A Perspective on Quantitative Assessment of Clinical Benefit-Risk at FDA: What Needs to Change and How to Move Forward

Robert T. O’Neill Ph.D.
Director, Office of Biostatistics
Office of Translational Sciences, CDER

FDA-PhRMA-Bio Workshop on ‘Assessing Drug Benefits-Risk in Regulatory Decisions: Framing the need, assessing the tools, and deciding next steps,’ November 6, 7, 2007
Outline of talk

- Framing quantitative benefit - risk assessment within the drug development and review process
- Benefit / risk first should focus on the science of drug safety assessment - addressing the asymmetry in safety/efficacy in RCT’s
- Infrastructure needs
- Tools needed - current and future
- Concluding remarks on the way forward
Conceptual Framework for Quantitative Benefit / Risk Analysis

- Severity vs frequency (rarity)
- Treatment vs. Prevention
- Acute vs. Chronic exposure / usage
- Where does data on risks derive from: Within RCT’s or external to RCT’s (e.g. liver failure)
- B/R changes with life cycle of medical products in the marketplace
- Dealing with multiplicity of benefits/risks and competing events
Issues in setting the framework

- Metrics of Benefit
- Metrics of Harm (Risk)
- Metrics of Benefit and Risk considered jointly
- B/R assessment as a function of exposure time
- Population vs. individual patient B/R
- B/R assessment at time of approval vs. life cycle
- Role of RCT’s in the estimation of B & R’s
Three situations

- Benefits and risks observed in clinical trials - evaluations based primarily on RCT’s

- Benefits observed in clinical trials; potential risks observed outside of the trials and not quantifiable

- Benefits and risk change over time, with multiple usage, and emerging information
What is different about the objectives and collection of safety (harms) data

- Endpoints may not be as precisely measured or adjudicated as in efficacy trials where there are a few pre-specified endpoints
  - impact of sensitivity/specificity on estimates and comparisons
  - exposure time may be critical to onset of events (dose, cumulative dose, mechanism of action - liver damage)

- Safety events can occur after withdrawal from exposure - follow-up criteria in study can lead to informative censoring

- Multiplicity of events, recurrent events and multiple different events per subject

- Counting events - coding, dictionaries, adjudication strategies, body systems
Consequences

◆ Safety endpoints are measured, collected, or followed with less accuracy

◆ After the fact the endpoints may get adjudicated, when it is too late to obtain other information that may be pertinent to the adjudication

◆ Power, study size, strength of conclusions impacted

◆ Reliance on patient level profile evaluations to determine medical plausibility
Benefits are ascertained and reported differently than harms in RCT’s.

Addressing the asymmetry in the collection of benefits and risks in order to better quantify the net benefits:

- Censoring of patients who withdraw from exposure
- Measurement of delayed or late side effects
- In treatment trials of subjects with symptoms always there is differential exposure
Extreme lack of sophistication in analysis and reporting of safety signals and risk
A medical/journal culture problem that carries over into drug development

- Estimates of event rates: Proportion (%) of N subjects with the event
- Estimates of relative risk and risk factors
- After the fact endpoint definitions
- Events per unit of time (e.g., rate per 100 person years)
- Hazard rate, hazard ratios
- Cumulative incidence
- Risk factor modification of hazard/cumulative incidence
- Composite vs. individual endpoint contributions for time to event (delayed) and for acute outcome
The medical literature and editors recognize a need to improve the reporting of safety outcomes.

Better Reporting of Harms in Randomized Trials: An Extension of the CONSORT Statement

John P.A. Ioannidis, MD; Stephen J.W. Evans, MSc; Peter C. Gotzsche, MD, DrMedSci; Robert T. O’Neill, PhD; Douglas G. Altman, DSc; Kenneth Schulz, PhD; and David Moher, PhD, for the CONSORT Group*

In response to overwhelming evidence and the consequences of poor-quality reporting of randomized, controlled trials (RCTs), many medical journals and editorial groups have now endorsed the CONSORT (Consolidated Standards of Reporting Trials) statement, a 22-item checklist and flow diagram. Because CONSORT primarily aimed at improving the quality of reporting of efficacy, only 1 checklist item specifically addressed the reporting of safety.

Considerable evidence suggests that reporting of harms-related data from RCTs also needs improvement. Members of the CONSORT Group, including journal editors and scientists, met in Montebello, Quebec, Canada, in May 2003 to address this problem. The result is the following document: the standard CONSORT checklist with 10 new recommendations about reporting harms-related issues, accompanying explanation, and examples to highlight specific aspects of proper reporting.

We hope that this document, in conjunction with other CONSORT-related materials (www.consort-statement.org), will help authors improve their reporting of harms-related data from RCTs. Better reporting will help readers critically appraise and interpret trial results. Journals can support this goal by revising Instructions to Authors so that they refer authors to this document.


For author affiliations, see end of text.
For definitions of terms, see Glossary.
*For a list of members of the CONSORT Group, see Appendix 1, available at www.annals.org.
The culture of safety analysis and planning is missing basic statistical sophistication

◆ The Cox 2 experience as an example
Rofecoxib

A critique of the quality of reporting basic statistical concepts and characterizing risk

Two different messages from the original article and the follow on article
GI events were the primary endpoint - there are no issues with reporting, estimates of incidence rates and analysis of primary outcome, another safety endpoint - This is not so for other safety endpoints, eg. cardiovascular
Decreasing patients exposed over time

Cumulative incidence of primary safety endpoint

Figure 1. Cumulative Incidence of the Primary End Point of a Confirmed Upper Gastrointestinal Event among All Randomized Patients.
General Safety

The safety of both rofecoxib and naproxen was similar to that reported in previous studies.²⁰,²¹ The mortality rate was 0.5 percent in the rofecoxib group and 0.4 percent in the naproxen group. The rate of death from cardiovascular causes was 0.2 percent in both groups. Ischemic cerebrovascular events occurred in 0.2 percent of the patients in each group. Myocardial infarctions were less common in the naproxen group than in the rofecoxib group (0.1 percent vs. 0.4 percent; 95 percent confidence interval for the difference, 0.1 to 0.6 percent; relative risk, 0.2; 95 percent confidence interval, 0.1 to 0.7). Four percent
A follow-up article on the same study

Same data

Risk of Cardiovascular Events Associated With Selective COX-2 Inhibitors

Debabrata Mukherjee, MD
Steven E. Nissen, MD
Eric J. Topol, MD

Atherosclerosis is a process with inflammatory features and selective cyclooxygenase 2 (COX-2) inhibitors may potentially have antiatherogenic effects by virtue of inhibiting inflammation. However, by decreasing vasodilation

JAMA. 2001;286:954-959
Reported as .4% in the original article - best estimate is 1.8% for 12 months exposure

**Figure 1.** Time to Cardiovascular Adverse Event in the VIGOR Trial

<table>
<thead>
<tr>
<th>Months of Follow-up</th>
<th>Rofecoxib</th>
<th>Naproxen</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4047</td>
<td>4029</td>
</tr>
<tr>
<td>2</td>
<td>3643</td>
<td>3647</td>
</tr>
<tr>
<td>4</td>
<td>3405</td>
<td>3395</td>
</tr>
<tr>
<td>6</td>
<td>3177</td>
<td>3172</td>
</tr>
<tr>
<td>8</td>
<td>2806</td>
<td>2798</td>
</tr>
<tr>
<td>10</td>
<td>1067</td>
<td>1073</td>
</tr>
<tr>
<td>12</td>
<td>531</td>
<td>514</td>
</tr>
</tbody>
</table>

Relative risk (95% confidence interval) = 2.38 (1.39-4.00); *P* < .001. VIGOR indicates Vioxx Gastrointestinal Outcomes Research.
The continuing story of the time course of the cardiovascular events

A level of sophistication not planned for nor fully understood but with very practical implications
Time-to-Event Analyses for Long-Term Treatments —
The APPROVe Trial
Stephen W. Lagakos, Ph.D.

The Adenomatous Polyp Prevention on Vioxx (APPROVe) trial\(^1\) compared rofecoxib with placebo in the prevention of recurrent colorectal polyps, but the researchers also collected data on adverse events after the initiation of treatment. There are several reasons why using such windows might be desirable. First, events occurring during treatment or the subsequent window period might be
Two messages depending upon event definitions (Different disease – placebo control)

**Figure 1.** Kaplan–Meier Estimates of the Cumulative Incidence of Confirmed APTC Events in the Rofecoxib and Placebo Groups, According to the Intention-to-Treat Principle.

1 bars represent 95 percent confidence intervals.

**Figure 2.** Kaplan–Meier Estimates of the Cumulative Incidence of Confirmed Thrombotic Cardiovascular Events in the Rofecoxib and Placebo Groups, According to the Intention-to-Treat Principle.

1 bars represent 95 percent confidence intervals.
Figure 1. Hypothetical 95 Percent Confidence Band for the Difference, $I_{36}(t) - I_p(t)$, between the Cumulative Incidence Curves for the Rofecoxib ($I_{36}$) and Placebo ($I_p$) Groups, Constructed from the Results of the APPROVe Trial.

Differences lying partly or completely outside the shaded region are inconsistent with the data. Differences lying wholly within the shaded region include the following: separation of the cumulative incidence curves in the two groups at times both before and after 18 months, and consistently higher or lower cumulative incidence in the rofecoxib group before 18 months.
Figure 2. Logical Inferences about the Cumulative Incidence Function, \( I_{12}(t) \), for a 12-Month Course of Rofecoxib, Based on Known Values for \( I_{36}(t) \) and \( I_p(t) \) That Are Identical for 18 Months before Diverging.

If monotonicity is assumed, so that \( I_p(t) \leq I_{12}(t) \leq I_{36}(t) \), then \( I_{12}(t) \) must equal \( I_p(t) \) for first 18 months and be somewhere in the shaded region after 18 months. The lower edge of the shaded region corresponds to the absence of an increased risk with the 12-month course; all other scenarios in the shaded region correspond to an excess risk with the 12-month course that occurs only after the discontinuation of treatment. If monotonicity is not assumed, nothing can be inferred about \( I_{12}(t) \) beyond month 12; however, the 12- and 36-month courses are identical for the first 12 months, so that, all other things being equal, \( I_{12}(t) \) must equal \( I_{36}(t) \), and thus \( I_p(t) \), through month 12. Although drawn as separate curves to be visually informative, the inferences are based on the assumption that the cumulative incidence functions overlap for the first 18 months.
Figure 3. Statistical Inferences about the Excess Risk, $I_{12}(t) - I_p(t)$, Associated with a 12-Month Course of Rofecoxib, Based on the Hypothetical Results of a Trial Comparing a 36-Month Course of Rofecoxib with Placebo.

The upper edge of the shaded region represents an upper 95 percent bound for $I_{36}(t) - I_p(t)$, constructed from the trial results. If monotonicity is assumed, this edge also represents an (at least) 95 percent upper bound for $I_{12}(t) - I_p(t)$. The assumption of monotonicity also implies that $I_{12}(t) - I_p(t) \geq 0$, so that the shaded region represents an (at least) 95 percent confidence band for $I_{12}(t) - I_p(t)$. If monotonicity is not assumed, nothing can be inferred about $I_{12}(t) - I_p(t)$ beyond month 12; however, since the 12- and 36-month courses are identical for the first 12 months, the first 12 months of the confidence band in Figure 1 also represents, all other things being equal, a confidence band for $I_{12}(t) - I_p(t)$ over this period.
Meta-Analysis of randomized trials
to evaluate low incidence of events

Psychopharmacologic Drugs Advisory Committee -
December 13, 2006
Briefing documents, clinical and statistical reports, background

Association between suicide attempts and selective serotonin reuptake inhibitors: systematic review of randomised controlled trials

Dean Fergusson, Steve Doucette, Kathleen Cranley Glass, Stan Shapiro, David Healy, Paul Hebert and Brian Hutton

BMJ 2005;330;396-
doi:10.1136/bmj.330.7488.396

Objective To establish whether an association exists between use of selective serotonin reuptake inhibitors (SSRIs) and suicide attempts.

Design Systematic review of randomised controlled trials.

Data sources Medline and the Cochrane Collaboration's register of controlled trials (November 2004) for trials produced by the Cochrane depression, anxiety, and neurosis group.

Selection of studies Studies had to be randomised controlled trials comparing an SSRI with either placebo or an active non-SSRI control. We included clinical trials that evaluated SSRIs for any clinical condition. We excluded abstracts, crossover trials, and all trials whose follow up was less than one week.

Results Seven hundred and two trials met our inclusion criteria. A significant increase in the odds of suicide attempts (odds ratio 2.28, 95% confidence 1.14 to 4.55, number needed to treat to harm 684) was observed for patients receiving SSRIs compared with placebo. An increase in the odds ratio of suicide attempts was also observed in comparing SSRIs with therapeutic interventions other than tricyclic antidepressants (1.94, 1.06 to 3.57, 239). In the pooled analysis of SSRIs versus tricyclic antidepressants, we did not detect a difference in the odds ratio of suicide attempts (0.88, 0.54 to 1.42).
Many other examples

- AVANDIA - meta analysis of short term trials and analysis of longer term trials
- Lotronex
- Analyses of observational data bases by many parties
Review Infrastructure

- **Process**
  - Prospective planning for collection and analysis of safety data - Statistical Analysis Plans (SAP’s)
  - Efficient and timely access to trial data, patient level data, during the review process
  - Analysis tools for comparative analysis, risk factor identification, time dependencies, patterns
  - Implement standards for clinical trial data (CDISC, HL7), electronic submission of clinical trial data, efficient storage and retrieval processes, and analysis programs for exploration and rigorous safety analysis
  - Formal quantitative benefit risk builds on this infrastructure
The Regulatory Review Infrastructure

What is currently the practice for quantitative benefit and risk evaluation - who does it, and what is the infrastructure to do so?
Guideline for Industry

Structure and Content of Clinical Study Reports

Efficacy evaluation of clinical outcomes in individual studies

Integrated Analysis of Efficacy of all evidence - all studies - not a formal meta-analysis - consideration of studies that are not supportive

July 1996
ICH E3
Reviewer Guidance
Conducting a Clinical Safety Review of a New Product Application and Preparing a Report on the Review

Additional copies are available from:
Office of Training and Communication
Division of Drug Information, HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane
Rockville, MD 20857
(Tel) 301-827-4573
http://www.fda.gov/cder/guidance/index.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

February 2005
Good Review Practices

III. SPECIFIC GUIDANCE ON THE CONTENT OF THE SAFETY REVIEW
7.1 Methods and Findings ................................................. 11
  7.1.1 Deaths ..................................................................... 11
  7.1.2 Other Serious Adverse Events ................................. 15
  7.1.3 Dropouts and Other Significant Adverse Events ........ 16
    7.1.3.1 Overall Profile of Dropouts ............................... 16
    7.1.3.2 Adverse Events Associated with Dropouts ........... 18
    7.1.3.3 Other Significant Adverse Events ....................... 19
  7.1.4 Other Search Strategies ........................................... 19
  7.1.5 Common Adverse Events ......................................... 19
    7.1.5.1 Applicant’s Approach to Eliciting Adverse Events in the Development Program 20
    7.1.5.2 Establishing Appropriate Adverse Event Categories and Preferred Terms 20
    7.1.5.3 Incidence of Common Adverse Events — Assessment of Various Databases 22
    7.1.5.4 Common Adverse Event Tables ............................ 23
    7.1.5.5 Identifying Common and Drug-Related Adverse Events ........ 23
    7.1.5.6 Additional Analyses and Explorations .................. 24
    7.1.6 Less Common Adverse Events ............................... 25
    7.1.7 Laboratory Findings .............................................. 25
    7.1.7.1 Overview of Laboratory Testing in the Development Program 26
    7.1.7.2 Selection of Studies/Analyses for Drug-Control Comparisons of Laboratory Values 26
  7.4 General Methodology ................................................ 44
    7.4.1 Pooling Data Across Studies to Estimate and Compare Incidence 44
    7.4.1.1 Pooled Data vs. Individual Study Data .................. 44
    7.4.1.2 Combining Data ............................................. 45
    7.4.2 Explorations for Predictive Factors ........................ 45
    7.4.2.1 Explorations for Dose Dependency for Adverse Findings 45
    7.4.2.2 Explorations of Time-Dependency for Adverse Findings 46
    7.4.2.3 Explorations for Drug-Demographic Interactions ........ 47
    7.4.2.4 Explorations for Drug-Disease Interactions ............. 48
    7.4.2.5 Explorations for Drug-Drug Interactions ................ 48
    7.4.3 Causality Determination ...................................... 49
Need to address data analysis and handling

Dealing with a massive amount of safety measurements, events, data on resolution of outcomes and summaries

- First steps in understanding what has been collected:
  - Visual displays and summaries
  - Conceptualizing the time dependencies
    - Cumulative exposure, interaction with other medications, covariates
Exploration vs. Confirmation of Adverse Event Profiles

Exploratory strategies

- Strategies for exploring associations, multiplicities, time dependencies, syndromes, coding disparities (MeDDRA), event combinations, composite outcomes.

- Strategies for displaying, visualizing and interpreting time dependent multiple outcomes and measurements.
Potential Approaches

All rely on electronic data accessible by tools and analysis programs and data formatting (standards)

- Visual graphics and informative displays
  - Individual subject case report profiles
- Summarizing patient outcomes by treatment group
- Comparisons of treatment groups with respect to patterns and event rates
  - Event history charts
- New measures of cumulative events - counting events and adjusting for duration of exposure
Review Tools

- WebSDM
- Patient Profile Viewer
- iReview
- eDish
- Qtech
- PKS
- Others under development
Exposure history of test treatment - time course

Time of occurrence and duration of ADE’s

Exposure history of concomitant medications - time course

Patient time line graph for clinical trial data, AE’s, med. exposures, labs, outcomes

Drug Experience

Adverse Events

Concomitant Medications

Primary outcome

Serious event

Lab data

Drug Experience (Time Interval)

Death (Time Point)

Adverse Events (Time Interval, by WITHDRAWN)

Missing Value

N

Y

Concomitant Medications (Time Interval)

Lipids (Time Point)

Ax19 Normal

Lab Tests (Time Point)

Ax19 High

Ax19 Normal

Ax19 Low

Ax19 Missing Value

Ax19 Missing Range Values

Urinalysis (Time Point)

Ax19 Normal

Ax19 Missing Value

SO1099 (Time Point)

Ax19 Normal

Ax19 Missing Value

B

Missing End

E

Missing Begin
## Data Exploration

<table>
<thead>
<tr>
<th>Preferred Term</th>
<th>COMPARATOR</th>
<th>Study Drug</th>
<th>Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nausea</td>
<td>23 (19.83%)</td>
<td>12 (10.00%)</td>
<td>35 (14.83%)</td>
</tr>
<tr>
<td>Diarrhoea NOS</td>
<td>21 (18.10%)</td>
<td>14 (11.67%)</td>
<td>35 (14.83%)</td>
</tr>
<tr>
<td>Anaemia NOS</td>
<td>18 (15.52%)</td>
<td>15 (12.50%)</td>
<td>33 (13.96%)</td>
</tr>
<tr>
<td>Vomiting NOS</td>
<td>15 (12.93%)</td>
<td>14 (11.67%)</td>
<td>29 (12.29%)</td>
</tr>
<tr>
<td>Constipation</td>
<td>14 (12.07%)</td>
<td>13 (10.83%)</td>
<td>27 (11.44%)</td>
</tr>
<tr>
<td>Hypokalaemia</td>
<td>15 (12.93%)</td>
<td>11 (9.17%)</td>
<td>26 (11.02%)</td>
</tr>
<tr>
<td>Oedema peripheral</td>
<td>16 (13.79%)</td>
<td>8 (6.67%)</td>
<td>24 (10.17%)</td>
</tr>
<tr>
<td>Pain in extremity</td>
<td>11 (9.48%)</td>
<td>11 (9.17%)</td>
<td>22 (9.32%)</td>
</tr>
<tr>
<td>Headache</td>
<td>12 (10.34%)</td>
<td>8 (6.67%)</td>
<td>20 (8.47%)</td>
</tr>
<tr>
<td>Urinary tract infection NOS</td>
<td>11 (9.48%)</td>
<td>8 (6.67%)</td>
<td>19 (8.05%)</td>
</tr>
<tr>
<td></td>
<td>8 (6.90%)</td>
<td>11 (9.17%)</td>
<td>19 (8.05%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CPK Range</th>
<th>COMPARATOR</th>
<th>Study Drug</th>
<th>Row Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 50</td>
<td>104</td>
<td>96</td>
<td>200.00</td>
</tr>
<tr>
<td>051 to 100</td>
<td>80</td>
<td>94</td>
<td>174.00</td>
</tr>
<tr>
<td>101 to 150</td>
<td>41</td>
<td>56</td>
<td>97.00</td>
</tr>
<tr>
<td>151 to 200</td>
<td>31</td>
<td>38</td>
<td>69.00</td>
</tr>
<tr>
<td>201 to 250</td>
<td>16</td>
<td>19</td>
<td>35.00</td>
</tr>
<tr>
<td>251 to 300</td>
<td>11</td>
<td>17</td>
<td>28.00</td>
</tr>
<tr>
<td>301 to 350</td>
<td>10</td>
<td>15</td>
<td>25.00</td>
</tr>
<tr>
<td>351 to 400</td>
<td>7</td>
<td>9</td>
<td>16.00</td>
</tr>
<tr>
<td>401 to 450</td>
<td>5</td>
<td>9</td>
<td>14.00</td>
</tr>
<tr>
<td>451 to 500</td>
<td>5</td>
<td>6</td>
<td>11.00</td>
</tr>
<tr>
<td>501 to 550</td>
<td>2</td>
<td>5</td>
<td>7.00</td>
</tr>
<tr>
<td>551 to 600</td>
<td>3</td>
<td>3</td>
<td>3.00</td>
</tr>
<tr>
<td>Above 600</td>
<td>5</td>
<td>17</td>
<td>22.00</td>
</tr>
<tr>
<td>missing</td>
<td>42</td>
<td>43</td>
<td>85.00</td>
</tr>
</tbody>
</table>

*Col Sum = 359.00, 427.00, 786.00*
SMQs under development...

- Anaphylactic reaction*
- Acute renal failure*
- Rhabdomyolysis/myopathy*
- Cardiac failure
- Torsades/QT prolongation*
- Haematopoietic cytopenias
- Haemorrhage
- Stevens Johnson syndrome
- Suicide/depression

[^Targeted for version 6.1 release of MedDRA]

Acute Renal Failure Narrow Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute renal failure</td>
<td>An abrupt fall in glomerular filtration rate.</td>
</tr>
<tr>
<td>Anuria</td>
<td>Complete absence of urine output.</td>
</tr>
<tr>
<td>Progressive renal failure</td>
<td>A gradual decline in renal function.</td>
</tr>
<tr>
<td>Renal failure acute</td>
<td>Rapid deterioration in renal function.</td>
</tr>
<tr>
<td>Dialysis</td>
<td>A method of replacing kidney function.</td>
</tr>
<tr>
<td>Renal failure acute on chronic</td>
<td>As above in chronic renal failure.</td>
</tr>
<tr>
<td>Renal failure neonatal</td>
<td>In infants.</td>
</tr>
<tr>
<td>Renal impairment NOS</td>
<td>Not otherwise specified.</td>
</tr>
<tr>
<td>Renal tubular disorder NOS</td>
<td>As above in NOS.</td>
</tr>
<tr>
<td>Renal transplant</td>
<td>Transfer of a healthy kidney.</td>
</tr>
<tr>
<td>Renal tubular necrosis</td>
<td>Severe damage to renal tubules.</td>
</tr>
<tr>
<td>Tubulointerstitial nephritis</td>
<td>Injury to the tubules and intertubular tissue.</td>
</tr>
</tbody>
</table>

SMQ Acute Renal Failure

**Definition:** Acute renal failure is a syndrome characterized by a relatively rapid decline in renal function that leads to the accumulation of water, crystalloid solutes, and nitrogenous metabolites in the body. Clinically significant acute renal failure is usually associated with a daily increase in serum creatinine and urea nitrogen levels (so-called ‘azotemia’) greater than 0.5 and 10 mg per deciliter, respectively. Oliguria, a rate of urine flow less than 400 ml per day, may be observed, but in some cases the urine output exceeds this limit (nonoliguric acute renal failure). Complete cessation of urine flow, anuria is relatively uncommon. Primary renal lesions in the majority of drug-induced acute renal failure are intrarenal, at the level of the vasculature or tubule. Secondary causes include infections, dehydration, and shock.

WebSDM – Sector Map Analysis

Each sector on the Sector map provides a visual representation of the clinical adverse event data within each system organ class (SOC)

Sector Map Analysis of Adverse Events

- A Sector Map analysis is a graphical method for detecting differences in adverse events rates between treatment groups
- The “map” is display of adverse events broken down into different sectors (rectangles)

Tiles within each sector can represent PTs, HLTs, or HLGTs.
Display of the duration of treatment and follow up and the timing of increases in ALT for the 500 patients in a study receiving a "low" and a "high" dose treatment.
Process changes

Allocate time in IND/NDA review process to discuss and understand prospective plans for collection and analysis of safety outcomes
Develop safety data analysis plan

- Rigorous ascertainment of key safety outcomes is essential
- Document real reasons for patient withdrawal from exposure and missing data
- Ensure follow-up of all subjects for the intended period of the study especially for AESI or SAE (mortality) until resolution even if subjects discontinue study drug.
- Describe the statistical analyses to be conducted including sensitivity analyses

C. George Rochester, Ph.D., RochesterG@cdr.fda.gov
Communicating the benefits and risks in a manner so that personal decisions can be made

What can be done differently

Example: PREMPRO and the Women’s Health Initiative
WARNING

Estrogens and progestins should not be used for the prevention of cardiovascular disease or dementia. (See WARNINGS, Cardiovascular disorders and Dementia.)

The Women’s Health Initiative (WHI) study reported increased risks of myocardial infarction, stroke, invasive breast cancer, pulmonary emboli, and deep vein thrombosis in postmenopausal women (50 to 79 years of age) during 5 years of treatment with conjugated estrogens (0.625 mg) combined with medroxyprogesterone acetate (2.5 mg) relative to placebo. (See CLINICAL PHARMACOLOGY, Clinical Studies and WARNINGS, Cardiovascular disorders and Malignant neoplasms, Breast cancer.)

The WHI study reported increased risks of stroke and deep vein thrombosis in postmenopausal women (50 to 79 years of age) during 6.8 years of treatment with conjugated estrogens (0.625 mg) relative to placebo. (See CLINICAL PHARMACOLOGY, Clinical Studies and WARNINGS, Cardiovascular disorders.)

The Women’s Health Initiative Memory Study (WHIMS), a substudy of WHI, reported increased risk of developing probable dementia in postmenopausal women 65 years of age or older during 4 years of treatment with conjugated estrogens combined with medroxyprogesterone acetate and during 5.2 years of treatment with conjugated estrogens alone, relative to placebo. It is unknown whether this finding applies to younger postmenopausal women. (See CLINICAL PHARMACOLOGY, Clinical Studies, WARNINGS, Dementia and PRECAUTIONS, Geriatric Use.)

Other doses of conjugated estrogens and medroxyprogesterone acetate, and other combinations and dosage forms of estrogens and progestins were not studied in the WHI clinical trials and, in the absence of comparable data, these risks should be assumed to be similar. Because of these risks, estrogens with or without progestins should be prescribed at the lowest effective doses and for the shortest duration consistent with treatment goals and risks for the individual woman.
For those outcomes included in the WHI "global index", the absolute excess risks per 10,000 women-years in the group treated with PREMPRO were 7 more CHD events, 8 more strokes, 8 more PEs, and 8 more invasive breast cancers, while the absolute risk reductions per 10,000 women-years were 6 fewer colorectal cancers and 5 fewer hip fractures. The absolute excess risk of events included in the "global index" was 19 per 10,000 women-years. There was no difference between the groups in terms of all-cause mortality. (See BOXED WARNING, WARNINGS, and PRECAUTIONS.)

<table>
<thead>
<tr>
<th>Event</th>
<th>Relative Risk PREMPRO vs Placebo at 5.2 Years (Nominal 95% CI)</th>
<th>Placebo n = 8102</th>
<th>PREMPRO n = 8506</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD events</td>
<td>1.29 (1.02-1.63)</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>Non-fatal MI</td>
<td>1.32 (1.02-1.72)</td>
<td>23</td>
<td>30</td>
</tr>
<tr>
<td>CHD death</td>
<td>1.18 (0.70-1.97)</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Invasive breast cancer b</td>
<td>1.26 (1.00-1.59)</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>Stroke</td>
<td>1.41 (1.07-1.85)</td>
<td>21</td>
<td>29</td>
</tr>
<tr>
<td>Pulmonary embolism</td>
<td>2.13 (1.39-3.25)</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>0.63 (0.43-0.92)</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>Endometrial cancer</td>
<td>0.83 (0.47-1.47)</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Hip fracture</td>
<td>0.66 (0.45-0.98)</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Death due to causes other than the events above</td>
<td>0.92 (0.74-1.14)</td>
<td>40</td>
<td>37</td>
</tr>
<tr>
<td>Global Index c</td>
<td>1.15 (1.03-1.28)</td>
<td>151</td>
<td>170</td>
</tr>
<tr>
<td>Deep vein thrombosis d</td>
<td>2.07 (1.49-2.87)</td>
<td>13</td>
<td>26</td>
</tr>
<tr>
<td>Vertebral fractures d</td>
<td>0.68 (0.44-0.98)</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Other osteoporotic fractures d</td>
<td>0.77 (0.69-0.86)</td>
<td>170</td>
<td>131</td>
</tr>
</tbody>
</table>

a: adapted from JAMA, 2002; 288:321-333
b: includes metastatic and non-metastatic breast cancer with the exception of in situ breast cancer

c: a subset of the events was combined in a "global index," defined as the earliest occurrence of CHD events, invasive breast cancer, stroke, pulmonary embolism, endometrial cancer, colorectal cancer, hip fracture, or death due to other causes
d: not included in Global Index
Concluding Remarks
The Science of Quantitative Benefit / Risk Assessment

- Appropriate quantification of benefit AND risk is just beginning to be understood and addressed - borrowing from other fields - need to understand what are limits and practical - must build on quantification of efficacy and harms

- Quantification of efficacy is more refined - 30 years of development - benefit may be a more difficult metric

- Quantification of safety (risk, harm) is way behind and needs a lot of attention to get where we are for efficacy - leads to asymmetry in benefit/risk quantification
Concluding Remarks

The infrastructure for quantitative benefit/risk assessment

- Will need to rely on standards for clinical trials, data formats, access, storage and retrieval
- Tools
- Data bases - role of epidemiologic observational studies - cross study comparative issues
- Framework, metrics, understanding, training
- A culture change - reflects medical literature and other practices