Error and Bias in Clinical Research

Edward A. Panacek, MD,MPH

SAEM Annual Meeting, May 2009

New Orleans, LA

Lecture

Session learning objectives

- Appreciate the inevitable role of error in all research studies
- Define major types of error in research
  - Random error
  - Systematic error (bias)
- Identify sources and sub-types of systematic error
- Define and identify confounding of study results and conclusions
- Techniques to avoid, reduce, or compensate for error in research

What is research?

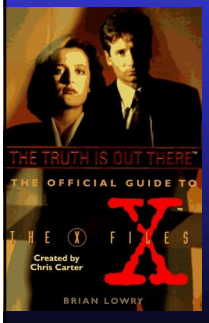
Let us think about this.....

What is research?

The NIH definition:

"A systematic investigation designed to develop or contribute to generalizable knowledge"

What is research?



In the purest, most generic sense, it is....

"A search for truth in the universe"

Truth in the universe

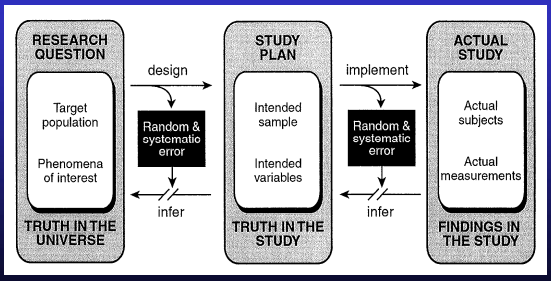
Can the investigator ever really know the absolute truth?

- Can we study the entire universe?
- Can our measurements be absolutely accurate?
- Can our data handling be error free?

For example. Suppose your research question was as simple as: What is the average, and range, for serum cholesterol in human beings on the planet earth?

- Could you ever test every human being?
  - No. You take a sample, so "sampling error" is possible
- Of all your samples, is it possible that none would be mishandled (lost, mislabeled, etc.)?
  - No, so "random errors" are likely, and possibly systematic errors
- When put through the lab, is that machine 100% valid and reliable?
  - No, so some "measurement error" is to be expected
- When the results are transcribed, are errors possible?
  - Yes, of course. So "random error" is likely

The relationship of research studies to the truth



External validity Internal validity

"Validity" of Research

- It is the degree to which a research parameter (test, survey, etc.) measures what it is intended to measure.
- It is the ability of research to find the truth.
- The threats to validity are "error".
- *"When bias creeps in...validity leaks out"*

When & where can error occur in research studies?

- Study design
- Study conduct
- Study measurements
- Data handling
- Data analysis
- Drawing conclusions

Potential impact of error on study results

- False negatives
- False positives
- Inaccurate effect sizes
  - Underestimates
  - Overestimates
- *Are there any other ways to be wrong??!*

Every research measurement has multiple potential ingredients

All study measurements have 3 potential components:

Truth

Random error

Systematic error

What are the main categories of research error?

- **Random error**
- **Systematic error**
  - **AKA: Bias**
- **Confounding**
  - Will cover this later

What is random error?

- **Deviations from the true value that occur in a random chance pattern, that cannot be precisely predicted**
- **Examples: sloppiness, simple mistakes, random chance effects**

What is systematic error?

"Any process at any stage of inference which tends to produce results or conclusions that differ systematically from the truth."

- To be distinguished from random error…

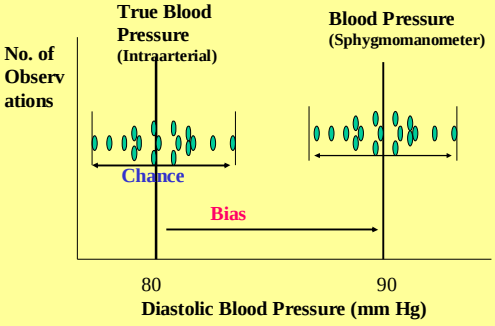
- Sackett DL. Bias in analytic research. J Chron Dis. 1979; 32:51-63.

Research as looking into a mirror analogy: curved mirrors and dirty mirrors

Types of error: The mirror image analogy

- **Consider observing or measuring something by looking at its' reflection in a mirror. This is much like doing research**
  - You can't see it directly
- **Bends or distortions in the mirror (like a carnival mirror) are the equivalent of *systematic error*. They are always there.**
- **Dirt on the mirror is like *random error*. Polishing the mirror (data cleanliness) can minimize the effects of random error.**

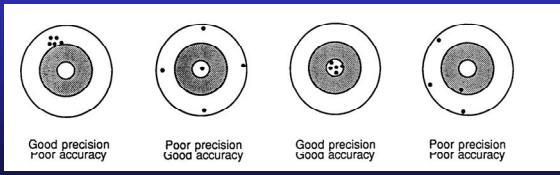
Example: Systematic error/bias vs chance



All studies make measurements.
Each of which have 2 performance parameters

- **Precision (reliability, consistency)**
  - **degree to which the same value is obtained when measuring the same thing repeatedly**
  - **affected by random error**
- **Accuracy (validity)**
  - **degree to which a measurement represents what it is intended to represent**
  - **affected by systematic error, e.g. observer bias, subject bias or instrument bias**

Using a target analogy



Types of biases

- **There are more than 100 different "biases" that have been described**
- **However, they are commonly divided in to two main categories:**
- **Selection Bias**
  - Occurs during subject enrollment
- **Information Bias**
  - Occurs during study measurements

References: J Kishore. A Dictionary of Public Health 2007

Selection Bias: Definition

A faulty assumption that occurs because there are systematic differences in characteristics between those who are selected for study and those who are not.

Systematic Errors: Selection Bias

- **Two main types:**
- **Ascertainment Bias** is systematic error resulting from failure to identify equally all categories of individuals who are supposed to be represented in a group
  - e.g., study based on specialty hospital.
- **(Non-Random) Sampling bias**
  - e.g. non-representative sample.

Sampling biases

- **All eligible subjects = the universe**
- **All studies take some form of sample**
- **Most statistical testing is based upon an assumption of having a true "random sample"**
- **Goal is a true "probability sample" where each candidate has an equal chance of study enrollment**
- **At a minimum, the study population should be a "representative" sample**

Selection Bias: systematic differences between those who are selected for study and those not

- Prevalence-incidence or survival bias: Selection of currently available, existing cases will miss fatal and short episodes, and may miss mild or silent cases
- Non-response bias: Differential rates of non-response to inquiries between cases and controls
- Membership bias: Membership in a group (blood donors, Army recruits) may imply a degree of health differing systematically from the general population
- Referral or admission rate bias: Cases who are more likely to receive advanced treatment (those with greater access to health care or co-existing illness) may distort associations with other factors
- *Sackett D, J Chron Dis1979; 32:51-63 and Schlesselman J, Case-Control Studies, 1982.*

Selection bias example

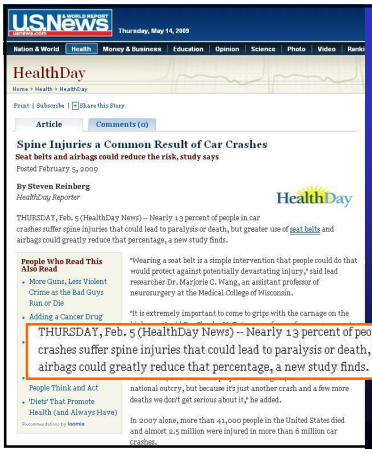
- Famous experiment by Yale psychologist Stanley Milgram in 1960's
- Asked people in Kansas to forward a letter to a target person in Massachusetts
- If did not know target person, then send it to someone they thought might know him
- His 1967 paper reported that it only took 5 jumps, on average, for letters to arrive
- Resulted in the "six degrees of separation" principle of connectedness

Selection bias example: A re-analysis of the study results

- According to Judith Kleinfeld,psychologist at the University of Alaska, Fairbanks, archives reveal that only 30 % of the letters actually reached their destination!
- Only those 30 % (selected sub-sample) were used in the calculation of the "5 steps" of separation

Gewolb, Josh. Random samples. *Science* 26 October 2001;294:777

Kleinfeld, Judith S. *Society*. Jan/Feb2002; 39(2):61-66)



Spectrum bias: A type of selection bias

THURSDAY, Feb. 5 (HealthDay News) -- Nearly 13 percent of people in car crashes suffer spine injuries that could lead to paralysis or death, but greater use of seat belts and airbags could greatly reduce that percentage, a new study finds.

Prevalence Bias

- A type of selection bias
- Occurs in the study design phase relating to subject enrollment

Example:

- Imagine your research question is: **"What is the prevalence of HBP in pts with cardiovascular disease?"**

Prevalence Bias- example

- Cohort Study

patients	Alive with CVD	Dead from Stroke	Alive with NO CVD
With HBP	50	250	700
Without HBP	80	20	900

- Case /Control Study

patients	Alive with CVD	Alive with NO CVD
With HBP	50	700
Without HBP	80	900

Selection bias example

Systematic Bias Introduced by the Informed Consent Process in a Diagnostic Research Study

Alice M. Mitchell, MD, MS, Jeffrey A. Kline, MD

Abstract

Objectives: To determine population characteristics, outcomes, and reasons for unsuccessful enrollment among potential study subjects approached for written, informed consent in a minimal-risk emergency department (ED) study. The authors hypothesized that the prevalence of venous thromboembolism (VTE) would be lower among study participants and that medical acuity and refusal to provide a blood sample would be the most common reasons for nonparticipation.

Methods: The authors requested prospective, written, informed consent for a blood sample and follow-up from consecutive ED patients undergoing evaluation for pulmonary embolism (PE) and recorded spontaneously stated reasons for refusal. VTE was diagnosed or excluded using a combination of D-dimer testing and selective computed tomography (CT) angiography of the chest with venography of the lower extremities. The primary outcome was defined by the number of CT scans positive for VTE among ED patients evaluated for PE.

Results: Over 16 weeks, 260 of 287 (91%, 95% confidence interval [CI] = 87 to 94%) eligible patients were approached and consent was obtained from 183 patients (64%, 95% CI = 58% to 69%). The prevalence of VTE was 6% among participants and 13% among nonparticipants (95% CI [of the differ-

Results: Over 16 weeks, 260 of 287 (91%, 95% confidence interval [CI] = 87 to 94%) eligible patients were approached and consent was obtained from 183 patients (64%, 95% CI = 58% to 69%). The preva-

ACADEMIC EMERGENCY MEDICINE 2008; 15:225–230 © 2008 by the Society for Academic Emergency Medicine

Information Bias

Occurs during data collection.
There are Five main categories.

1. Misclassification Bias
2. Observer/interviewer Bias
3. Recall Bias
4. Reporting Bias
5. Other information biases:
   - Hawthorne effect, loss to follow up,

Information Bias: systematic differences in data collection/reporting between cases and controls

- **Recall bias: Questions about specific exposures may be asked more frequently of cases, or cases may search their memories more intensively**
- **Family information bias: The flow of family information about exposures or illnesses may be stimulated by, or directed to, a new case in its midst**
- **Exposure suspicion bias: Knowledge of a patient's disease status may influence the intensity and outcome of search for exposure to a putative cause**
- **Instrument bias: Defects in calibration or maintenance of measurement instruments may lead to systematic deviations from true values**
- ***Sackett D, J Chron Dis1979; 32:51-63 and Schlesselman J, Case-Control Studies, 1982.***

MISCLASSIFICATION

Misclassification Bias: the erroneous classification of an individual, a value, or an attribute into a category other than that to which it should be assigned

- Often results from an improper "cutoff point" in disease diagnosis or exposure classification;
- Can result from poor diagnostic test performance

- 2 types of misclassification bias
  - differential (systematic)
  - non differential (random)

The effect of misclassification bias on study results and conclusions

TABLE 8-1. *Effect of nondifferential misclassification of alcohol consumption on the estimation of the incidence rate difference and incidence rate ratio for laryngeal cancer (hypothetical data)*

	Incidence rate ($\times 10^5$ yr)	Rate difference ($\times 10^5$ yr)	Rate ratio
No misclassification			
1,000,000 drinkers	50	40	5.0
500,000 nondrinkers	10		
Half of drinkers classed with nondrinkers			
500,000 drinkers	50	20	1.7
1,000,000 "nondrinkers" (50% are actually drinkers)	30		
Half of drinkers classed with nondrinkers and one-third of nondrinkers classed with drinkers			
666,667 "drinkers" (25% are actually nondrinkers)	40	6	1.2
833,333 "nondrinkers" (60% are actually drinkers)	34		

Effect of misclassification bias

Non-differential misclassification

- **Bias is towards the null**
  - **= more conservative**
  - **i.e. weaker strength of association between the IV and DV**

Differential misclassification

- **Bias direction is unpredictable and potentially much more serious**

Recall Bias
Non Campus Mentis

"History, as we know, is always bias, because human beings have to be studied by other human beings, not by independent observers of another species."

Anders Henriiksson (ed), *Non Campus Mentis*, NY, Workman Publishing Co., 2003, chapter 1

Requirements for a "bias free" Case-Control study

- Cases are representative of all those in the study base who develop the disease
- Controls are representative of all those in the study base at risk of developing the disease and eligible to become cases and be detected in the study
- Collection of risk factor and exposure information is the same for cases and controls
- Ancestral geographical origins and predominant environmental exposures of cases do not differ dramatically from controls

Key Requirements for a Bias-Free Case-Control Study

- Cases are representative of all those in the study base who develop the disease
- Controls are representative of all those in the study base at risk of developing the disease and eligible to become cases and be detected in the study

Selection Bias

- Collection of risk factor and exposure information is the same for cases and controls
- *Ancestral geographical origins and predominant environmental exposures of cases do not differ dramatically from controls*

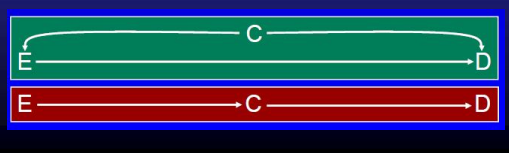
Information Bias

What is "Confounding"?

- **From the Latin *confundere*, to mix together**
- **"The distortion of the apparent effect of an exposure on risk, brought about by the association with other factor[s] that can influence the outcome"**
- **A Dictionary of Epidemiology by John Last, 1995.**

Confounder

- **An extraneous variable that distorts the observed relationship between the study independent and dependent variables**
- **Must have independent associations with the IV and the DV**
  - **i.e. is unevenly distributed between the study groups**
- **Is not an intermediate step on the "causal pathway"**



Confounding variables

- **Confounding variables are always extraneous, but not all extraneous variables are confounding**
  - **An EV associated with both the IV and DV**
- ***Can partially or completely alter the study conclusions if not taken into account !***
  - **e.g. can result in a conclusion that A causes B, when truth is that A has no effect on B**

Imagine a research question: "Which hospital has better CABPG surgery outcomes?"

PROBLEM 1

Here are data on the outcome of surgeries performed in one of two hospitals during a particular month. The data are stratified by the overall health status of the patient prior to the surgery.

	Good health			Poor health	
	Good outcome	Poor outcome		Good outcome	Poor outcome
Hospital A	1440	360	Hospital A	30	270
Hospital B	540	60	Hospital B	300	600

PROBLEM 1

Here are data on the outcome of surgeries performed in one of two hospitals during a particular month. The data are stratified by the overall health status of the patient prior to the surgery.

	Good health			Poor health	
	Good outcome	Poor outcome		Good outcome	Poor outcome
Hospital A	1440	360	Hospital A	30	270
Hospital B	540	60	Hospital B	300	600

Good surgery outcomes

Hospital A Good health : 80% Poor health: 10%

Hospital B : 90% : 33%

Now, with all patients combined, calculate the good outcome rates by hospital

a. Make a 2x2 table of surgery outcome by hospital.

	Good outcome	Poor outcome	Total
Hospital A	1470	630	2100
Hospital B	840	660	1500

Hospital A: 70% **Hospital B: 56%**

Here are data on the outcome of surgeries performed in one of two hospitals during a particular month. The data are stratified by the overall health status of the patient prior to the surgery.

	Good health			Poor health	
	Good outcome	Poor outcome		Good outcome	Poor outcome
Hospital A	1440	360	Hospital A	30	270
Hospital B	540	60	Hospital B	300	600

Good outcomes, stratified by health

Hospital A Good health : 80% Poor health: 10%

Hospital B : 90% 33%

a. Make a 2x2 table of surgery outcome by hospital.

	Good outcome	Poor outcome	Total
Hospital A	1470	630	2100
Hospital B	840	660	1500

Hospital A: 70% Hospital B: 56%

How is this possible?

- **Confounding by an extraneous variable!**
- **Outcome is confounded by health status, which is maldistributed between the hospitals**

PROBLEM 1

Here are data on the outcome of surgeries performed in one of two hospitals during a particular month. The data are stratified by the overall health status of the patient prior to the surgery.

	Good health			Poor health	
	Good outcome	Poor outcome		Good outcome	Poor outcome
Hospital A	1440	360	Hospital A	30	270
Hospital B	540	60	Hospital B	300	600

The profound effects of confounding

- **In this case, a confounding variable (baseline health status) completely reversed the study conclusions regarding which hospital has the better CAPG surgery outcomes.**
- **This is because that variable was not evenly distributed between the 2 hospitals, and it had a stronger association with out come than did the hospital type**

Examples of extraneous variables that could be confounders

- Smokers in an asthma study
- Current steroid use in an asthma study
- Prior analgesic use in a pain study
- Prior BP meds in trauma VS study
- Stimulant drugs use in trauma VS study

It is said that the randomization process is designed to control for confounders

Is that true?

A study of the role of confounding

Observational Study of Intravenous versus Oral Corticosteroids for Acute Asthma: An Example of Confounding by Severity

Sunday Clark, MPH, Thomas Costantino, MD, Gail Rudnitsky, MD, Carlos A. Camargo Jr., MD, DrPH

Abstract

Objectives: To demonstrate the problem of "confounding by severity" using the example of intravenous (IV) versus oral corticosteroids for the treatment of acute asthma. Double-blind, randomized trials have clearly demonstrated that IV and oral corticosteroids have comparable efficacy. **Methods:** Using a standardized protocol, 64 emergency departments enrolled 1,847 patients, aged 18–54 years, with acute asthma. Because route of corticosteroid therapy was not randomized, potential confounders of the association between corticosteroid route and hospital admission were controlled for by multivariate logistic regression and stratification. **Results:** Among the 1,193 patients, 383 (32%) received IV corticosteroids and 810 (68%) received oral corticosteroids. The two groups differed markedly at baseline, with patients receiving IV corticosteroids having more severe asthma. Overall, patients receiving IV corticosteroids were more likely to be admitted or experience a relapse event within 48 hours (51% vs. 19%; p < 0.001). On multivariate analysis, patients receiving IV corticosteroids remained more likely to be admitted or experience a relapse event within 48 hours (odds ratio = 2.6; 95% confidence interval = 1.2 to 6.0). **Conclusions:** In this observational study, patients with worse asthma exacerbations were more likely to receive IV corticosteroids as compared with oral corticosteroids. Although we controlled for many markers of asthma severity, we were unable to completely control for baseline differences between the IV and oral corticosteroid groups. Observational research continues to serve as an important tool for describing problems and for understanding many exposure–disease associations. For examining the impact of treatments on adverse outcomes, randomized trials are often required to avoid intractable confounding by severity. **Key words:** asthma; corticosteroids; confounding by severity. ACADEMIC EMERGENCY MEDICINE 2005; 12:439–445.

Trying to adjust/control for suspected confounders

TABLE 2. Relation between Route of Corticosteroid Therapy (Intravenous versus Oral) and Risk of Hospital Admission or Relapse within 48 Hours of ED Presentation (*n* = 1,193)

	Odds Ratio	95% CI	p-value
Model 1 (adjusted for age)	4.4	3.4, 5.8	<0.001
Model 2 (adjusted for age + initial PEF)	5.0	3.6, 6.9	<0.001
Model 3 (adjusted for above + 13 factors)*	3.1	2.0, 4.7	<0.001
Model 4 (adjusted for above + corticosteroid variables)†	2.6	1.2, 6.0	0.02

PEF = peak expiratory flow.
*Multivariate model includes age, route of corticosteroid therapy, initial PEF, and 13 factors (gender, insurance status, recent use of home nebulizer, history of intubation for asthma, recent use of other asthma medications, recent corticosteroid treatments, interaction of corticosteroid treatment and intubation, hospital admission for asthma during past year, duration of symptoms, severity of asthma symptoms during past 24 hours, respiratory rate, inhaled β-agonist treatments during entire ED stay, and change in PEF).
†Corticosteroid variables include time of ED corticosteroid treatment and dose of corticosteroid treatment. Each intravenous dose of methylprednisolone was multiplied by five fourths to create a dose equivalent to one dose of prednisone.

groups. Observational research continues to serve as an important tool for describing problems and for understanding many exposure–disease associations. For examining the impact of treatments on adverse outcomes, randomized trials are often required to avoid intractable confounding by severity. **Key words:** asthma; corticosteroids; confounding by severity. ACADEMIC EMERGENCY MEDICINE 2005; 12:439–445.

Known confounders should be formally addressed in the design and analysis phases

Randomization is designed to minimize the impact of "unknown" confounders

Handling potential Confounders

- Conduct literature review to ascertain currently known risk factors
- Collect data on known risk factors and other potential confounders
- Compare cases and controls in prevalence of potential confounders: "Table 1"
- Identify associations of potential confounders with risk factor of interest

- *Aschengrau and Seage, Essentials of Epidemiology in Public Health, 2003.*

Dealing with Confounders

In design:

- Randomize
- Restrict: confine study subjects to those within specified category of confounder
- Match: select cases and controls so confounders equally distributed

In analysis:

- Standardize: for age, gender, time
- Stratify: separate sample into sub-samples according to specified criteria (binning?)
- Multivariate analysis: adjust for many confounders
- *Aschengrau and Seage, Essentials oin Public Health, 2003.f Epidemiology*

The goals of properly done research

1. Minimize the effects of **random error**.
2. Eliminate sources of **systematic error**.
3. Identify and adjust for possible **confounding**

All are designed to bring the investigator closer to the real truth.

Steps to minimize types of error in research studies

- **Random error**
- **Systematic error (bias)**
- **Confounding**

Steps to minimize random error

- **Meticulous attention to detail in performing the study**
  - **Dual data entry**
- **Reliability measurements**
  - Kappa calculations, etc.
- **Sample size calculations**
  - **To minimize Type II errors**
- **P value testing/confidence intervals**
  - **To minimize Type I errors**

Steps to minimize Systematic error

- **Knowledge of proper study design that minimizes potential sources of bias**
- **Selecting a proper study design**
- **True probability sampling**
- **Randomization (when able)**
- **Blinding (when relevant and able)**

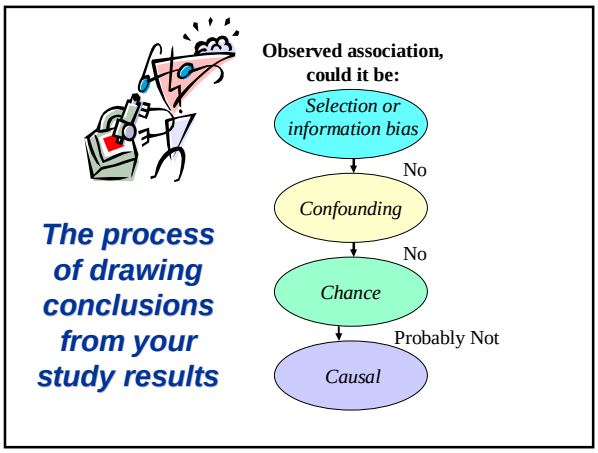
Steps to minimize confounding

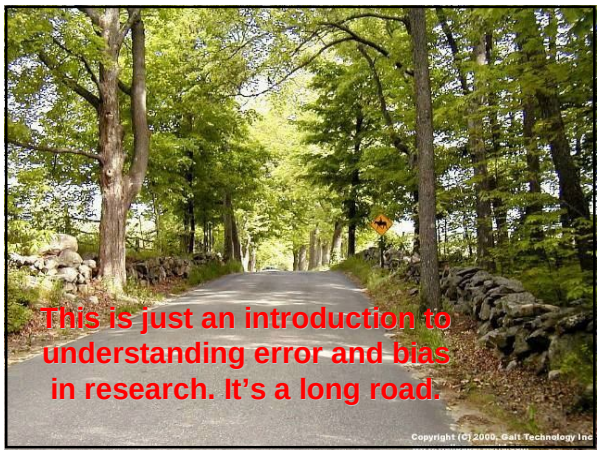
Design phase:

- **Predict extraneous variables that may function as confounders**
- **? Restrict those patients or**
- **Measure those variables precisely**
- **Randomize**

Analysis phase:

- **Stratified analyses**
- **Adjusted (regression) analysis**


Observed association, could it be:
Selection or information bias
No
Confounding
No
Chance
Probably Not
Causal
The process of drawing conclusions from your study results


This is just an introduction to understanding error and bias in research. It's a long road.
Copyright (C) 2000, Galt Technology Inc