as well.

DR. GROSS: Jackie Gardner?

DR. GARDNER: I thought it was a sobering calculation that Curt did with averaging five drugs

over 20 million people and coming up with the fact that four million people have to be exposed. I wonder if one of the things we should talk about helping with is no matter how good the system is if people don't act on what they are seeing, then the system hasn't failed but we have to do the next step. And, somewhere short of four million exposures, how do we get to that? When David was doing his active surveillance, both the trees and the other charts, it seems that there isn't a wisdom about at what point do you take action and i wonder if we should think about that and, rather than finding ways to trigger the reporting, try to figure out recommendations for when, having seen a signal, one acts rather than waiting until four million people have to be exposed.

One of the things that we have been talking about that FDA has just done is put up the drug watch web site for emerging risks but even there, as I look at that, somebody has to decide that this is worth putting on there for us to watch for. So, I think we need to think about helping to

direct the action from signals rather than how to get more signals.

DR. GROSS: Yes, I agree. I think that makes a lot of sense. I think we would probably need to review several examples in which a signal was identified and then see what it took to confirm the signal as a real problem and was there anything else that could have been done.

DR. GARDNER: Right, and I think the confirmation process maybe is what takes a long time and we don't want to rush, as the media would do, with heightened awareness of something that turns out not to be a problem. If we could figure out at what point, short of a large epidemiologic study in a database that would test the hypothesis, that we think we should be taking some action, that would be a useful discussion over the next two days.

DR. GROSS: Curt Furberg?

DR. FURBERG: There is also a system in New Zealand where they record the first use, similar to the U.K. system. But they don't work

with the physician; they work with the patient. So, they call the patient up, the first 5,000 or 10,000 after three or six months, and get direct information about efficacy and adverse effects, and if the patient is unable to take the call they may find out that they passed away so you get mortality as well. They have picked up cases early on, to answer your question. The only case I remember is arrhythmic death and asthma that came from there. But there are several examples from there that are in the literature.

The difficulty in dealing with U.S.A. and comparing countries is that it all depends on when the drug is introduced in a country. So, you need to control for the approval date and since that is shifting now in the U.S.--we are the first country to approve a drug in--what?--60 percent. All new drugs are introduced first in the U.S. We should be the first ones to discover it. In the past it was Europe, and so on. So, it complicates it but it is a very interesting question.

DR. GROSS: Annette Stemhagen?

DR. STEMHAGEN: There are a couple of things. In terms of the role of the manufacturer, many times the reports that come to the

manufacturer aren't because somebody is reporting an adverse event; it is because they are calling to request information on how to treat the patient or something about the drug. So, if we are going to rely on increased reporting, however it is, there has to be additional education with the providers, and so on, because changing the system without that part--the reasons people report aren't primarily because it is an adverse event, although in many cases, if it is significant, it will be but it is more for request for information. So, I think we need to think about that.

I guess the other thing is we are talking about changes in the kinds of reports. We don't really know yet what the new safety regulations are going to be. We know what the drafts were that Dr. Trontell alluded to in terms of what that is going to do. Is it going to sort of reduce some of the noise by not requiring an non-serious labeled

except in certain instances? But is it also changing the definition of serious an unexpected where we might have more cases? So, I think we also are sort of poised here for some change in what the distribution of cases in AERS is going to look like once we get these new regulations and we don't know what they are going to be. I am sure you won't be able to tell us what the timing on that is but I am going to ask anyway.

DR. TRONTELL: The FDA is still reviewing the comments on the very extensive proposed SADR rule and I can't give you a time.

DR. GROSS: A number of suggestions have been made and will continue to be made over the next day. Maybe one thing that would be a good idea would be, when this committee meets again, to get a follow-up on our recommendations. It also would be interesting to look at examples of problems that have been identified and time periods until decisions were made, and what was involved in that. I don't think we have that information before us and would need it to give you more help

and more specific recommendations as to how you could supplement AERS right now to get to an answer more expeditiously, if that is possible. Robyn Shapiro?

DR. SHAPIRO: I am really sorry but it just is, to me, counter-intuitive to think that if you didn't have more and better reports from people that you wouldn't have more to work with in a more timely way. It is just counter-intuitive to me that that is not true.

So, another example, forget the doctors and the busy practice, how about the hospitals? They must report now to the OPR whenever a patient is about to die, and they do that. Why do they do that? They do that because they are going to lose JCAHO accreditation if they don't and/or subject themselves to a survey by the state on behalf of CMS. So, why can't those be sticks that could be used to at least enhance the number of reports about which there could be response to assure the quality of those reports? I just don't get it.

DR. GROSS: To echo what you are saying, I

think it is a reasonable idea. I think one of the differences between what you are proposing and what is already on the books is that very often, let's say mandatory reporting of various infectious diseases to state health departments is--well, it is mandatory but there theoretically a specific diagnosis has been made. Here, in what we are talking about, when an adverse event is reported we are not sure whether it is causally related to the drug or not so we would just have to define what gets reported and not say that a causal association is necessary before a report can be made. Elizabeth?

DR. ANDREWS: I think from what I have heard today and what I have observed in my experience is that the spontaneous reporting system is good only for--mainly for detection of rare and serious adverse events. So, I think the quality is not enhanced by volume of reports nor by making it mandatory but by sharpening the focus so that for those events that do represent real signals, they are identified earlier with richer detail for the

case review, and the system is not encumbered by a lot of superfluous activity and requirements that add little to our knowledge. The system is good for events that have been identified as possibly related to an exposure, which means that it really isn't good for a number of other things, as was alluded to before, such as events of long latency, events that may be recognized by the physician or the prescriber of the medication. It isn't particularly useful, despite data mining techniques, for quantifying the evidence, looking for risk factors and particularly high risk populations, although some evidence may be found from case reports.

But I would echo the comments of others that I would like to see resources devoted to looking at other ways of identifying signals that can be identified with the spontaneous reporting system and quantifying using the resources that are now available that weren't when the system was originally in place. So, my recommendation for improving the system is probably making it more

sharply focused and spreading resources into areas of greater need.

DR. GROSS: How would you make it more sharply focused?

DR. ANDREWS: I haven't reviewed the regulations that Dr. Stemhagen was just mentioning, but I know that quite a lot of effort goes into processing non-serious and labeled events from which we learn very little. So, I would focus on the rare and the serious events and probably that is about it.

DR. GROSS: Okay. Art, did you have a comment?

MR. LEVIN: I have a few. Robyn, much as I am a fan of mandates, with all my experience with mandated systems, they are sort of voluntary systems at the end of the day because the ability to enforce mandates is really difficult and you end up with all the same problems, maybe to a slightly different degree but you have tremendous underreporting and you have tremendous problems with standard definitions, and all of the things

that come in a voluntary system also come in a mandatory system.

I guess another thing--I am still going back to the beginning here and I think Lou's point is well taken. I guess thinking about what Syd Wolfe said, maybe we have to learn from the failures as well. I mean, it would be helpful to understand where the AERS system has failed to identify serious harm or the potential for serious harm for patients and then sort of do a forensic on that and say, you know, could we have done anything and, if we could, what would we have done to make the system more capable of success rather than failure in that particular case. Whether people think that is a valuable exercise, to sort of look back and do autopsies, as Syd suggested, with an eye to saying could we figure out how that could be improved or not. If we can't figure that out, why would we want to put resources into the system, any more resources?

DR. GROSS: It was suggested to me that it is time to put our nickel down for specific

recommendations for the FDA. Anne wants to say something before nickel time.

DR. TRONTELL: Does that mean I get to speak for free? I would actually like to ask the committee members, Lou and Art and Dr. Manasse as well. It is not entirely clear to us where we would look and what we would call failures. I think there is any of a number of metrics that might be applied to say at what apparent pace has the agency taken action on a safety problem. On a very crude level, back to the technical issues in front of us today, some of the greatest challenges for the agency in taking regulatory action is when the cases are muddy; when the background rate of the adverse event in the treated population is unclear or debatable; where we have less than complete information. If it were clear and obvious, I would like to think the agency acts in due course on such information.

So, process questions aside, if we were to bring back some root cause analysis would anyone volunteer instances where they might think they saw

the system in apparent failure?

MR. LEVIN: Let me say failure may be too strong a word. I think it goes back to Curt's point of timeliness, the inability to work in a timely fashion, therefore, putting lots of people at risk. Would there have been a way to shorten that time-frame and put less people at risk? So, sort of system failure, not Failure with a capital "F".

DR. GROSS: So, for number 1(a), the AERS system is most effective for pointing out serious and rare events; 1(b), is--

DR. MITCHELL: And, Peter, acute.

DR. GROSS: Acute, okay. Acute, serious and rare events, and where it is less effective is for adverse events that have a long latency period, such as cancer, or high background rates, such as heart attack. Anybody have any additions to those two before we move on?

DR. PLATT: Maybe just saying everything else.

DR. GROSS: Everything else.

DR. HENNESSY: Do we need to be explicit that we are talking about signal generation and that they usually, although there are some

exceptions, don't provide convincing evidence of a cause-effect relationship, or is that inherent in the question?

DR. GROSS: Maybe we will get to that a little later. Item 1(c), what can the FDA do to improve the AERS system? Unfortunately, I didn't keep an inventory but from memory I will just mention a few things, and then I would appreciate it if you would add what I have left out. One of the suggestions was that there be mandatory reporting of serious adverse events at the practitioner level or at any level.

The second was to review past examples of safety signals that have been identified and then acted on and review the time sequence of that, and see whether any suggestions could be made to shorten that time interval before a decision was made.

The next suggestion was that there be

regional reporting and perhaps a pilot program be tested. What else has been suggested?

DR. PLATT: Perhaps you could go round the table so that we could sort of give a quick restatement of the things we think are most important.

DR. GROSS: I don't want to, you know, restate everything. I mean, if anybody wants to add just raise your hand. Richard?

DR. PLATT: It seems to me that FDA could do three things to improve, I would say, the cost effectiveness of the system. One is to discontinue the attempt to use it for everything other than serious and rare events. The second is to move a lot of the very thoughtful work about signal detection to much more suitable data environments which could be good receptacles for the very thoughtful work that is being done. Finally, to concentrate a smaller amount of its resources on a much smaller target of signals. And, since we didn't say it under 1(a) it should really be confined to signal generation. I think there

should be no attempt to quantitate signals. As soon as there is a concern of a serious event coming in through the system, I think FDA should move it to a defined population environment.

DR. GROSS: Anyone else?

DR. SELIGMAN: Could I just probe you a little bit, Rich, on your first recommendation? How would you stop using it for all things other than the acute, rare and serious?

DR. PLATT: I suppose I would, first of all, change the request, please, only send us these kinds of things. Then I would try to develop a filtering system so that among the many you will receive anyway you can acknowledge them but not devote resources to them.

DR. GROSS: Ruth Day?

DR. DAY: About the regional reporting, Henri, I would have some reservations about whether that would delay the timeliness of being able to use the information so I would just like to put that in as a potential reservation since you listed it, Peter.

DR. GROSS: Okay. Annette?

DR. STEMHAGEN: What Rich was talking about, again, we don't know what the regulations

are going to be, but trying to get the noise out of there of the non-serious, labeled--not only the FDA resources but industry resources as well, we spend a lot of time on a lot of the reports that really aren't giving us the serious events. I know there is a lot of discussion in the new regs about less effort required on those reports but I think that is really going to be very important.

DR. GROSS: Lou?

DR. MORRIS: Is there any formal system for grading signals? I have heard about strong signals, weak signals but is there any formal system in which FDA grades a signal?

DR. AVIGAN: The answer is no because it turns out that different kinds of safety problems are distinct in their character. It is very difficult to create sort of a one-size-fits-all with regards to is the problem the number of cases that was seen; is the problem the severity; is it

susceptibility of a particular population. So, strength of signal is different for different kinds of safety problems. So, we do have operating principles based upon how we put a case series together and how we try to sharpen our assessment of where the risk may lie. We also have to, to some extent, articulate what the uncertainty risk is as information is coming to us. Part of what we are dealing with is risk uncertainty.

One of my concerns about precluding using AERS as a quantitative measure in all cases is that we deal often with an interim period where this is uncertainty. We see a signal. There is a clear problem at hand, but we don't have a grasp of the extent of the problem or its depth and we need more experience to know. And, this is a period which may be different in different cases, and I think this was alluded to by other speakers. It is important to try to shorten that period as much as possible with various methodologic strategies so that we do come to an answer so that we can make a disposition. But the case of valdecoxib, as we

heard this morning, was a case where the reporting rate, which was basically passive data with some denominator of usage, was so extraordinarily different than in the other cases that we couldn't make an informed judgment based on an imprecise measure.

So, again, I think that one of the hardships that we face as analysts is this kind of having to make regulatory decisions in a data environment where over time the emerging risk is potentially becoming more sharply focused but there is a necessary time lag, and we have to develop a strategy for how to deal with that emerging period.

DR. MORRIS: This morning I was talking to someone outside the panel and I heard that, I believe it was terfenadine and Stevens-Johnson syndrome--that there was a two-year lag between when the first two cases were identified and when there was another case. It seems to me that even though it is infrequent, it is so severe that those two years should not have occurred and that type of signal should have been worked up much faster

rather than just passively waiting for another signal. And, if there was some system for grading it, not in terms of uncertainty but in terms of severity, that might at least prioritize where we think the problems lie.

DR. SELIGMAN: I think the issue of how to grade and/or respond to signals has been well discussed and documented both in U.S. and internationally. I think the best thing you find is probably in the recent guidance document that was published and provided where we detail those safety signals that warrant further investigation. We were pretty explicit as to the situations where we think such investigation--

DR. MORRIS: But don't you think it is pretty broad? I mean, that is pretty broad.

DR. SELIGMAN: What is pretty broad?

DR. MORRIS: The guidelines.

DR. SELIGMAN: Well, it says more unlabeled adverse events, especially serious.

DR. MORRIS: Well, within that--

DR. GROSS: I am going to interrupt. I

don't think we are going to resolve this right now. One more comment from Elizabeth and then we are going to move on to active surveillance.

DR. ANDREWS: I was going to echo the comment that for the improvement in the system. I think streamlining to focus more on the acute, serious and rare events would make a lot of sense; that the effort that is considerable and a very good effort that has been devoted to methodologies for looking at large databases be applied to more systematic review of quantification of events looking for risk factors and identifying populations at high risk; and looking for systems that can do a better job of identifying signals relating to events of long latency.

DR. GROSS: Let's more on to the active surveillance section. The first question is how can active survival systems be used to augment the currently available FDA systems for safety signal detection and risk characterization? Anyone want to comment? Allen:

DR. MITCHELL: This is again where my

confusion came up earlier because it seems to me that there has been a lot of discussion today about what are considered active surveillance systems, but without any real reference to historically productive systems, you know, whether they are based on computerized data or based on case control surveillance approaches. It seems to me it is difficult to respond to this question if, in fact, active surveillance is the larger palette of colors that is available or should be available to the agency. So, it is hard to respond without the benefit of the discussion tomorrow and I just find myself in a difficult situation.

DR. GROSS: Jackie Gardner?

DR. GARDNER: Maybe I take a broader view of active surveillance. I guess it seems to me to dovetail with the conversation we have just had it, and that is, David told us this morning that active surveillance requires a hypothesis to initiate it effectively. So, maybe this is where things come together. After the first two Stevens-Johnson syndromes with terfenadine maybe we should have

gone or could have recommended going to some kind of, however defined, active surveillance, prospective, current system based on exposure and looking forward to try to pick things up faster in the available databases that we have heard about and that will only get better with increased prevalence of electronic medication records.

DR. GROSS: Yes, Sean?

DR. HENNESSY: There are a number of diseases that are commonly due to drugs and perhaps ongoing case control surveillance for outcomes, like Stevens-Johnson syndrome, agranulocytosis, liver disease, would be fruitful in detecting drug causes of those events more quickly than is currently done.

DR. GROSS: Richard?

DR. PLATT: I think that it would be helpful to be clear about the use of the term active surveillance. I think it is a problem for us. I think there is tremendous opportunity to do systematic questioning, often in the absence of hypotheses, in defined databases that I think by

the definitions we are using is called active surveillance. I think there is a great future for that. I think the active identification of cases by the stimulated activity of specific individuals has yet to be shown to be very useful in a general sense. I think that we would have to know a lot more to know that it is worth investing resources except in very limited circumstances. What might those circumstances be? Specific quantitation of event rates in known high risk situations. But I think that as enthusiastic as I am about actively using new resources that are available, I am skeptical about the value of putting time and effort into trying to train observers to work in open environments to detect more cases of events of interest.

DR. GROSS: So, we all seem to agree that this is an approach that aborning and we need more information.

DR. SELIGMAN: Rich, I believe you said earlier that defined population data resources is the future?

DR. PLATT: Oh-oh, I am being set up here.

DR. SELIGMAN: I just circled that quote. I just want to make sure I got it right.

DR. PLATT: Well, I will stick my neck out further and say I think it should be a big piece of the present.

DR. GARDNER: Richard, could you explain what you meant about provoked individuals and whatever you said?

DR. PLATT: I think that it is difficult to ask individuals in the course of their normal activities to do the kind of attribution that would ordinarily be needed to provide useful case identification that would allow us to do reasonable epidemiology studies. I would be concerned that the limits on that kind of case detection will tend to force us into channels of confirming things we already know. So, absent really controlled circumstances, like clinical trial environments, I think that doing active surveillance for cases of interest is going to be problematic.

DR. GARDNER: So, are you making a

distinction between stimulating reporters to collect more information about individual cases from faster utilization of databases in which information is going in routinely, such as your own? Are you thinking that active surveillance doesn't happen using your databases?

DR. PLATT: So, that is the definitional problem that I think we have using the term active surveillance. If we say passive surveillance is what AERS is, that is, spontaneous reporting, I think there is a tremendous leap forward the agency could make by making more use of data that is collected now during the routine delivery of healthcare and making thoughtful use of that information.

DR. GARDNER: Exactly my point too.

DR. PLATT: But I am contrasting that from the idea of saying active surveillance is training individuals to go out and look in healthcare environments for cases and collecting them individually. I think the effort would be much better spent in trying to understand how to use

routinely collected healthcare data to identify potential cases which could then be reviewed in systematic ways.

DR. GROSS: Ruth Day had a comment, and then we will move on to 2(b).

DR. DAY: Just a brief comment. We are talking about active versus passive systems as if it is a total dichotomy, and I think there may be a continuum here because we did hear this afternoon about some active systems which still have the element of voluntariness in it, say, in emergency departments, and so on and so forth. So, I don't think it is a clear dichotomy but there are places along the way, and perhaps someone at the FDA or elsewhere might figure out what the appropriate variables are in order to then scale the different tools along the continuum so that appropriate selection could take place as events warrant.

DR. GROSS: Item 2(b) is what types of drug products or safety problems are best suited to active surveillance methods? Anne Trontell?

DR. TRONTELL: I think we were

deliberately a bit broad in defining active surveillance. It may actually lead in to this current question, which is are there particular settings or networks, as Dr. Hennessy was suggesting--are emergency departments a good place to pick up agents that cause anaphylaxis since that is where you have relatively enriched presentation of such cases; liver failure that might present to transplant centers; some of the examples that we heard this morning. If people might speculate or think of other systems--pheresis centers, as we heard, for people who present with TTP. There may be places. The question logically arises in many clinician's minds that this may have an extraneous source and efforts are made to look for a drug-related cause.

DR. GROSS: Curt Furberg?

DR. FURBERG: I think our concern obviously is underreporting and the active surveillance is clearly a step forward. I agree that the focus really should be, as stated by Richard and others, medical records and physicians

but we shouldn't forget about the patient as a source of information. There are studies showing that the yield may increase by a factor of ten. Studies are reported in the literature that you pick up more events. Maybe the additional ones are less severe or I don't know, but it is worth some effort to do that concurrently with whatever you are doing so that we have a good sense of how much we are missing.

DR. GROSS: I sense a confusion over definitions. I wonder if someone would distinguish active from passive surveillance and be a little more specific than we have been so far. We all seem to think we know what we are talking about but we kind of blur the two. Passive surveillance involves active intervention to get the report and active surveillance often can be passive because it is voluntary. So, anyone? David?

DR. GRAHAM: I think that is part of the dilemma. We don't have a good definition for what active surveillance is. Passive surveillance is basically AERS and what we identify and capture

from that, and what are the purposes for which AERS is intended and its application in terms of postmarketing safety. Then the question is what other ways can one use to identify safety signals or to go beyond identifying them to preliminarily quantitating them. It would either be an adjunct to a passive system, a complement to it or something that could actually do better than it does it, or that could be done in sort of an ongoing, prospective way. You can view the passive surveillance as sort of a radar system. It is continually on, 24 hours a day, 365 days a year. The reports are coming in and the question is, you know, when do you see something that you think is a signal. So, there is no clear definition about what active surveillance is, what parameters you would use to define it, what domain within which one would conduct it. That is one of the reasons why it is there for the committee. We have struggled internally with a definition of active surveillance and basically, for myself, I see it as something that can either complement passive

surveillance or be an intermediary step between passive surveillance and an actual hypothesis testing situation.

DR. GROSS: Could we reduce it to this, that passive surveillance depends on a voluntary effort and active surveillance is involuntary in the sense that you are data mining a system where various information is automatically collected and you are trying to make decisions whether there are signals there or not?

DR. GRAHAM: I think that that would be one approach, passive being things that basically come to us, in a sense, unsolicited and active being things that we do sort of in an intentional, prospective way, and it can be any of the things that you have just described. And, what defines active surveillance is probably very broad in terms of the different approaches that are possible.

DR. MORRIS: There is an analogy in consumer marketing where passive information search is what people process whatever comes to then; active search is when they go out and solicit

information. It seems to me that is what you are saying, that it is really a process of information gathering. In one case it comes to you; in the other case you are going out and soliciting. I think most of it is voluntary and I don't think active and passive are the key issues here. I think it is the value of the information and the biases in the information more than how you solicit the information.

DR. GROSS: I have a suggestion. Just like the Institute of Medicine struggled with a definition for performance measures and outcomes and structure and process, it might be worthwhile for the FDA to struggle with a definition of passive surveillance and active surveillance so when we have to discuss it we have an agreed upon definition that we are talking about. Allen?

DR. MITCHELL: I think we are on the right track but it seems to me we are just hung up on semantics. It really seems we are talking about spontaneous reporting versus some kind of directed effort, and that directed effort could be something

utilizing existing data sets, data sets going forward, or--and I find it missing from the entire conversation--data sets that were, in fact, created for the purposes of drug surveillance.

I do feel the need to mention case control surveillance which had its origins, as Susan Jick pointed out, back in the early '70s when reviewing in-hospital data but data that included histories of use of medications prior to admission. Dr. Hanenman, from Finland, observed that among chronic aspirin users there was a deficit of MIs. That observation prompted a 26-hospital study which Dr. Dennis Sloane coordinated, which was a case control study specifically focused on MIs and appropriate controls, with the focus being exposure to aspirin prior to admission. That was the first real demonstration of the protective effects of aspirin. But apart from the finding, it also began what became case control surveillance, which is an activity that a number of groups have taken part in, where you don't necessarily have hypotheses and, in fact, you typically don't. So, so you can

focus case control surveillance on rare diseases such as agranulocytosis or aplastic anemia or anaphylaxis, the way Sean mentioned. You could focus it on birth defects, which has been a good part of my career. You could focus it on admissions for MI, cancer, any number of outcomes and, at the same time, survey, if you will, the vast range of exposures that may be related.

Those kinds of studies, interestingly enough, had FDA support--I am going back now--in the 1970s. FDA really saw the value in that approach. FDA's budget and priorities changed and Lynn Rosenburg, in our group, who has been championing this in recent years sort of focused the effort on cancers with appropriate control groups and has made a huge number of findings, many of which were first-time findings and many of which were corroborated either in clinical trials or other data sets.

So, I find it a little concerning that we are focusing an awful lot of attention appropriately on clinical data sets, if you will,

databases that are derived from medical care, but not focusing on data sets that were specifically developed for the purposes of evaluating drug safety. Whether you want to call it self-interest or public health interest, I don't want that to get lost in the conversation.

DR. GROSS: Since we have an answered 2(a) and 2(b), I am going to read 2(c).

DR. MANASSE: Peter, before you do that I would like to address something in 2(b), if you don't mind.

DR. GROSS: Well, let me read this and then we will be thinking of it all together. How might active surveillance systems for drug safety problems be used most efficiently, that is, with greater specificity and sensitivity? Henri?

DR. MANASSE: Well, I wanted to specifically look at what types of safety problems are best suited to active surveillance, and I think it also ties to this issue of specificity and sensitivity, and direct our attention at the outpatient drug benefit for both the elderly and

dual eligibles, that is, Medicare and Medicaid patients that goes into effect January 1. So, some 35 million Medicare eligibles will have access to pharmaceutical agents and be paid for in part by the government. The dual eligibles, that is, elderly Medicaid patients are going to be shifted into that program as well. There is planned on-line claims adjudication that will immediately, at the pharmacy level or the dispensing level, determine eligibility, co-pay and all those typical adjudication issues, and there is a component to the law that directs the movement to electronic prescribing. All of that says to me that that population as a problem population with respect to the use of medications is perhaps ripe for very careful prospective determination of how active surveillance might be laid over that entire program.

Now, this begins January 1. CMS, as you might imagine, is working very, very hard to get this put together, and it seems to me to be timely and necessary to put together a very deliberative

active surveillance program so that we can build a better safety net through signal detection in that particular patient population.

DR. GROSS: So, you are adding demography to this, which I think is a great idea and now is the time to do it. Anybody else? Richard?

DR. PLATT: I am thinking of a reply to Anne's question about focusing surveillance in specific areas. My enthusiasm for that would be very high if it were possible to have systematic solicitation from the providers about their attribution of events as possibly being drug related. As an example, we are testing a system, an EMR system whereby pediatricians who see children within two weeks of immunization and who enter a diagnosis that is not on the white list are prompted to say could this diagnosis be related to a vaccine adverse reaction. If the answer is yes, then the question is do you want to submit a VAERS report?

So, it would be the "systematicness" of the solicitation that would make focused

surveillance in specific environments worthwhile. My concern is that if all you do is say we are interested in those environments but wait for the clinicians to initiate the connection, you basically still have a passive system, a passive system in a few places.

DR. TRONTELL: If I can comment, I think that is a very constructive remark. Perhaps it wasn't emphasized in my recollection from Dr. Mendelsohn's presentation but, in fact, the pilot work that was done by him and Budnitz was to alert the coders to the terms used by clinicians in describing what might be an adverse event side effect--look for rash; look for anaphylaxis--getting again at some of these settings where in the drug-induced liver injury network the clinicians are sensitized to the liver being a not infrequent target of drug-induced toxicity. So, in an individual who doesn't have an otherwise readily explained cause for their liver injury or liver failure they do an active systematic seeking of other potential causes that

could be drugs.

DR. PLATT: I am just concerned that merely telling physicians in general that this is something to be aware of will give sort of uninterpretable results. That is my concern, that we should find a way to prompt physicians in real time, and I think that the technology is taking us to a place were we can do that.

DR. GROSS: We have a lot of expertise around the table here and I am going to drag the answers out of you, no matter what! How about if we take Henri's approach? He defined demography, the elderly, Part D, to set up a surveillance system. Are there any other demographic characteristics or groups where you think we could set up a surveillance system because, you know, certain safety issues may be more common in one demography than another? Jackie?

DR. GARDNER: Yes, and it is one that we have dealt with in the past and it came up again today, and that is the VA. The VA has had an extensive linked surveillance system for a long

time. When this committee, working with another, spent a great deal of effort looking at long-acting opioids we kept saying, well, if there is ever a population in which we might have experience with this it would be the VA. But we didn't have VA data. This morning it came up again. We hear again and again, and I am sure it is true just as with CMS, that there are logistic issues and perhaps even cross-agency issues, and if there is anything this committee could do in a recommendation that would help overcome some of those logistic and cross-agency barriers, I would like to suggest that we try to find a way that we can help do that with both CMS and VA because that has to be more cost effective than generating new databases. Although I do support Allen completely because I think that is an excellent system.

DR. GROSS: Let me ask the FDA what healthcare databases--if you are permitted to answer this, what healthcare databases do you have access to, like some of the big HMOs? I won't mention any names but can you tell us that?

DR. TRONTELL: I think we can describe them without necessarily giving names, though we certainly could do that if requested. We have

population-based databases where we have administrative claims data and drug exposure. In those we have varying degrees of access to medical records to validate those outcomes and exposures. In addition to those generalized resources, we also have done some work with longitudinal or electronic medical records. We have access to one large database where we are learning the ropes of that particular system so we can look, again, at drug exposures and outcomes, and also look at it over time and others where, in fact, some of the ancillary data may be available where there is actually digitized diagnostic information, laboratory data. So, the databases vary somewhat from one situation to the next in how deep their data might be that is readily accessible electronically. We have a number; we are exploring what other ones we might also add. Is that enough to give you an answer?

DR. GROSS: Yes. So, as Allen mentioned earlier, maybe we can't answer this question completely without the presentations that are supposed to be made tomorrow. I guess NEISS is also another system that may give you some useful information on active surveillance.

I think we have kind of gotten out of this all we are going to get out of it this afternoon. Why don't we go on to drug utilization? Based upon the presentations today, what are the priority areas for FDA to expand or improve its use of drug use data?

You know, as we go through these questions I think it is going to be very important for this group to try to answer the specific questions. We have a lot of expertise, a lot of general ideas but the FDA is really asking for specific answers and we are going to be most useful if we can give specific answers. Do you have a specific answer?

DR. MORRIS: I have a specific question.

DR. GROSS: Well, we have a lot of questions. I think we need some answers.

DR. MORRIS: Well, it will lead to an answer.

DR. GROSS: If you will give the answer after the question.

DR. MORRIS: It is for Allen or Richard. One of the big differences in terms of types of active systems is whether you use existing databases which have ICD-9 codes versus setting up specific registries or something like that where

you get much more specific information about adverse events, or at least can get it even with passive systems where you have adverse event reporting. My question is how specific does the information need to be to be useful as a surveillance system? Will an ICD-9 code give us enough information that we don't need to worry about MedDRA type codes, or do we need both?

DR. MITCHELL: Well, in case control surveillance you typically start out with ICD-9. You are not relying on anyone taking the initiative to report. You will typically go to a hospital and identify, by going through their records, patients

with X, Y or Z diagnoses. It can be on an ICD-9 code or it can be based on, as we have done for certain neonatal conditions, a review of neonatal intensive care unit logs, with subsequent confirmation. I think that is highly effective. But I think you can use any number of approaches. What you just don't want to do is have, with whatever approach you are using, a lot of contamination by misclassification. You need some assurance that the diagnostic entity that you are studying is, in fact, what you think it is.

DR. MORRIS: So, if we are just looking for signals still, if we have an ICD-9 code as a signal, what I am hearing is that is specific enough to allow you to make the decisions you need to make to work up that problem further, and you don't have to get a MedDRA code to say I have a signal.

DR. TRONTELL: The ICD-9 codes actually vary, not in a statistical sense of specificity but, for instance, there is no ICD-9 code for Torsade de pointes, which is of concern if you

wanted to look for Stevens-Johnson syndrome. The ICD-9 is not fine enough in its coding to distinguish those, but that is clearly the mechanism that is used for reimbursement and that is why these data systems--it will depend on the particular safety problem that you are looking at.

DR. PLATT: I think that ICD-9 codes are clearly not sufficient for certain important kinds of questions. So, we have typically gone to registries or case control studies for that. There is an important convergence though that I think is coming from the fact that electronic medical records are able to serve many of the purposes that we have assigned to registries. So, I think it is a very rapidly shifting terrain in terms of what will be available for use by the agency.

DR. AVIGAN: Can I just expand that question because I think that is a very important, insightful question? From an operational point of view, if the medical record adjudication is a critical step in winnowing away the chaff from the wheat and you are only interested in one percent of

what the ICD-9 code actually indicates, then that requires a fair amount of medical record review. If this becomes sort of the framework of FDA signal detection, then we have to make sure that we are embarking upon a general procedure that will be workable from the point of view of manpower and also from the point of view of accessibility of medical records, with all the issues around access to medical records which is implied by kind of a general national surveillance system.

So, I think the idea is very interesting but I would ask the panelists to consider this paradigm from the point of view of practicality and also what the hurdles are around that.

DR. GROSS: So, we now know anybody who wants to identify rare and unusual events will not get satisfaction from the ICD-9 coding system. I have had that experience multiple times myself.

Number three, priority areas, anyone have some suggestions for what the FDA should focus on? Henri?

DR. MANASSE: We heard today that the

Medicaid program has utilized drug utilization review. If I recall correctly, there are three major objectives in that effort. One was to find outliers with respect to prescribing practices that may have been inappropriate. The other was to take a look at certain drugs that may either have been high safety risk, high expense or otherwise problematic. Thirdly, to use these data to help improve practice.

I think that the experience, at least that has been published, has been quite mixed depending on how states manage these. But, in effect, the DUR programs have aided the Medicaid programs to, on the one hand, perhaps save money and, on the other hand, perhaps improve some level of quality. It seems to me that the FDA has an ethical obligation that when it is collecting data and signals that may not be the basis for pursuing the withdrawal of a drug from the marketplace to use that data to improve practice and to enhance safety.

Now, that may be a new sort of philosophy

but I want to interject it in this dialogue because as we improve the signal detection systems, as we move to these active surveillance systems, if we are picking up information that could be quickly distributed to the practice community, and that data used in rapid cycle improvement, we may even get as good as what aviation does with signal detection.

You recall that when there was a rudder problem on the 737, Southwest Airlines quickly pulled all their aircraft back, inspected all the rudders and, on assurance that the rudders were fine, put the aircraft back in the air. It wasn't because a 737 crashed. I would advise that we begin to look at how this data can, in fact, be distributed to the practice community. Miss Holquist, for example, gave that presentation this morning about a lotion being mistaken for either eye drops or ear drops. If every pharmacist in every pharmacy in this country knew of that particular problem, I am sure that they would have a little discussion with the patient at the

counter. So, I would like us to think about that.

DR. GROSS: Well, from your mouth to God's ears that we should be as good as the airline industry and as error free! Any other comments? Allen?

DR. MITCHELL: I couldn't go back to the office if I didn't speak in favor of a priority--not the priority--being the safety of medications in pregnancy, specifically with respect to birth defects. It is approaching 50 years since thalidomide and we still do not have any kind of system in place in this country to systematically identify human teratogens. So, I just put that out.

DR. GROSS: We have a lot of work to do. Any other priority areas besides teratogenic drugs?

DR. PLATT: I would like to second the notion that job one ought to be getting effective access to CMS data, not just to drug exposure data but to fully linked CMS claims data. I think that would be a transforming event for our society. So, I think if you have to set priorities for FDA, I

think that would be at the very top of the list.

DR. GROSS: Good comment. Anyone else? Ruth Day?

DR. DAY: I am not sure whether this comes up under number three or number four, but I think special attention to OTC switch products would be useful. There are drugs that used to be prescription and are now, or will be in the future, becoming OTC. Although a lot may be known about adverse events, we now have a new factor of self-selection and whether consumers understand whether it is all right for them to take these products or not, and now to use them safely and effectively. Since it is OTC, it is a little harder to track but I think that this is an important area to consider, especially within the first couple of years of transfer from prescription to OTC status.

DR. GROSS: How would you marry OTC drug use and serious adverse events?

DR. DAY: I think some new tools are going to have to be developed for that, and direct

contact with the patient because the physician isn't going to know, and so forth.

DR. GROSS: Yes, they often don't tell us about OTC drugs.

DR. MITCHELL: If I could just respond to that question, I am going to sound like an old saw but there are case control surveillance, in particular, queries of OTC use as well as prescription, as well as herbals. So, there are already systems in place that can speak, and have spoken, to OTC use. I agree that it is an important area.

DR. GROSS: Would you comment on those systems?

DR. MITCHELL: Pardon?

DR. GROSS: Comment on the OTC drug use systems?

DR. MITCHELL: Again, in case control surveillance, because you are in contact with the patient--it has its limitations, of course, but it also has its strengths because now you have the final common pathway for all exposures. So,

patients who have MIs or who have cancer or who have fractures of the femur can be routinely queried at some point prior to admission about the wide range of both prescription and over-the-counter drugs, and over-the-counter drugs can be quite broad in this definition. In that way you can identify issues of risk related to, let's say, non-steroidals. You can identify, as was done, issues of safety or protection related to aspirin in MI. It is really methodologically no different, except that is has a lot more statistical power because the drugs tend to be used more commonly, many of them, particularly when they are switched so when they go from prescription to OTC the use tends to bump up fairly dramatically. So, there are approaches extant that speak to that, and that is clearly a limitation, as you point out, Peter, what physicians systematically screen patients for OTC use and then record it.

DR. GROSS: Yes, Curt?

DR. FURBERG: I would like to see more formal interactions with regulatory agencies in

other countries. I don't know what is going on. It is not transparent; it is not communicated. How do we hear about it? Well, we read in the newspapers that in Europe they have withdrawn a drug, we hear it and we don't understand why. I think that should be brought up with a committee like this one and we should have a broader discussion about those issues.

DR. GROSS: Good point. Anyone from the FDA want to comment on whether that is being done now or anticipated?

DR. SELIGMAN: At the agency level there is an Office of International Programs that is responsible for regular communication with regulators. In addition, we have within the Office of Drug Safety regular video conferences with the EMEA as well as Canada, Australia and New Zealand. These are often focusing on emerging drug safety problems, cases that were seen, and opportunities for exchange of information. We don't really formally publicize these meetings because they are, indeed, pre-decision with an opportunity to sort of

share information. But those kinds of interactions do occur quite regularly.

DR. GROSS: Sean?

DR. HENNESSY: An important use of drug utilization data that we haven't talked about today but does goes on and remains important is evaluating the effectiveness of programs to try to improve prescribing, sort of process measures. For example, studies showing the effectiveness or ineffectiveness, as it were, of programs to reduce co-prescribing of terfenadine with drugs that it interacts with, and cisapride and drugs it interacts with.

DR. MORRIS: Following along that same vein, one of the things we heard today is the NEISS data and the concept that three-quarters of hospitalizations due to drugs are due to unintentional overdose. I don't know what that means. I am sure that there are lots of explanations but in terms of drug utilization data per se, I would suggest we don't need to know more basic information that the quality of drug use data

that we heard about today probably exceeds the quality of the numerator data that we would use for it. But more specific data on how people use drugs and how that results in some kind of negative outcome I would find very useful because, as Ruth said, getting back to the behavioral side, and self-selection issues, and remembering issues, and getting to quantify that I think would be very useful in terms of patient safety.

DR. TRONTELL: Can I just comment on Lou's remark? Bear in mind that the data that Dr. Mendelsohn reported was of a pilot. So, the representativeness of those particular findings is not--

DR. MORRIS: But that is a nationally representative sample.

DR. TRONTELL: But what was given there was from the pilot to expand it to a nationally representative sample. That is now in the process of being analyzed.

The other thing to bear in mind is that the classification scheme, as we work with CDC, may

undergo some refinement and unintentional overdose might include individuals who have an excess of pharmacologic action that may not necessarily represent an overdose of the product, so an individual who has excessive action of warfarin, and so forth.

We don't want to belabor it. That was a study to show that by educating the coders we would do a better job of picking up these events.

DR. MORRIS: I think we need more information.

DR. GROSS: Robyn?

DR. SHAPIRO: You are not going to like this--

DR. GROSS: That is why you are here, to keep us on our toes.

DR. SHAPIRO: Okay, good! I would like to step back and take a look at what risk management and drug safety are about. For me, the two key components are minimizing risk and assuring that risk is reasonable in relationship to benefit. I mean, I think that the crux of the conversation

here is that we don't have sufficient ways to pick up signals in time to assure that we are minimizing risk to the extent that we would like to do it.

So, the FDA approves "me too" drugs and, by definition, the benefits are the same as some other drug already on the market. Yet, we are opening ourselves up to the possibility of unknown risks. So, is that risk with that "me too" drug reasonable in relation to benefit, which is nothing, which is no superiority? So, for me, the whole topic of approval of "me too" drugs is very relevant to this conversation.

DR. GROSS: Stephanie?

DR. CRAWFORD: I have just a quick comment. Another priority area that just jumps out is one of the settings where there is probably a lot of data and we just need some creative thinking to figure out how to incorporate it is inpatient data. Perhaps meeting with JCAHO representatives, ASHP representatives and some others--I really think is possible because the data is collected at those systems levels--just to think about what is

needed in a market basket or cohort of these groups. I think the inpatient side could really go up with some important lessons for us.

DR. GROSS: I think we are getting close to the end here. Are there any parting words of wisdom? If not, I want to thank you all for your excellent input. It has been a very interesting session--

DR. PLATT: I think we haven't talked about number four yet, at least some of us didn't realize we had moved to number four.

DR. GROSS: Well, I figured we hadn't answered any of the others so what made you think we were going to answer number four? No, actually some very good comments have been made; sorry to be cavalier. But if you have a succinct answer to number four--not a question but an answer, I would love to hear it.

DR. PLATT: Well, themes we have talked on today are three broad areas in which I think there are surveillance opportunities, and those are in the use of the burgeoning linked claims data

systems. I will stick by my prior comment that the CMS data system is the most important but I think that the constellation of federal data sets ought to be next on the list. That is, there should be every reason for you not to take no for an answer with regard to the VA or the Department of Defense or other such systems. Then, I think in the private sector there is a substantial number of data systems.

At the second level, I think that the installed base of electronic medical record systems is now large enough that it is worth the agency's attention and that will only improve. But I think you can do a substantial number--

DR. GROSS: I am sorry, what did you say number two is?

DR. PLATT: Electronic medical systems. The installed base is obviously nearly as large as it is for claims systems so you have less power but there are certain kinds of things that are now possible to do in the records of the millions of people who are cared for by electronic medical

record systems.

Then, I would urge you to take Curt Furberg's and Ruth Day's comments to heart about the possibilities of direct outreach to patients. I think the technology for automated interactive, computer-driven telephone systems is another technology that deserves a lot of attention. I think it would be very straightforward and inexpensive now for the agency, for instance, to do stratified outreach questions by telephone to thousands of individuals exposed to new drugs and acquire a lot of useful information very quickly and at low cost.

DR. GROSS: Thank you, Richard. That was very helpful. Allen?

DR. MITCHELL: I need to acknowledge that there may be a conflict here in terms of my interests but, in fact, back in '98 we began a random digit dial ongoing survey of the U.S. population to identify, in the previous week, exposures to prescription and over-the-counter and herbal products. That has been ongoing and has

been published, and the agency is aware of it but, again, it is using the consumer as the final common pathway to provide information on use, duration of use, as well as perceived indication and so forth. So, again, it is just a resource that is there in addition to others.

DR. GROSS: So, that is part of number three by Richard. Good! Well, thank you all very much. I will see you bright and early at 8:00, but don't go because Shalini Jain has one last comment.

MS. JAIN: Two quick things. You can't leave your valuables but if you would like to leave your meeting materials if you are coming tomorrow, we will be locking up the room so just take whatever you need in your rooms or whatever is valuable. In addition, anyone participating in the evening program, we will be meeting in the lobby at 5:45. Thanks.

[Whereupon, at 4:32 p.m., the proceedings were adjourned, to reconvene on Thursday, May 19, 2005 at 8:00 a.m.]

- - -