Similarly, a health care system that is driven by robust comparative clinical evidence will save lives and money. One success story is Cochrane Collaboration, a nonprofit group that evaluates medical research. Cochrane performs systematic, evidence-based reviews of medical literature.

**Treatment**

Randomization
- Å 1923, introduced by Fisher for agriculture
- Å 1948, Hill, streptomycin for Tuberculosis
- Å 1962, thalidomide crisis
- Å 1966, the first dental RCT

**Sequence generation**

<table>
<thead>
<tr>
<th>Acceptable</th>
<th>Unacceptable</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Table of random numbers</td>
<td></td>
</tr>
<tr>
<td>- Computer generated randomization list</td>
<td></td>
</tr>
<tr>
<td>- Coin toss</td>
<td></td>
</tr>
<tr>
<td>- Systematic or alternate treatment assignment (ABAB, ABAB, ...)</td>
<td></td>
</tr>
<tr>
<td>- Hospital numbers</td>
<td></td>
</tr>
<tr>
<td>- Date of birth</td>
<td></td>
</tr>
</tbody>
</table>

**Allocation concealment**

- Å No foreknowledge
- Å The decision to accept or reject a participant should be made, and informed consent should be obtained from the participant, in ignorance of the next assignment in the sequence ð

**Periodontal trials**

- 7% of the trials reported allocation concealment
Question

- RCTs evaluating treatments for acute MI:
  - If you ______ the investigator towards treatment assignment: 8.8% chance of finding significant results
  - If you do not ______ the investigator towards treatment assignment: 24.4% chance of finding significant results

- Non RCTs: 58.1% report significant results


Clinical Trials 101

About 1/3 of the clinical trials published in elite medical journals appear not to ensure that patients are assigned to different treatments by chance.

Attempts by physicians to circumvent randomization are not isolated events; they’re part of an endemic problem stemming from ignorance.

Science, 1994

Cherry picking

The impact of this is like introducing a drug like penicillin.

USA Today, February 14, 1991

Gram-negative + Shock

<table>
<thead>
<tr>
<th></th>
<th>Death +</th>
<th>Death -</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA-1A</td>
<td>18 (33%)</td>
<td>36</td>
</tr>
<tr>
<td>Placebo</td>
<td>27 (57%)</td>
<td>47</td>
</tr>
</tbody>
</table>

$60 \text{ N=101 patients}$

NEJM

“Treatment of gram-negative bacteremia and septic shock with HA-1A human monoclonal antibody against endotoxin. A randomized, double-blind, placebo-controlled trial. The HA-1A Sepsis Study Group.”

A new RCT

"Treatment of Septic Shock with Human Monoclonal Antibody HA-1A A Randomized, Double-Blind, Placebo-Controlled Trial "

Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials

On spurious subgroup effects
<table>
<thead>
<tr>
<th>Study</th>
<th>Subgroup Benefit</th>
<th>Prior Hypothesis</th>
<th>Subgroup Benefit</th>
<th>Prior Hypothesis</th>
<th>Confirmed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barber</td>
<td>Tachycardia (&gt; 100 beats)</td>
<td>no</td>
<td>Barber</td>
<td>Tachycardia</td>
<td>no</td>
</tr>
<tr>
<td>MIAM</td>
<td>High risk patients</td>
<td>no</td>
<td>MIAM</td>
<td>High risk patients</td>
<td>no</td>
</tr>
<tr>
<td>Anders</td>
<td>Beneficial &lt; 65 yrs; harmful &gt; 65 yrs</td>
<td>no</td>
<td>Anders</td>
<td>&lt; 65; &gt; 65</td>
<td>no</td>
</tr>
<tr>
<td>Hjalm.</td>
<td>Heart rate &gt; 65 beats/min</td>
<td>no</td>
<td>Hjalm.</td>
<td>HR &gt; 65 beats/min</td>
<td>no</td>
</tr>
<tr>
<td>Wilh.</td>
<td>Electrical or mechanical complications</td>
<td>no</td>
<td>Wilh.</td>
<td>Complications</td>
<td>no</td>
</tr>
<tr>
<td>Multi</td>
<td>Anterior myocardial infarction</td>
<td>no</td>
<td>Multi</td>
<td>Anterior MI</td>
<td>no</td>
</tr>
<tr>
<td>Taylor</td>
<td>Beneficial &lt; 6 month of MI; otherwise harmful</td>
<td>no</td>
<td>台</td>
<td>&lt; 6 m; - if tx &gt; 6m</td>
<td>no</td>
</tr>
<tr>
<td>Beta-b.</td>
<td>Electrical or mechanical complications</td>
<td>no</td>
<td>Beta-b.</td>
<td>complications</td>
<td>no</td>
</tr>
<tr>
<td>Yusuf</td>
<td>B-blockers without ISA</td>
<td>no</td>
<td>Yusuf</td>
<td>B-blockers without ISA</td>
<td>no</td>
</tr>
</tbody>
</table>

### ISIS-2 trial (ISIS Collaborative Group)

- 17,187 patients with acute MI
- Randomization to placebo, aspirin, streptokinase, aspirin + streptokinase
- Aspirin reduced mortality 23%
- Streptokinase reduced mortality 25%
- Aspirin + streptokinase reduced mortality 42%

Ideas generated by data are unreliable

"For example, subdivision of the patients in ISIS-2 with respect to their astrological birth sign appears to indicate that for patients born under Gemini or Libra there was a slightly adverse effect of aspirin on mortality (9% SD 13 increase NS) while for patients born under all other astrological signs there was a strikingly beneficial effect (28% SD 5 reduction; p <0.00001)."

ISIS-2 Collaborative Group, Lancet, 1988

Rachel Nowak, Science, 264, p. 1538 1994
HIV vaccination

Next, the researchers injected a lethal strain of SIV into two animals. Neither monkey became infected with the new virus. "These animals were strongly protected—as well protected as we have seen with any live attenuated strains that we have studied," Desrosiers said.

Desrosiers’s talk bowled over many researchers. "I was jumping out of my seat," says Susan Barnett, the principal investigator on an AIDS vaccine project at Chiron, a biotech company in Emeryville, California.

Next, the researchers injected a lethal strain of SIV into two animals. Neither monkey became infected with the new virus. "These animals were strongly protected—as well protected as we have seen with any live attenuated strains that we have studied," Desrosiers said.

Desrosiers’s talk bowled over many researchers. "I was jumping out of my seat," says Susan Barnett, the principal investigator on an AIDS vaccine project at Chiron, a biotech company in Emeryville, California.
Coronary Drug Project

- Goal: to evaluate the efficacy of clofibrate
- Mortality among subjects compliant with clofibrate: 15%
- Mortality among subjects not compliant with clofibrate: 24%
- The difference is highly significant (P<0.00011)

Coronary Drug Project

- Mortality among subjects compliant with placebo: 15%
- Mortality among subjects not compliant with placebo: 28%
- The difference is highly significant (P<0.00011)

Massaging ñCòin PICO

Exploration of exposures
Pre-trial defined exposure-comparison led to negative findings; explore other exposure comparisons

Coffee and Cancer of the Pancreas

- The relative risk associated with drinking > 2 cups per day was 2.7 (1.6 to 4.7).
- Coffee accounts for 50% of pancreatic cancers.
Massaging ðOèin PICO

Exploration of outcomes

The Effect of Vitamin E and Beta Carotene on the Incidence of Lung Cancer and Other Cancers in Male Smokers

Beta Carotene Cancer Prevention Study Group
The Alpha-Tocopherol

The largest individually randomized cancer prevention trial ever conducted is the SELECT trial.12

There were statistically non-significant increased risks of prostate cancer in the vitamin E group (P=.06)

JAMA, January 7, 2009;Vol 301, No. 1
Outcome

- Maternal periodontal disease is associated with an increased risk for low birth weight.
- A cohort of 1,115 healthy pregnant women were enrolled at less than 26 weeks' gestation and followed until delivery.
- Pre-eclampsia

Pocket depth versus attachment level

Cherry Picking?

Measures of clinical significance?

<table>
<thead>
<tr>
<th>Year</th>
<th>Clinical Significance</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1970</td>
<td>Pocket depth</td>
<td>If the pocket depth is more than 5 mm, the chances for failure are so great that there is an obvious indication for surgical pocket elimination. J. Waerhaug</td>
</tr>
<tr>
<td>1970-1992</td>
<td>Attachment loss</td>
<td>The most desirable basis for evaluation of results after any treatment is to measure the level of attachment. S.P. Ramfjord</td>
</tr>
<tr>
<td>1992</td>
<td>Pocket depth</td>
<td>Clinical attachment level was included as a safety assessment. J. Periodontol 2001 p. 1537</td>
</tr>
</tbody>
</table>

Measures of clinical significance?

<table>
<thead>
<tr>
<th>Year</th>
<th>Clinical Significance</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1970</td>
<td>Pocket depth</td>
<td>If the pocket depth is more than 5 mm, the chances for failure are so great that there is an obvious indication for surgical pocket elimination. J. Waerhaug</td>
</tr>
<tr>
<td>1970-1992</td>
<td>Attachment loss</td>
<td>The most desirable basis for evaluation of results after any treatment is to measure the level of attachment. S.P. Ramfjord</td>
</tr>
<tr>
<td>1992</td>
<td>Pocket depth</td>
<td>Clinical attachment level was included as a safety assessment. J. Periodontol 2001 p. 1537</td>
</tr>
</tbody>
</table>

Measures of clinical significance?

<table>
<thead>
<tr>
<th>Year</th>
<th>Clinical Significance</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1970</td>
<td>Pocket depth</td>
<td>If the pocket depth is more than 5 mm, the chances for failure are so great that there is an obvious indication for surgical pocket elimination. J. Waerhaug</td>
</tr>
<tr>
<td>1970-1992</td>
<td>Attachment loss</td>
<td>The most desirable basis for evaluation of results after any treatment is to measure the level of attachment. S.P. Ramfjord</td>
</tr>
<tr>
<td>1992</td>
<td>Pocket depth</td>
<td>Clinical attachment level was included as a safety assessment. J. Periodontol 2001 p. 1537</td>
</tr>
</tbody>
</table>

Systematic Review: Local Controlled Release Antimicrobials

Hanes, Ann Perio, 2003
What if neither PD, nor AL, nor recession are affected by the affected treatment?

The results of this study also highlight the limitations of CAL as a measure of periodontal regeneration. This point is illustrated by the comparable gains in CAL achieved in test and control groups despite significant differences in radiographic evidence of bone fill.

Procrusteus

- Post-study generated hypotheses
  - Opportunistic data torturing
  - Procrustean data torturing
- Biological rationales are fitted to explain the observed results. Yusuf compared post-study generated hypotheses as betting on the horse after the race is over
- Post-study generated hypotheses should be regarded with suspicion (need for confirmation)

Blinding (masking) and placebo/nocebo effects
Earliest description of placebo effect?

I never knew any advantage from electricity in palsies that was permanent. And how far the apparent temporary advantage might arise from the exercise in the patient’s journey, and coming daily to my house, or from the spirits given by the hope of success, enabling them to exert more strength in moving their limbs, I will not pretend to say.

Placebo effects

Responses observed in patients which are due to factors other than the treatment itself.

- Expectancy and belief of patient and physician
- Optimism, strong conviction, persuasive abilities of the therapist concerning anticipated positive responses

Size of Placebo effects

Five treatments were identified with the following characteristics:

- Strong positive reports by at least 2 groups of investigators
- At least one well controlled negative study
- Treatment has been abandoned
- Classification of outcome as excellent, good, and poor is possible

Clinical Psychology Review, 1993

Responses to Ineffective Treatments

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Participants</th>
<th>Treats</th>
<th>DFF</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td>1002</td>
<td>27</td>
<td>-0.07</td>
<td>-0.40 to -0.15</td>
</tr>
<tr>
<td>Obesity</td>
<td>126</td>
<td>6</td>
<td>-0.04</td>
<td>-0.02 to 0.12</td>
</tr>
<tr>
<td>Asthma</td>
<td>81</td>
<td>3</td>
<td>-0.03</td>
<td>-0.03 to 0.14</td>
</tr>
<tr>
<td>Hypertension</td>
<td>129</td>
<td>7</td>
<td>-0.32</td>
<td>-0.78 to 0.13</td>
</tr>
<tr>
<td>Insomnia</td>
<td>100</td>
<td>5</td>
<td>-0.66</td>
<td>-0.66 to 0.13</td>
</tr>
<tr>
<td>Anxiety</td>
<td>257</td>
<td>6</td>
<td>-0.06</td>
<td>-0.31 to 0.18</td>
</tr>
</tbody>
</table>

NEJM 2001; 344:1594-1602
Study of the Therapeutic Effects of Intercessory Prayer (STEP) in cardiac bypass patients: a multicenter randomized trial of uncertainty and certainty of receiving intercessory prayer.

Am Heart J. 2006 Apr;151(4):934-42

Blinded intercessory prayer (604)
Blinded no intercessory prayer (597)
Unblinded intercessory prayer (601).

Cardiac Complications

<table>
<thead>
<tr>
<th></th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Blinded</td>
<td>352 (59%)</td>
<td>249</td>
</tr>
<tr>
<td>Blinded</td>
<td>315 (52%)</td>
<td>289</td>
</tr>
</tbody>
</table>

Major events and 30-day mortality were similar across the 3 groups.

If nocebo/placebo-effects are real

- Double-blind clinical trials
  - Patient
  - Clinician: Performance bias (Expertise-based RCTs)
  - Examiner: Ascertainment bias
  - Statistician: Data analyst bias
- Elimination of placebo/nocebo responders from trials
- Hard endpoints
Knee Surgery

• Arthritis surgery
• 225,000 procedures/year
• Increase in pain in lavage and debridement group

NYT, July 11, 2002

Parkinson’s Disease

• Parkinson Disease
• Use of Placebo surgery in controlled trials of a cellular-based therapy of Parkinson’s disease

NEJM, 1999

Migraine Headache

<table>
<thead>
<tr>
<th></th>
<th>Closure</th>
<th>Sham</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFO</td>
<td>+ 73</td>
<td>- 74</td>
</tr>
<tr>
<td>Closure</td>
<td>3 (4%)</td>
<td>71</td>
</tr>
<tr>
<td>Sham</td>
<td>3 (4%)</td>
<td>70</td>
</tr>
</tbody>
</table>

A Randomized Trial of Vertebroplasty for Painful Osteoporotic Vertebral Fractures

NEJM
Volume 361:557-568 August 6, 2009 Number 6

Backpain

<table>
<thead>
<tr>
<th></th>
<th>Surgery</th>
<th>Vertebroplasty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Worse</td>
<td>7 (20%)</td>
</tr>
<tr>
<td></td>
<td>Better</td>
<td>28</td>
</tr>
<tr>
<td>Sham</td>
<td>5 (14%)</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td></td>
</tr>
</tbody>
</table>

Trephination and pain

“The mock trephination procedure mimicked the actual trephination procedure except that the perforator only penetrated gingival tissue (not bone) and the file handles were placed on the gingival tissue and turned to mimic the trephination procedure. The spoon excavator was placed through the gingival perforation and moved across bone in a scraping motion. All instruments were used in the same sequence and for the same time period as during the actual trephination.”

OOO, 2000; 90: 507-13
Active Albuterol or Placebo, Sham Acupuncture, or No Intervention in Asthma

![Graph showing improvement in asthma](image1)

Arch Intern Med. 2009 May 11;169(9):858-66.

A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain.

![Graph showing pain reduction over time](image2)

Eliminate Placebo responders

![Image of WSJ article](image3)

No Masking and endpoint choice

- North American Symptomatic Carotid Endarterectomy Trial Collaborators
- 50 clinical centers throughout the United States and Canada
- Carotid endarterectomy versus aspirin for prophylaxis against stroke
- Endpoint: TIAs versus morbidity

![Image of Nature article](image4)

The “softer the endpoint, the more masking is needed

- Antibiotic prophylaxis and flare-ups
  - Morse and colleagues find a reduction of flare-ups from 20% to 2% associated with antibiotic prophylaxis
  - Two independent investigators did not find an effect
- Flossing and interproximal caries
http://www.consort-statement.org/

- Approval by institutional review board
- Funding of trial
- Randomized controlled trial Number used to register trial at inception
- Cluster randomization
- Equivalence trials

Item 12a - Statistical methods used to compare groups for primary and secondary outcomes

It is customary to present a 95% confidence interval, which gives the range expected to include the true value in 95 of 100 similar studies.

The principal comparisons between the chewing and nonchewing regimens used one-tailed statistical tests of hypotheses.

THE EFFECT OF CHEWING SUGAR-FREE GUM AFTER MEALS ON CLINICAL CARIES INCIDENCE JADA

Item 19 - All important harms or unintended effects in each group

http://www.prisma-statement.org/

IOM Sets Out “Gold Standard” Practices for Creating Guidelines, Systematic Reviews

http://www.sealedenvelope.com/power_binary.php
<table>
<thead>
<tr>
<th>TITLE</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>1</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>1</td>
</tr>
<tr>
<td>Introduction summary</td>
<td>1</td>
</tr>
<tr>
<td>Methods</td>
<td>1</td>
</tr>
<tr>
<td>RESULTS</td>
<td>1</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>1</td>
</tr>
</tbody>
</table>

### METHODS

#### Final study selection

- Final studies selected for data extraction: 40 studies.

#### Analysis procedures

- Data extraction on: 40 studies.

#### Synthesis of results

- Synthesis of results: final analysis.

### RESULTS

#### Final analysis

- Final analysis: final analysis.

#### Final study selection

- Final studies selected for data extraction: 40 studies.

#### Analysis procedures

- Data extraction on: 40 studies.

#### Synthesis of results

- Synthesis of results: final analysis.

### REFERENCES

- Verges et al.
- Xiong et al.
- Vettore et al.
Serious doubts about the existence of an association between periodontal disease and low birth weight. A likely association.

Important risk factors were not taken into account in any study on periodontal disease and adverse outcomes in pregnancy.

é silenceé.

Conclusion: The balance of studies does not demonstrate a relationship between sugar quantity, but a moderately significant relationship of sugar frequency to dental caries.

An ‘A’ paper (26) gave data on a cross-sectional study carried out on a population of Spanish school children aged 5–15 years old.

Ethics and RCTs

Beneficence
Respect for persons

- Informed consent
- "To tell the truth, all of the discussion today about the patient's informed consent still strikes me as absolute rubbish." (Hill)

Pediatric Research

- Men > Women > Children > Pregnant women
- 1996 workshop on inclusion of children (NICHD and American Academy of Pediatrics)
- 1997 NIH plans to develop a policy for the inclusion of children
- March 6, 1998

Justice

- Equitable distribution of burdens and benefits
  - Investigation of birth control medication in Africa
  - Septic shock trials

The use of block sections in the premolar region to investigate the effect of allografts and autografts (X and Y, 1979) in Bogota.
Equipoise

- Equipoise often is a big hurdle in RCT design
  - HRT: "The view that HRT prevented heart disease was so entrenched that during the planning of WHI, some argued that it would be unethical to deny some women the drug and give them a placebo (Science Vol. 279, 2002)."
  - flecainide, encanaide: "The major concern seemed to be whether it was ethical to randomize a subject to placebo after finding a drug that 'works' for that patient. (Washington Public Health, 1992)."
  - Calcium channel blockers: Compared with diuretics users, the adjusted risk for myocardial infarction was associated with about a 60% mortality increase among users of calcium channel blockers (Psaty et al., JAMA, 1995).
  - Dental examples: Scaling & Root planing, diagnostic radiation, safety of water fluoridation

While the risk of infective endocarditis in patients with mitral valve prolapse is 1 / 1.1 million dental procedures, the guidelines have been revised 9 times since 1955

Hypothesis Testing and Error rates

P-values and The role of chance

Toothbrush experiment

- 30 subjects
- 5 have bad oral hygiene
- random assignment to 2 treatment groups
- The two treatments under investigation are identical

By chance, 4 individuals with bad oral hygiene may fall into one of the experimental groups leading to an odds ratio of 4*14/1*11 or 5.1

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>0.07</td>
</tr>
<tr>
<td>5%</td>
<td>0.16</td>
</tr>
<tr>
<td>10%</td>
<td>0.22</td>
</tr>
<tr>
<td>50%</td>
<td>1.00</td>
</tr>
<tr>
<td>90%</td>
<td>4.33</td>
</tr>
<tr>
<td>95%</td>
<td>6.33</td>
</tr>
<tr>
<td>99%</td>
<td>12.70</td>
</tr>
</tbody>
</table>
Type I error rate

Â Assume individuals are randomly assigned to two identical treatments, A and B.
Â Since the treatments are identical, the expected odds for failure are identical in both groups, i.e., OR = 1.
Â However, due to random variability OR will vary around 1. E.g., individuals in group A may fare better or worse than individuals in group B (OR ≠ 1)

Type I error

Â When the observed OR is large (by chance!), we may mistakenly conclude that the treatments differ.
Â We may make a false-positive conclusion
Â Our rule for how small the probability for this mistake should be is the α level or level of significance

Type I error rate: 5%

Â If a trial with identical treatments is repeated 100 times, a mistaken conclusion of a significant treatment effect will be made 5 times (e.g., OR < 0.16 or > 6.33)
Â There will be five false positive findings and 95 true negative findings
Â By a 5% chance, we may stumbled onto a difference and conclude that treatments differ.

Type II error rate

Â Assume individuals are randomly assigned to two treatments that differ: the experimental is better than the control by a specified amount.
Â Since the experimental treatment is better, the odds for failure should be lower in the experimental group, i.e., OR < 1.
Â However, due to random variability OR may approach 1 or even be larger than one!

Type II error

Â When the observed OR is close to 1 or larger than one, we mistakenly accept the notion that the treatments are identical (OR=1). Our rule for how small the probability for such a mistake should be is the β level
Â We made a false negative conclusion
Â When we accept that OR=1, we make a type II error: there was a positive effect of treatment A on the outcome, but by chance we missed this difference.
Â Power= 1 - β

Type II error rate: 20%

Â If the trial is repeated 100 times, and the experimental is in truth better than the control, a mistaken conclusion of no treatment effect will be made 20 times.
Â There will be 20 false negative findings and 80 true positive findings
Type II error

- The attainable power (1-\(b\)) in a study is often the major limiting factor in the design of clinical studies.
- The power of a study depends on:
  - Sample size
  - Incidence

Major Caveat

- Type I and type II error are calculated based on the assumption that we know that true difference between the exposed group and the control group.
- Of course, the truth is not known.

Example 1

- Assume 100 new products are investigated
- 20 truly are effective
- 80 are non-effective
- How many false-positive and false-negative answers do you expect in the 100 trials when the type I error rate is 5% and the type II error rate is 20%?

Example 2

- Assume 100 new products are investigated
- 5 truly are effective
- 95 are non-effective
- How many false-positive and false-negative answers do you expect?
Example 3

Assume you are dealing with an extremely challenging disease: i.e., melanoma. 1000s of treatment have been evaluated, none have been found to be effective. How many false-positive and false-negative answers do you expect?

<table>
<thead>
<tr>
<th>Truth regarding the exposure</th>
<th>EFFECTIVE</th>
<th>NOT EFFECTIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective</td>
<td>8 (80%)</td>
<td>59 (5%)</td>
</tr>
<tr>
<td>Not effective</td>
<td>2 (20%)</td>
<td>941 (95%)</td>
</tr>
</tbody>
</table>

59 out of 67 positive findings (88%) are false positive findings
2 out of 943 negative findings (0.2%) are false negative findings

False-positive error rates is a major concern in clinical studies.

Just as the sphinx winks if you look at it too long, so, if you perform enough significance tests you are sure to find significance, even when none exists.

Jerome Cornfield, 1976

Type I and II error rate

For chronic disease, where the identification of effective treatments that are better than the standard are infrequent, there is a need for stringent control over type I and type II errors. High false positive error rate. Ideally, both the type I and the type II error rate should be less than 5%.