

Evaluating Harms of Medical Interventions in Infectious Diseases

John H. Powers, MD FACP FIDSA

Assistant Clinical Professor of Medicine

George Washington University School of Medicine

University of Maryland School of Medicine

Introduction

- **Defining adverse events**
- **Differences between evaluating harms and evaluating benefits**
- **Basic principles of evaluating harms**
 - **Frequency**
 - **Nature of harm**
 - **Evaluating causality**
 - **Comparison to benefits essential**
 - **Ongoing evaluation of risk-benefit, not one time thing**
- **Risk communication and risk perception in antimicrobial usage**

Defining Terms

- **Side effect = effect of intervention that is not the principal effect for which intervention was chosen; may be desirable or undesirable**
 - Nebeker JR et al. Ann Intern Med 2004;140:795-801.
- **Adverse Event (or Experience) = undesirable effect with use of intervention whether or not drug related**
- **Adverse Reaction or Adverse Drug Reaction = undesirable effect reasonably associated with use of the intervention. Includes signs, symptoms, lab values, vital signs, ECG, etc.**

Evaluating Effectiveness vs. Harms

- Evaluating Efficacy

- Testing a hypothesis in randomized trials
- Randomized trial designed to ascribe causality to interventions prescribed
- Statistical testing based upon single comparison of primary endpoint
- Events common (all patient either a success or failure)
- Evaluate individual studies
- Detailed CRF in trials

- Evaluating Harms

- Hypothesis often not known at time of initiation of trials
- Causality ascribed in post-hoc manner and potential confounders
- Multiple comparisons of interest
- Events may be less common or rare
- Pool data across studies
- Spontaneous reports (AERS) often not as detailed

Evaluating Risks

- Starting point in law since 1938 is drug is not safe, need evidence of potential harms compared to benefits
- Medical interventions not “innocent until proven guilty” given *a priori* knowledge that all drugs have some harm
- “Safety” implies substantive absence of all harm, yet all drugs associated with some potential harms
 - CONSORT Ann Intern Med 2004;141:781-88.

Regulatory Standard for “Safety”

- *“adequate tests by all methods reasonably applicable to show whether or not such drug is safe for use under the conditions prescribed, recommended, or suggested in the proposed labeling”*
 - Food Drug and Cosmetic Act, sec 505 (d)
- What tests are “adequate”?
- Note link to conditions of use – safe in one clinical setting might not be safe in another setting

Overall Considerations

- **Not testing a hypothesis when evaluating safety, rather looking for a hypothesis to test in future studies**
- **Search widely and then further evaluate**
- **Assessments based on many pieces of information making a coherent whole to avoid false-positive signals**
 - **Pre-clinical in vitro and animal studies**
 - **Healthy volunteers**
 - **Clinical trials**
 - **Post-approval evaluations from case series, case control, cohort and randomized trials**
- **P-values (hypothesis testing) and statistical significance less important when searching for signals related to safety**

Statistical vs Clinical Significance

- *“No formal tests of significance can answer those questions. Such tests can, and should, remind us of the effects that the play of chance can create, and they will instruct us in the likely magnitude of those effects. Beyond that they contribute nothing to the ‘proof’ of our hypothesis.....Yet too often I suspect we waste a deal of time, we grasp the shadow and lose the substance, we weaken our capacity to interpret the data and to **take reasonable decisions** whatever the value of P. And far too often we deduce ‘no difference’ from ‘no significant difference.’ Like fire, the chi-squared test is an excellent servant and a bad master.”*
- **Austin Bradford Hill, The environment and disease: association or causation? Proc Roy Soc Med 1965; 58:295-300.**

Overall Considerations

- **“Absence of evidence is not evidence of absence”**
 - Hartung J et al. *Anesthesiology* 1983;58:298-300.
- **Reporting of harms in medical literature generally suboptimal**
 - Ioannidis J et al. *JAMA*. 2001 Jan 24-31;285(4):437-43
- **Data on safety often obtained outside of randomized clinical trials**
 - Randomized trials less susceptible to random error, systematic error (bias and confounding)
 - Need to use other forms of non-randomized data to acquire information on potential harms

Overall Considerations

1. Frequency of harms

2. Nature of harm

- Serious adverse events vs non-serious
- Conditions of use – dose and duration of exposure
- Timing of onset
- Ability to monitor for adverse event to mitigate potential harm

3. Evaluating causality of harm

4. Qualitative as well as quantitative comparison of benefits to harms

- Other interventions without that adverse event available to treat or prevent disease

1. Frequency of Harm

- Most data used to evaluate risks come from passive, voluntary reporting systems (Adverse Event Reporting System, AERS)
- Can report through MedWatch
<https://www.accessdata.fda.gov/scripts/medwatch/medwatch-online.htm>
- Estimated only 1-10% of adverse events are reported therefore analyses of frequency of AEs often serious underestimates
- FDAAA of September 2007 requires FDA to develop better ways of obtaining data

1. Frequency of Harm

- Sample size of trials for approval based on numbers needed to demonstrate effectiveness
- Trial sample sizes too small, too short duration, too narrow a population to observe evidence of harms
- “Rule of threes”:
 - absence of observed harm allows ruling out with 95% confidence rate of 3 divided by number of subjects studied
 - e.g. Not observing an event in 3000 patients allows one to rule out rate as high as 3/3000 or 0.1%
- When exposing millions of patients, a “low” rate may mean substantial absolute number of patients and morbidity
 - Rumble CL. N Engl J Med 1975;292:372-3.
 - Hanley JA et al. JAMA 1983;249:1743-5.

2. Nature of Harm

Serious Adverse Reaction

- **Adverse experience at any dose that results in any of following outcomes:**
 - death
 - life-threatening adverse experience,
 - inpatient hospitalization or prolongation of hospitalization
 - persistent or significant disability/incapacity
 - congenital anomaly or birth defect
- **Other events which jeopardize patient or subject and may require medical/surgical intervention to prevent outcomes listed in definition e.g. bronchospasm treated in ER or home**

2. Nature of Harm

- **Seriousness - Non-serious/common events and serious/uncommon events are of most interest**
 - Interventions with serious/common events unlikely to be developed
 - Unlikely to see serious/uncommon events in studies
- **Severity is grading of events within levels of seriousness**
 - e.g. a rash may be considered a non-serious AE but severe when it covers a large percentage of the body surface area (ICH-E2A)
- **Data from less serious adverse events may be predictors for rarer, more serious adverse events**
 - Asymptomatic increases in liver tests may be signal for more serious adverse liver toxicity when intervention more broadly used
 - Certain kinds of rash may be premonitory signal for Stevens-Johnson Syndrome

2. Nature of Harm

- **Timing of onset – does AE occur after single dose or after multiple doses**
- **Conditions of use – does AE occur at normally prescribed dose or at supra-therapeutic doses?**
- **Monitoring – can one monitor for AE in order to take action to mitigate risk?**

3. Evaluating Causality

- **Hill's nine considerations for evaluating causality (causation more likely if following present):**
 1. **Strength of association – larger effects more likely to represent true findings**
 2. **Consistency – observed by different persons in different places**
 3. **Specificity – no other likely explanation**
 4. **Temporality – effect occurs after cause with expected delay**
 5. **Biological gradient – greater exposure leads to greater effect**
 6. **Plausibility – knowledge of mechanism (limited by current knowledge)**
 7. **Coherence – agreement with epidemiological and lab findings (but such data does not nullify epidemiological associations)**
 8. **Experiment – testing hypothesis in experiment where investigator controls exposure/medical intervention**
 9. **Analogy – reference to similar situations**

Table 2. Association between Hypoglycemia-Related Hospital Visits and Recent Antibiotic Use.*

Variable	Case Patients no. (%)	Controls	Univariate Odds Ratio (95% CI)	Adjusted Odds Ratio (95% CI)
All patients	788	3791		
Fluoroquinolones				
Gatifloxacin	61 (7.7)	77 (2.0)	4.4 (3.0–6.3)	4.3 (2.9–6.3)
Levofloxacin	114 (14.5)	341 (9.0)	1.7 (1.4–2.2)	1.5 (1.2–2.0)
Moxifloxacin	24 (3.0)	162 (4.3)	0.8 (0.5–1.3)	0.8 (0.5–1.3)
Ciprofloxacin	209 (26.5)	1075 (28.4)	1.1 (0.9–1.3)	0.9 (0.8–1.1)
Cephalosporins†	62 (7.9)	397 (10.5)	0.9 (0.6–1.2)	0.9 (0.6–1.2)
Macrolides‡	318 (40.4)	1739 (45.9)	1.0	1.0
Patients with diabetes	724	3473		
Fluoroquinolones				
Gatifloxacin	57 (7.9)	73 (2.1)	4.3 (2.9–6.2)	4.2 (2.8–6.3)
Levofloxacin	104 (14.4)	307 (8.8)	1.8 (1.3–2.3)	1.5 (1.2–2.0)
Moxifloxacin	22 (3.0)	149 (4.3)	0.8 (0.5–1.3)	0.8 (0.5–1.3)
Ciprofloxacin	195 (26.9)	1002 (28.9)	1.1 (0.9–1.3)	0.9 (0.7–1.1)
Cephalosporins†	55 (7.6)	362 (10.4)	0.8 (0.6–1.1)	0.8 (0.6–1.1)
Macrolides‡	291 (40.2)	1580 (45.5)	1.0	1.0
Patients without diabetes	64	318		
Fluoroquinolones				
Gatifloxacin	<6	<6	6.4 (1.4–29.9)	9.0 (1.3–63.4)
Levofloxacin	10 (15.6)	34 (10.7)	1.7 (0.7–3.9)	2.1 (0.7–6.0)
Moxifloxacin	<6	13 (4.1)	1.0 (0.2–4.5)	1.7 (0.2–11.8)
Ciprofloxacin	14 (21.9)	73 (23.0)	1.1 (0.6–2.3)	1.2 (0.5–2.9)
Cephalosporins†	7 (10.9)	35 (11.0)	1.2 (0.5–3.0)	2.3 (0.8–6.7)
Macrolides‡	27 (42.2)	159 (50.0)	1.0	1.0

Park-Wyllie et al NEJM 2006;354:1-10

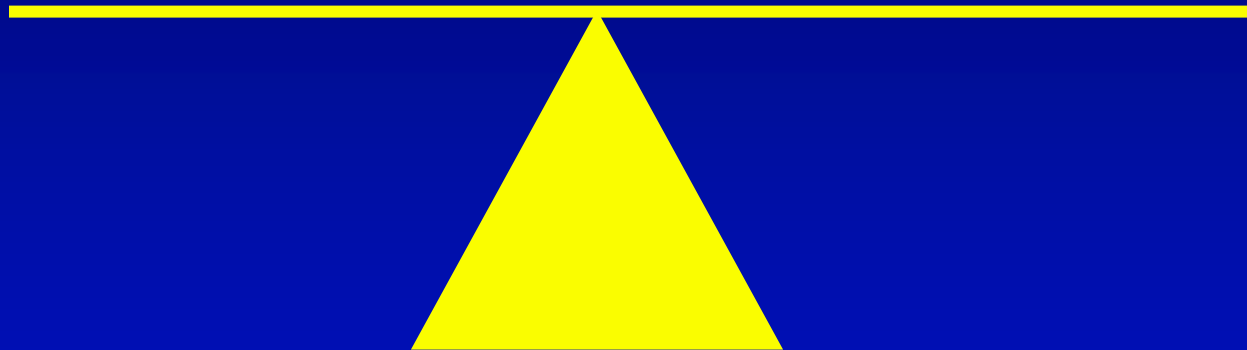
4. Balancing of of Risks and Benefits

Drug Benefit

Adverse Events

**Prevention
of Death**

**Nausea, Vomiting
Headache, Sniffles**



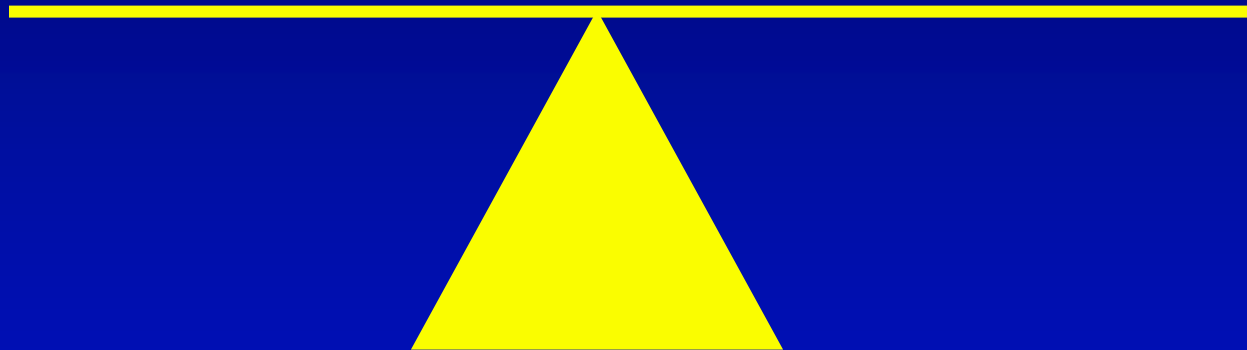
4. Balancing Risks and Benefits

Adverse Events

Drug Benefit

Death

**Decrease in Nausea, Vomiting
Headache, Sniffles**



Learning from History

- *“If the drug that killed one person in ten thousand was of only minor use therapeutically, it might still be judged to be unsafe, whereas the drug that killed one in a thousand persons, if it had marked and undisputed therapeutic value it would still be a safe and valuable drug”*
 - J.J. Durett, Chief, Drug Division, FDA, December 1938
- **Safety and effectiveness dependent upon conditions of use – not just if a drug “works” but *in whom, when, how used* and on *what outcomes***

Learning from History

- *“I think you will be interested in some of the implications that arise from the observations recorded. I refer particularly to the 105 deaths associated with the consumption of the drug...in a hundred instances the drug was administered on a physicians prescription” for causes such as “Bright’s disease, bichloride and mercury poisoning, renal colic and backache.”*
 - Theodore Klumpp, FDA director Drug Division, December 29, 1939 referring to deaths from elixir of sulfanilamide

Learning from History

- *“For the most part, sulfapyridine should be used only for patients who are seriously ill. I doubt the advisability of using the drug for patients who have influenza, the common cold, sinusitis or tonsillitis. In such cases, this treatment may be worse than the disease, not only much more uncomfortable, but more dangerous.”*
 - H. Corwin Hinshaw, Proceedings of Staff Meeting of the Mayo Clinic, 1939;14:771.
- **Studies on prevention failed to show benefit of antimicrobials**
 - Robertson O. Newer Knowledge Concerning the inception of pneumonia and its bearing on prevention. Ann Intern Med 1943;18:12.

Learning from History

- *“And now I am starting to use [sulfa] prophylactically. And why not? It has not been proven to work that way! Not scientific you say! Remember we are front line soldiers; when we see the enemy we do not have to wait for orders from headquarters through a long line of red tape. We must go for him, without waiting for the attack!...Then why not get the jump on those tough little bacteria? Kill them before they get a foothold.”*

**William McIlwaine. Virginia Medical Monthly
1941:68:410-1.**

“It Can’t Hurt”

- 23 yo female nursing student, recently engaged in normal health with no underlying illness
- Prescribed antibiotic in ER for “persistent cold” (thought it would be faster than going to family physician)
- Dx of “acute bronchitis in ER” and rx filled 42 minutes later
 - Day 3 – nausea abdominal pain (liver failure)
 - Day 7 – incoherent, Stevens-Johnson syndrome
 - Day 8 –liver transplant
 - Day 10 - death



www.fqresearch.org

Example: hepatotoxicity with telithromycin vs acetaminophen

- **Telithromycin**

- Unproven benefit in acute exacerbations of chronic bronchitis, sinusitis, failed in pharyngitis trials
- Acute onset of hepatic failure
- Hypersensitivity reaction after single administration at usual dose
- Frequency unclear given underreporting and short history of usage
- Occurs in healthy persons
- Can't monitor for event

- **Acetaminophen**

- Effective for pain control in a wide range of diseases both serious and self-resolving
- Onset of hepatic failure
- Cumulative toxicity after multiple doses and when administered at greater than usual dose
- More data on frequency given long history of usage
- Occurs more commonly with underlying liver disease
- Can monitor LFTs

Concept of “Risk”

- **Risk = potential negative impact on some asset/characteristic of value arising from some present process or future event**
- **Differentiate risk from uncertainty**
 - Risk implies a measurable value
 - Uncertainty implies something that is not measured
- **Confusion occurs when there is uncertainty about measurement of risk; often the case with new drugs**

Concept of “Risk”

- Risk consists of two factors:
 - Impact: death vs less serious morbidity
 - Probability: likelihood of event occurring
- Probability refers to outcomes in groups of subjects, not outcome in an individual
 - Probability of outcome in an individual is either 0% or 100%
 - “You have a 1 in 100 chance of an adverse event” is an incorrect statement
- If probability of event is 1 in 100, who is the one and who are the other 99? Correct answer : no one knows
- Try to narrow down who is most at risk of event to decrease probability of event (age, gender, other baseline characteristics)

Risk Communication vs Perception

- **Risk communication** – clear factual description of nature, frequency, severity, description of at risk characteristics of potential adverse events and potential ways to mitigate risk
- **Risk perception** – subjective judgments people make about the characteristics and severity of risk
- **Study of risk perception** arose out of observation that experts and lay people often disagree about risk of various technologies and natural hazards

Addressing Risk Perception

- Dealing with risk perception requires understanding and education
- *Not* communicating risk at all is not appropriate way to deal with risk perceptions
- Appropriate communication is not “scaring” people and patients have *prima facie* right not to take a medication (first do no harm)
- “Optimism bias” in evaluating effectiveness compared to safety:
 - Early, unconfirmed reports of potential effectiveness often accepted at face value and translated into treatment guidelines – combination therapy in community-acquired pneumonia
 - Early reports of potential safety signals dismissed or “need more data” before communicating potential risk
 - Over-valuation of unclear benefits without addressing risk

Continuing Process

- **Information is not static; need continual reassessment of both safety and effectiveness based on new information**
- **Assessment of risk-benefit balance may change with new information – same risk may become less acceptable if effectiveness changes (resistance) or new evidence on harms**
- **Need ongoing reassessments of what defines “resistance” and its clinical impact since often little information at time of approval**
- **Same risk may be more acceptable in serious and life-threatening diseases than in self-resolving diseases**

Conclusions

- **Evidence for benefits and harms is not based on opinion or anecdotes, but on reasoned application of scientific criteria**
- **Not testing hypothesis in evaluating risks so need to search broadly and then further evaluate signals**
- **Take into consideration both nature as well as frequency of harms realizing limitations on frequency data**
- **Need to balance the potential benefits and harms based on conditions of use**