



An Evidence-based Medicine (EBM) Seminar

Evidence-based Medicine: Empowering Patients, Providers & Payers to Improve Quality and Safety in Health Care

Disclosures: For Profit Companies

Delfini Group: Michael E. Stuart, MD & Sheri Ann Strite

For-profits In the last 12 months:

§ Amgen

§ Astellas

§ Genentech

§ Members of the editorial board for DynaMed,
EBSCO Publishing (unpaid)



What We Are Going to Cover

- Why Evidence-based Medicine and Critical Appraisal Are So Important for Patient Care
- Critical Appraisal Sampler
- Requirements for an Evidence-based Approach
- What Patients Need for Patient-centered Decision-making
- What Policy Makers Need for an Evidence-based Approach

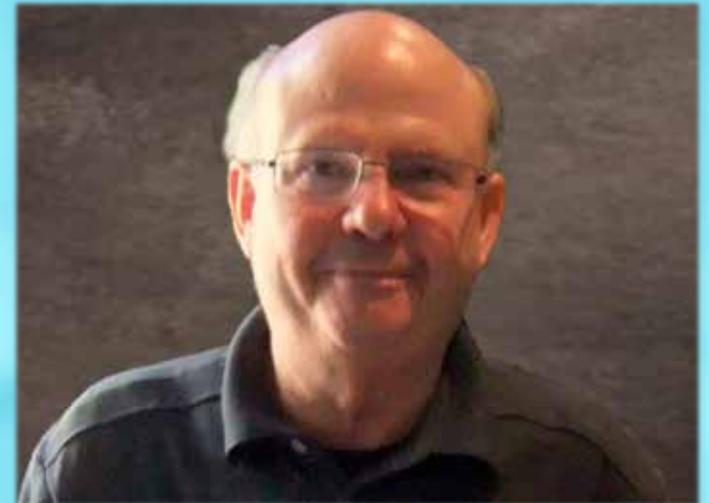
Michael E. Stuart MD & Sheri Ann Strite

- Evidence-based clinical QI experts
- Medical information scientists & evidologists
- Consultants and trainers
- Best known for training program in simplified approach to critical appraisal - 11+ years
- Guideline facilitators
- EBQI facilitators
- Text book authors & contributors



Mike Stuart MD

- Clinical faculty University of Washington
- 30 years Family Practice at Group Health Cooperative in Seattle
- Headed medical education and evidence-based clinical improvement depts
- 20 years EBM practitioner
- Deep clinical guideline development expertise



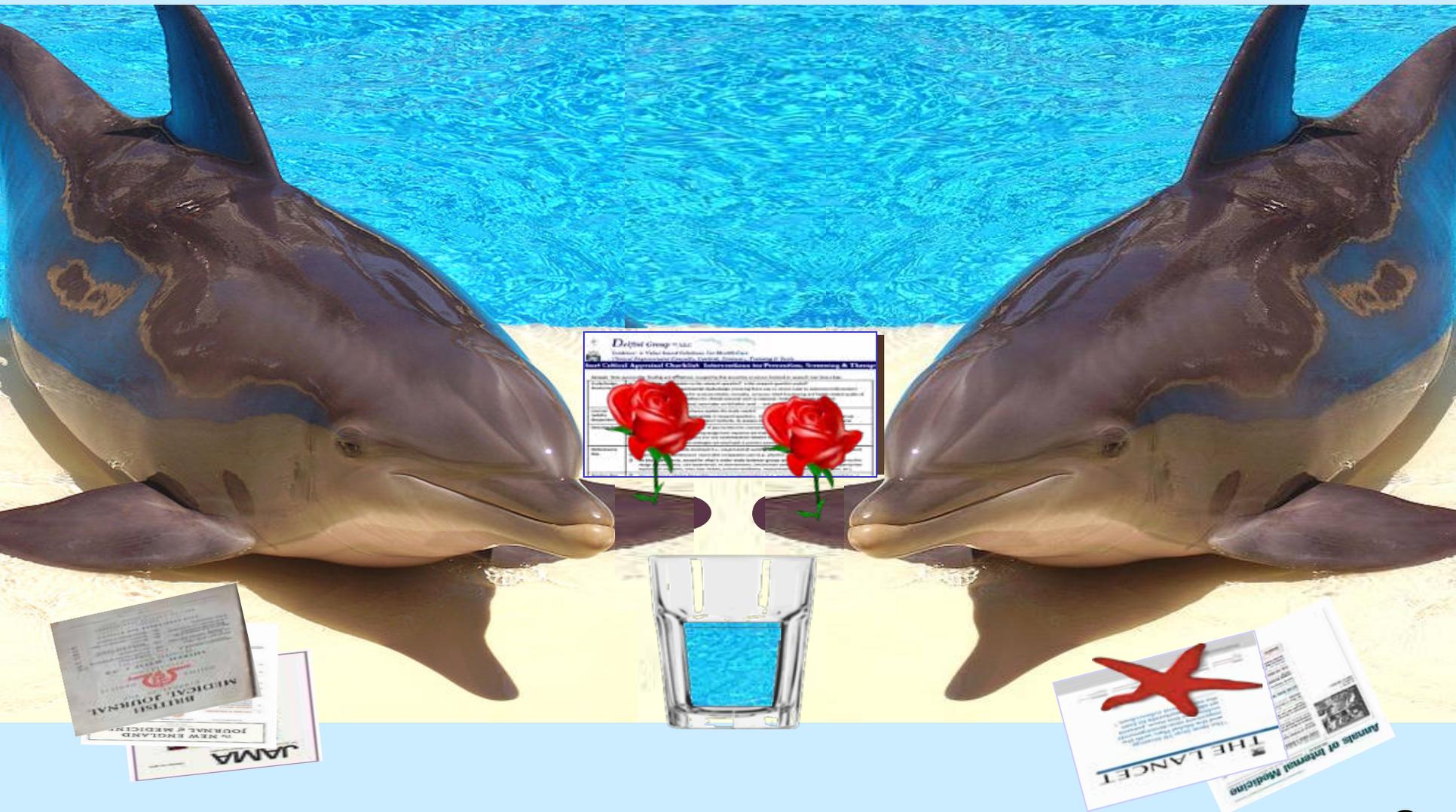


Sheri Strite

- 20 years at Group Health in various research-related roles
- 5 in medical education and clinical improvement department
- 12 in EBQI & evidology including taught EBM 3+ years in UCSD medical school
- No clinical training



Healthcare Information is Like a Big Shipwreck & We Want to Help Rescue You from Drowning





**Evidence- and Value-based Solutions for Health Care
Clinical Improvement Consults, Content Development,**

The EBM Information Quest: Is it true? Is it useful? Is it usable?

Delfini Co-founders: Michael E Stuart MD, President & Medical Director . Sheri Ann Strite, Principal & Managing Partner

Free resources at
www.delfini.org

Quick Picks



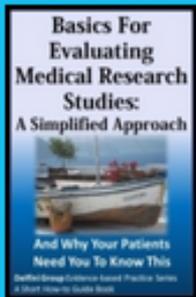
[Why Critical Appraisal Matters](#)



[Services](#)



[Seminars](#)



[Books](#)

PRACTICAL EVIDENCE-BASED PRACTICE RESOURCES & SERVICES FOR HEALTHCARE QUALITY

Helping people improve healthcare decision-making through "applied" evidence-based clinical quality improvement and communications with a focus on value and patient-centered care.

[DelfiniGram™: GET ON OUR UPDATE LIST](#) 

• [Resources](#)

We are a public service entrepreneurship and make much of our work freely available online.

• [Solutions](#)

Read about the [problems with healthcare information](#) and our solutions.

• [Help](#)

See our [services](#).

To navigate our website—

- See our Quick Picks to selected sections to the left.
- See our Main Menu to the right.
- Beyond our Main Menu, you will find our Quick Navigator to some of our key resources.

Also see [Delfini Group Publishing](#).

What people say about Delfini—

At Home



[Read Our Blog...](#)

Menu

Use of our website implies agreement to our [Notices](#). Citations for references available upon request.

- Home
- What's New
- Blog
- Seminars
- Services
- Delfini Group Publishing
- Resources
- Sample Projects
- Notices
- About Us & Our Work
- Testimonials
- Other
- Site Search

Quick Navigator

change... real success... improved patient care... valuable...

"...thanks for presenting this valuable information in a fun and memorable way..."

"...I thought the subject was going to be boring and not much use...I was very wrong...I cannot thank you and Sheri enough."

"I know a couple hundred people...who have got to hear your message!"

"Best one-day educational program I've experienced in 25 years!"

Delfini evidence reviews are consistently praised as useful, rigorous and transparent. Delfini average training program scores are 4.8 out of 5.0.

"... gifted facilitators and passionate educators... have rarely encountered presenters so well-prepared and well-versed in their subject matter... both are engaging, up-beat and empowering... extremely generous with their knowledge and expertise..."

See all of our [Testimonials](#).



[Updates & Contact Info](#)



[Free Online Tools](#)



[Free Online Tutorial](#)



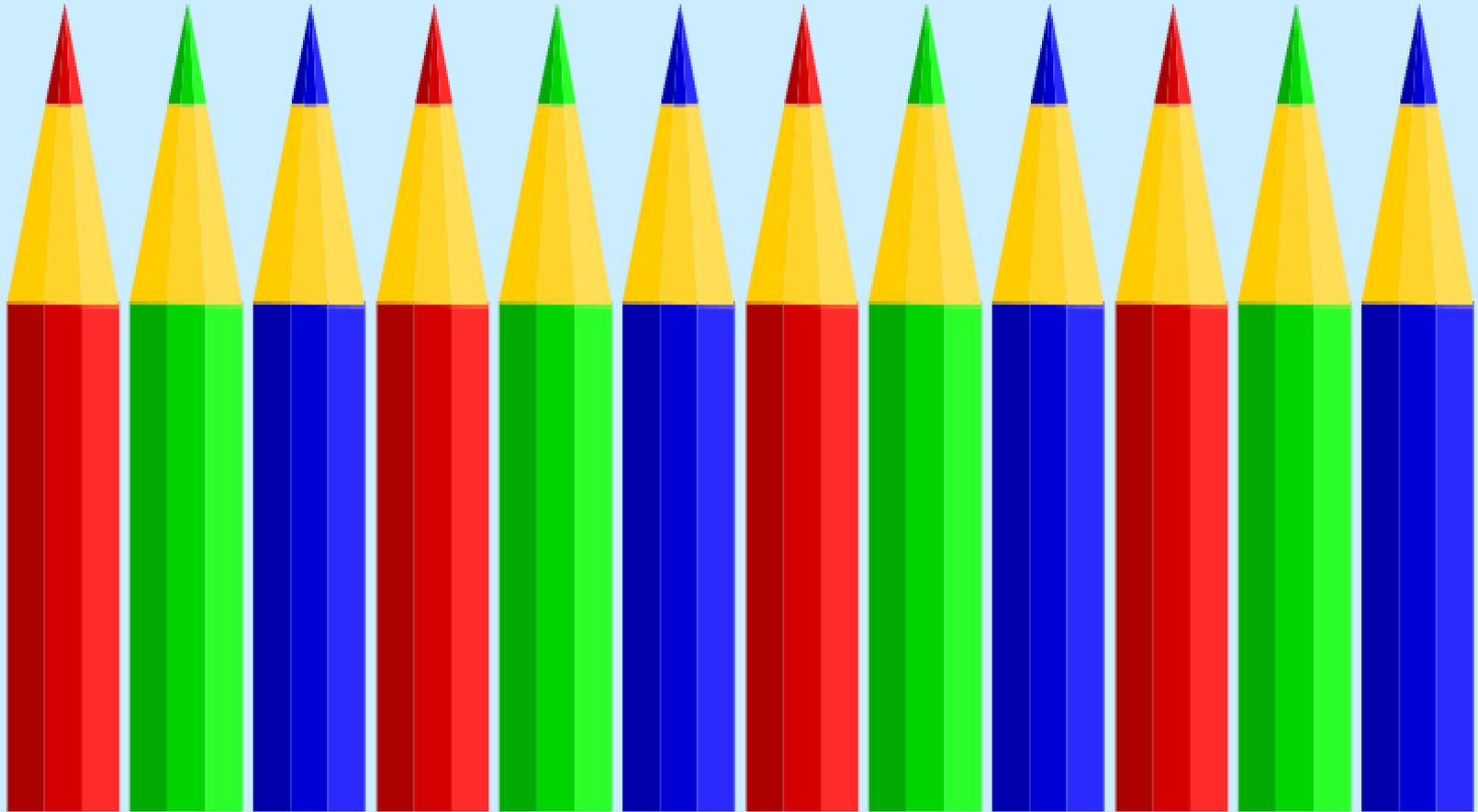
[Delfini Blog](#)

[Testimonials](#)
[Other](#)
[Site Search](#)
[Contact Info](#)
[Updates](#)

.....
[Quick Navigator to Selected Resources](#)

- . [Tools & Educational Library](#)
- . [The Healthcare Information Problem](#)
- . [Tutorials](#)
- . [DelfiniClick™—commentaries](#)
- . [On the Same Page™—patient decision support](#)
- . [Evidence Messaging Scripts](#)
- . [Delfini Publications](#)
- . [Medical Leader Interviews](#)
- . [Glossary](#)
- . [Recommended Reading](#)
- . [Recommended Web Links](#)
- . [Delfini Critical Appraisal Flier](#)
- . [Get DelfiniGram™ Updates](#)

A Couple of Colorful Points Before We Start...



We've Taken Notes For You

Understanding Medical Literature



Discuss this Hypothetical

You read in the newspaper that you are 65% more likely to survive a crash if you are wearing your seatbelt than if you are not...

Delfini's Definition of EBM

“Evidence-based medicine (EBM) is the use of the scientific method and application of valid and useful science to inform health care provision, practice, evaluation and decisions.”

Delfini's Definition of Critical Appraisal

- A scientific evaluation of evidence (e.g., research data) to appraise **validity** (closeness to truth) and **usefulness** (e.g., generalizability to one's own patients or circumstances, meaningful benefit, etc).
- Critical appraisal = critical appraisal concepts + clinical knowledge + **critical thinking**

What A Critical Appraisal Looks Like

Delfini Critical Appraisal Case Study

HYPOTHETICAL CASE STUDY: MYOCEPTIMAB PREVENTS CARDIOVASCULAR MORBIDITY

Critical Appraisers & Date: Sheri A. Strite & Michael E. Stuart MD, Delfini Group; April Fool's Day 2000-Any Year

PUBLISHED ABSTRACT

Background

Elevated myoreactive protein has been demonstrated to be associated with increased risk of myocardial infarction (MI). Myoceptimab is an inhibitor of myoreactive protein and has been shown to reduce myoreactive protein levels.

Methods

We conducted a randomized, double-blind trial in the Beaverton University Heart Care Center to assess the efficacy and safety in patients ages 55 and older who were at increased risk for cardiovascular events and had elevated myoreactive protein levels above 4 mg/L on two separate occasions. Patients were randomly assigned to receive 60 mg of myoceptimab (29 patients) or placebo (35 patients) daily for 6 months.

The study outcome was cardiovascular morbidity as defined by mean reduction of elevated levels of myoreactive protein, onset of new angina, admission to the hospital for any cardiovascular-related condition, myocardial infarction, stroke, claudication, heart failure or cardiovascular death).

Results

who are at increased risk for cardiovascular events.

CRITICAL APPRAISAL

- Study size: small
- Primary endpoint: questionable composite
- Randomization: not truly randomized; patients assigned to groups by study consent date
- Concealment of allocation: no details
- Baseline characteristics: slightly higher rate of angina in the placebo group
- Blinding: insufficient details and no indication of blind assessment
- Intergroup differences: participating cardiologists were not restricted in patient management so as to replicate real-world conditions; no details of co-interventions reported between groups
- Attrition: less than 1 percent
- Safety, including long term harms, is uncertain
- Results: questionable clinical significance, selective reporting and *post-hoc* results

What's the Goal of Medical Research for Interventions?



Does this really matter?

If I take this pill, what might happen to me?



Q: Why do we want good science for medical decision-making?

1. Determine cause & effect
2. Assess likelihood of effects (probability)

Is it True? Is it Useful?

Are the reported results—

§ Reliable and

§ Clinically useful?

§ Reliability = assess validity
("closeness to truth") by assessing bias,
confounding and chance

§ Clinical usefulness = assess meaningful
clinical benefit

We “Read A Study” By Critically Appraising It

Your Question—

§ Are the reported results likely to be true and clinically useful?

Utilize—

§ Critical appraisal concepts +

§ Clinical knowledge +

§ Critical thinking ✨



Is This All Just Theoretical?



Does this really matter or can't I just—

- 1 Rely on the FDA?
3. Read only the best journals?
4. Avoid industry research
5. Find a reliable source?

Facts—

1. Can't rely on FDA
2. There is a general lack of skills
3. Much industry work is important and all researchers have biases; all studies, flaws
4. No 1-stop resource

Quality Science & Critical Appraisal Skills Make A Difference



§ Roughly 58,000 US lives lost in Vietnam War

<http://www.digitalhistory.uh.edu/modules/vietnam/index.cfm>

§ Roughly 42,000 women with advanced breast cancer subjected to autologous bone marrow transplant and high dose chemotherapy—over 9,000 died from treatment; RCTs showed no benefit and cost \$3.4 billion

Mello MM, Brennan TA. PMID: 11558695

Caveats

- § Focus is on **superiority studies** of **therapies** with emphasis on **efficacy** unless otherwise stated
- § We believe that concluding that all agents in a class have similar effects is dangerous
- § There are exceptions to everything we say
- § Judgment is required
- § Best answer to everything:
IT DEPENDS!



The Landscape

- § Primary Studies: Original research
- § Secondary Studies: Studies of studies such as systematic reviews and meta-analyses
- § Secondary Sources: Information sources that reference primary or secondary studies

Introducing PICOTS & CI

Patient/population (i.e.,
Condition)

Intervention

Comparators

Outcomes

Timing

Setting

§ PICOTS useful for—

§ Framing question

§ Describing & summarizing studies (primary and secondary)

§ Assessing heterogeneity of studies

§ Synthesizing evidence

§ Forming clinical recommendations and decision support

§ CI for searching



Some Starting Terms & Concepts

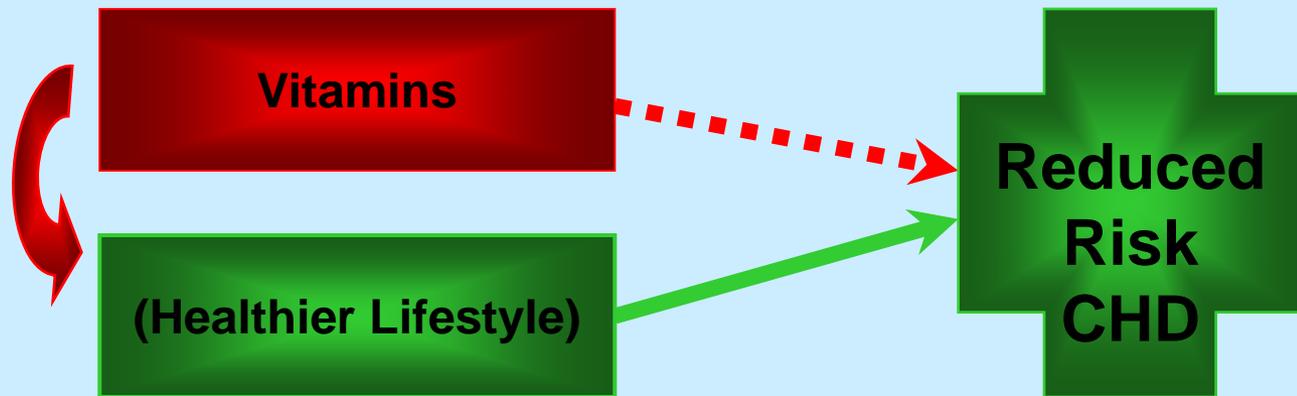
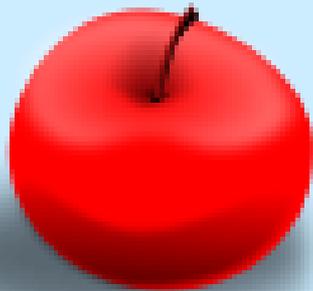
- § **Validity** = “closeness to truth”
- § **Bias** = anything that “systematically” leads away from truth
- § **Outcomes** = what we are interested in studying (e.g., reduction in mortality)
 - § Synonyms = **endpoint** or **outcome measure**
- § **Confounder** = a factor other than what you are studying that might be responsible for or affect your study’s results
- § **Number-needed-to-treat** = The number of patients who need to be treated in order for one patient to benefit over that patient taking the comparator agent within the study time period.

Confounders: Known & Unknown

A variable (the confounder) falsely appears to be the cause of the outcome instead of the true cause

Is taking vitamins responsible for reduced CHD risk? Or do people who take vitamins have healthier lifestyles?

Example



Evidence-Grading

- § An evidence grade is a **summary expression** of the quality and usability of the evidence
- § Many systems exist—understand the **meaning** behind the grade and whether the **criteria** are valid
- § What is being graded? (study, outcomes, overall strength of the evidence)
- § **Delfini System**
 - § A, B, BU and U (uncertain)
 - § U is not used for efficacy
 - § U may be used for safety
- § 1-Pager is available



We Love the PMID Number!

www.ncbi.nlm.nih.gov/pubmed?term=21081725

NCBI Resources How To My NCBI Sign In

PubMed.gov
U.S. National Library of Medicine
National Institutes of Health

Search: PubMed
21081725 Search Clear

RSS Save search Limits Advanced search Help

Display Settings: Abstract Send to:

JAMA
FULL text at

JAMA. 2010 Nov 17;304(19):2127-8; author reply 2128.
Importance of blinding in randomized trials.
Strite SA, Stuart ME.

Comment on
JAMA. 2010 Aug 18;304(7):793-4.
PMID: 21081725 [PubMed - indexed for MEDLINE]

+ Publication Types, MeSH Terms

+ LinkOut - more resources

Related citations

Minimizing bias in randomized trials: the importance of blinding. [JAMA. 2010]

Practical aspects of randomization and blinding in randomized clinical trials. [Arthroscopy. 2003]

Review Deliberate ignorance: a systematic review of blinding in nursing clinical trials. [Nurs Res. 2011]

Review Who is blinded in randomized clinical trials? A study of 200 trials and a : [Clin Trials. 2006]

Review Participant expectancies in double-blind randomized placebo-controlled tri: [Clin Trials. 2010]

See reviews...
See all...

How to Read a Forest Plot



Graphic Display: Point Estimate, CI and Summary Diamond

These are several studies reported in a meta-analysis (some studies are removed, so this is not meant to total correctly) — this is just a sampler.

Odds Ratio (95% CI)

Study A n = 50

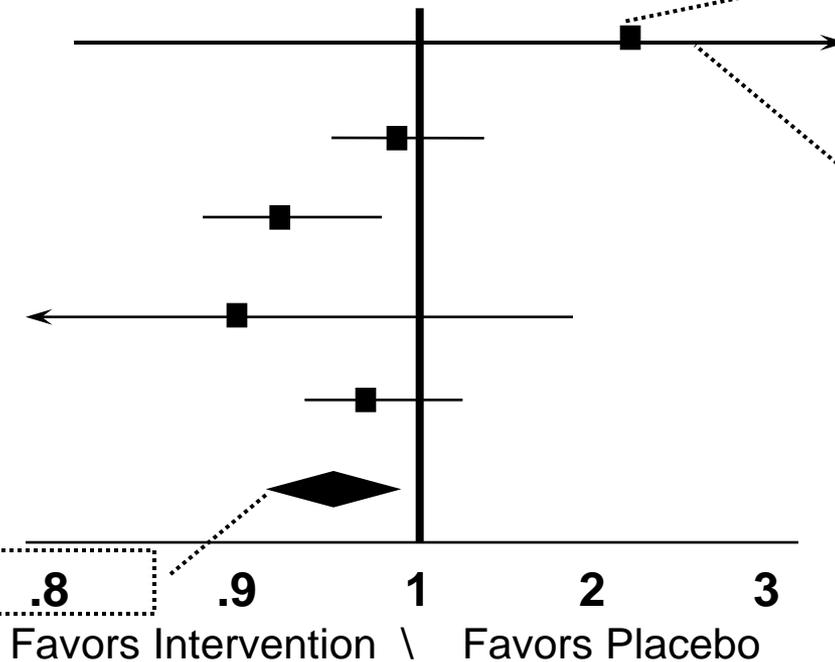
Study B n = 4500

Study C n = 1500

Study D n = 500

Study E n = 4000

Total n = 15000



This **square** is the study result (ie, the **point estimate**)

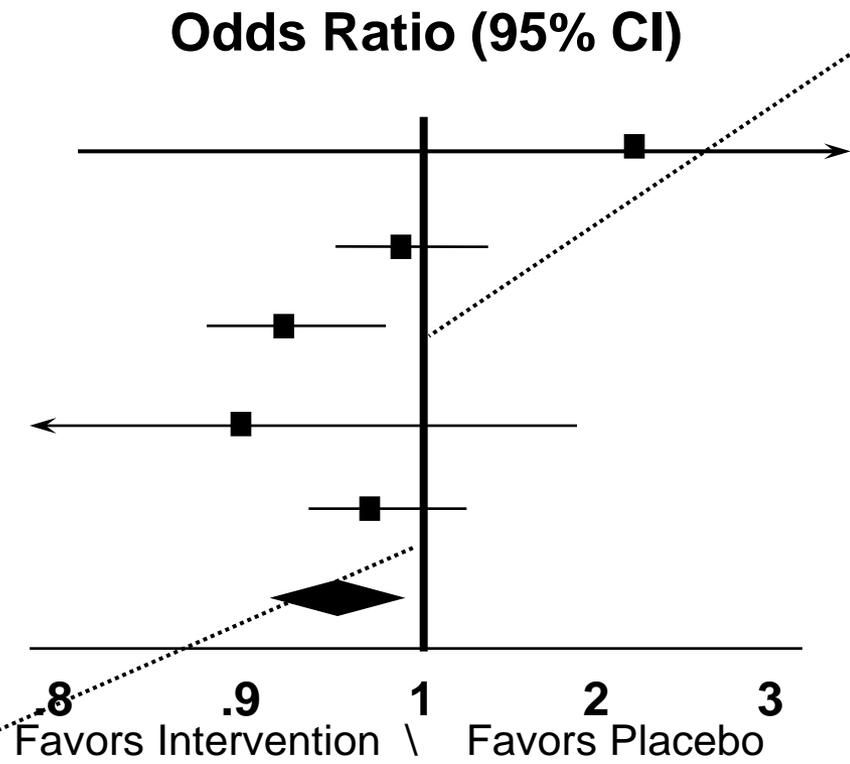
This **line** is the **confidence interval** (ie, a statistically calculated range of equally plausible study results given a **margin for chance**)

The **summary diamond**

.8 .9 1 2 3
Favors Intervention \ Favors Placebo

Favors Intervention \ Favors Placebo & The Line of No Difference

Study A n = 50
 Study B n = 4500
 Study C n = 1500
 Study D n = 500
 Study E n = 4000
 Total n = 15000



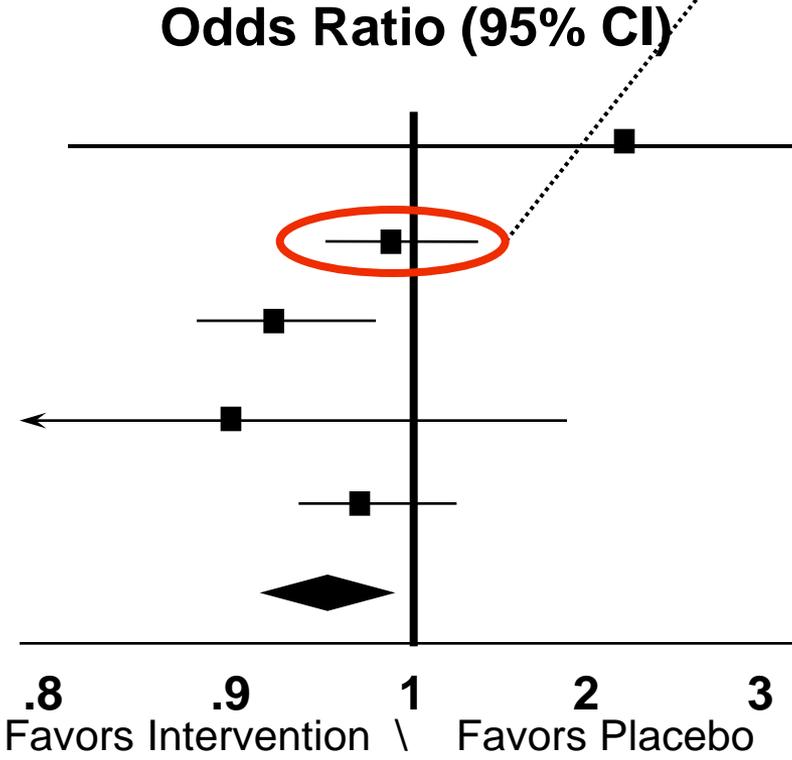
This **center line** is the **line of no difference**. Results to the right favor placebo in this example. Results to the left favor the intervention.

Synonyms:

- Line of no difference
- Line of no effect
- Infinity
- Unity

Non-Statistical Significance

Study A n = 50
 Study B n = 4500
 Study C n = 1500
 Study D n = 500
 Study E n = 4000
 Total n = 15000



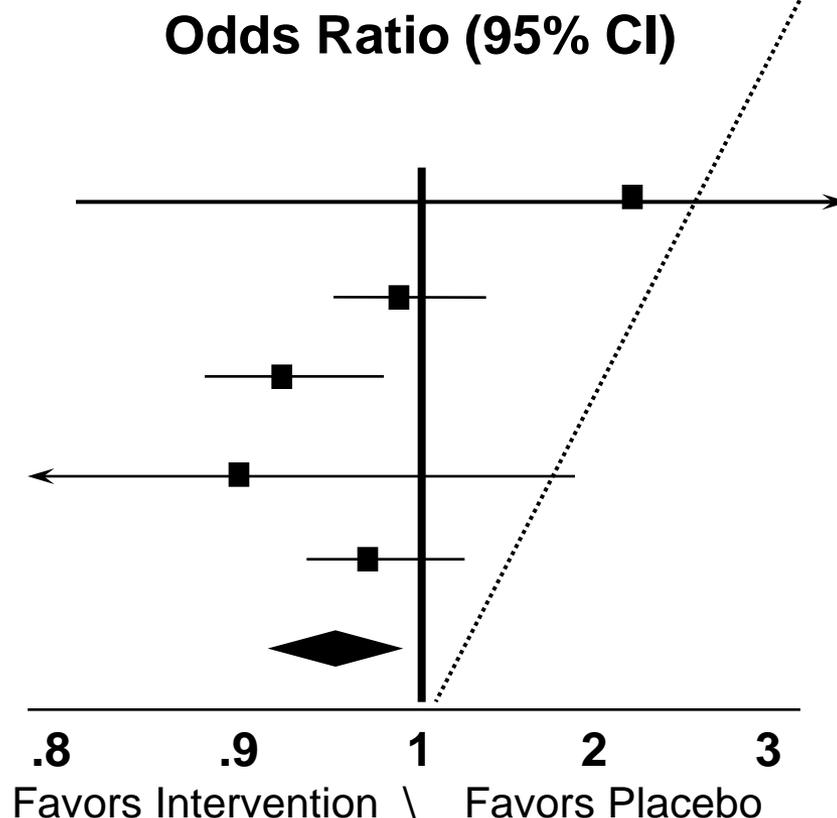
Therefore, it is statistically plausible, within 95% certainty in a valid study, that Study B may favor the placebo or Study B may favor the intervention.

This is not possible. Thus, the results of Study B are **not statistically significant**.

Anything **touching this line** means the results are **not statistically significant** because it is not possible to favor both placebo and intervention.

You Need to Know the Numerical Value for “No Difference”

Study A n = 50
Study B n = 4500
Study C n = 1500
Study D n = 500
Study E n = 4000
Total n = 15000



The line of no difference equals 0 or 1 depending upon the measure of outcome used.

No difference for a **percent** is expressed as **zero**.

ARR and **RRR** are expressed as **percentages**. Therefore, if these were used, this number would be zero.

No difference for a **ratio** is 1:1. So for **odds ratio** or **relative risk ratio** (aka **relative risk**), this number **equals 1**.

Why We Are Here

- § Critical appraisal matters
- § 1-pager on Critical Appraisal for Superiority Trials of Therapies
 - § Short Checklist

The Primary Reason To Do This...



What is it that we all want from our health care?

1. That we don't die prematurely (mortality).
2. That we don't suffer from conditions or diseases that we can avoid (morbidity).
3. That our health issues do not detract from our quality of life (health-related quality of life).
4. That we do not experience unpleasant symptoms from our health issues (symptom relief).
5. That health issues do not interfere with our daily activities (emotional, physical and mental functioning).

Health care professionals refer to these five items as "health care outcomes."

Let's Talk Outcomes

- Identify five examples of outcomes that are not one of the following:
 - morbidity; mortality; symptom relief; emotional, mental or physical functioning; and, health-related quality of life
- Discuss the importance of this question, including what are these five outcomes, and why is it important to be able to distinguish outcomes that are not one of these five.

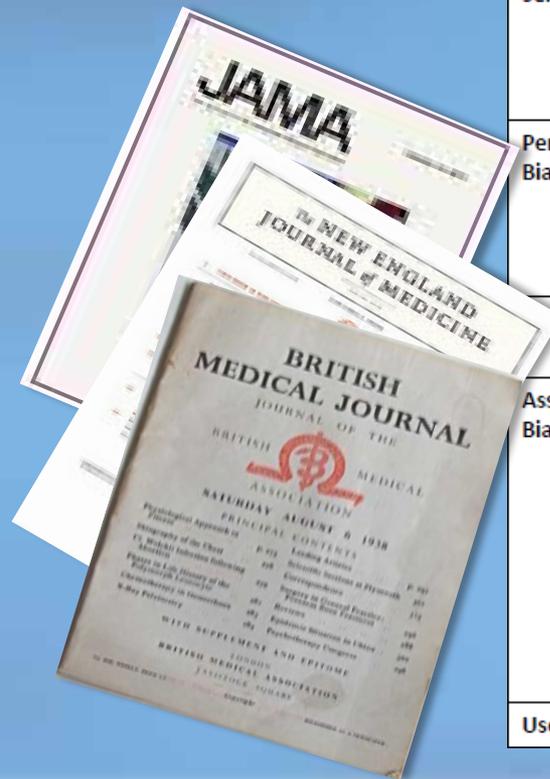
Let's Talk Outcomes

Evaluate the following composite endpoints:

- Safety: All-cause mortality, MI, stroke and rash.
- Safety: All-cause mortality, MI, stroke and admission to the hospital for cardiovascular problem.
- Diabetes: HbA1c, advanced retinal disease as determined by an ophthalmologist, end-stage renal disease (ESRD), cardiovascular outcomes, all-cause mortality

Why Critical Appraisal Skills Make a Difference for Quality Patient Care

Study Design Assessment	<ul style="list-style-type: none"> <input type="checkbox"/> Is the design appropriate to the research question? Is the research question useful? <input type="checkbox"/> For efficacy, use of experimental study design (meaning there was no choice made to determine intervention) <input type="checkbox"/> Clinically significant area for study (morbidity, mortality, symptom relief, functioning and health-related quality of life) and reasonable definitions for clinical outcome such as response, treatment success or failure <input type="checkbox"/> If composite endpoints used, reasonable combination used — and used for safety if used for efficacy
Internal Validity Assessment	<ul style="list-style-type: none"> <input type="checkbox"/> Can bias, confounding or chance explain the study results? <input type="checkbox"/> Ensure prespecified and appropriate 1) research questions, 2) populations to analyze, and 3) outcomes
Selection Bias	<ul style="list-style-type: none"> <input type="checkbox"/> Groups are appropriate for study, of appropriate size, concurrent and similar in prognostic variables <input type="checkbox"/> Methods for generating the group assignment sequence are truly random, sequencing avoids potential for anyone affecting assignment to a study arm and randomization remains intact <input type="checkbox"/> Concealment of allocation strategies are employed to prevent anyone affecting assignment to a study arm
Performance Bias	<ul style="list-style-type: none"> <input type="checkbox"/> Double-blinding methods employed (i.e., subject and all working with the subject or subject's data) and achieved <input type="checkbox"/> Reasonable intervention and reasonable comparator used (e.g., placebo) <input type="checkbox"/> No bias or difference, except for what is under study, between groups during course of study (e.g., intervention design and execution, care experiences, co-interventions, concomitant medication use, adherence, inappropriate exposure or migration, cross-over threats, protocol deviations, measurement methods, study duration, changes due to time etc.)
Attrition Bias	<ul style="list-style-type: none"> <input type="checkbox"/> Might attrition, including missing data, discontinuations or loss to follow-up, have resulted in distorted outcomes?
Assessment Bias	<ul style="list-style-type: none"> <input type="checkbox"/> Assessors are blinded <input type="checkbox"/> Low likelihood of findings due to chance, false positive and false negative outcomes <input type="checkbox"/> Non-significant findings are reported, but the confidence intervals include clinically meaningful differences <input type="checkbox"/> Intention-to-Treat Analysis (ITT) performed for efficacy (not safety) (all people are analyzed as randomized + reasonable method for imputing missing values which puts the intervention through a challenging trial or reasonable sensitivity analysis) or missing values are very small. <input type="checkbox"/> If time-to-event analysis performed, appropriate, transparent and unbiased. <input type="checkbox"/> Analysis methods are appropriate and use of modeling only with use of reasonable assumptions <input type="checkbox"/> No problems of selective reporting
Usefulness	<ul style="list-style-type: none"> <input type="checkbox"/> Clinically significant area + sufficient benefit size = meaningful clinical benefit (consider efficacy vs effectiveness)



Comment on This Statement

“To demonstrate the efficacy of coronary artery bypass or angioplasty when compared to medical treatment, a reasonable study would be one in which people undergoing bypass are compared to those who were managed with medical treatment.”

Comment on This Statement

“An appropriate study method would be to compare patients receiving surgical treatment with patients treated medically in a different health care facility.”

- Yes or no, and why?”

Comment on This Statement

- Bill Clinton is admitted to Hospital A for a coronary artery bypass graft (CABG) surgery. A newspaper reports—*alors, mon Dieu!*—that our man, Bill, has been admitted to the hospital which has the highest rates of mortality resulting from CABG surgery in his metropolitan area.
- His hospital comes in at 3% as compared to 2% and 1% respectively for the two other primary hospitals in his area—Hospitals B and C.
- The difference is statistically significant.
- Bill has just given you \$1,000,000 (that he does not have) to advise him. But it is his life, hey!
- Should he switch hospitals?

A Goal of Clinical Trials: Causality

- Are outcomes of interest due to the intervention or due to some other factor (bias or chance)?

From Critical Appraisal to Patient Care

Once we have critically appraised the trial in a systematic way and determined that the study is valid, we—

- Evaluate the results for clinical usefulness by evaluating both benefits and harms.
- Patients benefit from effective therapeutic interventions when benefits outweigh harms.
- Bottom Line: We cannot know if an intervention is likely to be effective without critically appraising the studies reporting the results.

The Goal of Clinical Trials: Outcomes

- So this is an overarching question that guides our assessment of results from a *valid* study:
 - What is the probability of benefit or harm?
 - "External validity" is the term used when asking this question "for my patient" or "my population?"

The Goal of Clinical Trials: Differences

- We look at the number of events that occurred in a study to calculate how many people benefited or were harmed out of all of those treated and compare those numbers between study arms.

The Goal of Clinical Trials:

Safety

- Safety information is usually quite limited, and months or years may pass before we know about safety—if ever.
- So patients should generally always be advised that this can be an area of uncertainty.

The Successful Clinical Trial Boiled Down

- The scientific method requires, in most instances, that there are **at least two concurrent groups** for study in clinical trials.
- At the most basic level, one intervention is pitted against another intervention in an experimental study, which includes placebo or "usual care," and the outcomes in the groups are compared.

The Successful Clinical Trial Boiled Down

- You need to isolate what you are comparing—meaning everything else in your experiment must be completely the same except for the interventions under study
- You need a group of people that makes sense to study
- Randomization is a method to make study groups as similar as possible

The Successful Clinical Trial Boiled Down

- You want to eliminate any choice of treatment by patients or caretakers because people choose treatments for different reasons
- The reason for choosing a treatment may result in other kinds of choices that will make the comparison groups different
- Choice is eliminated through random assignment

The Successful Clinical Trial Boiled Down

- Once the study is under way, you want to be sure that the results are not distorted by some other difference between the groups

The Successful Clinical Trial Boiled Down

- You want to successfully measure the resulting outcomes of interest
- You want those outcomes to be of importance to patients
- And you want them to be true (e.g., not distorted by bias, not a result of chance and not be a false negative because your groups were too small to experience outcomes of interest)

The Successful Clinical Trial Boiled Down

- Again, clinical trials are about comparing the resulting differences between the groups.
 - When we talk about "research results," we are talking about the "difference in outcomes between the groups."
- And you want those differences to be of sufficient size to matter—which is context-dependent and a matter of judgment.

The Successful Clinical Trial Boiled Down

- That's it! That's a quality clinical trial in a nutshell.
- How you achieve this has a lot of other pieces to it—such as how blinding can be so important in helping to ensure that the groups are treated the same.

Comment on This Statement

“To demonstrate the efficacy of coronary artery bypass or angioplasty when compared to medical treatment, a reasonable study would be one in which people undergoing bypass are compared to those who were managed with medical treatment.”

Comment on This Statement

“An appropriate study method would be to compare patients receiving surgical treatment with patients treated medically in a different health care facility. Yes or no, and why?”

Watershed Moment 1991 Study Design

circ.ahajournals.org/content/91/1/245.full

DONATE HELP CONTACT AHA HOME

Circulation

American Heart Association

User Name
User Name
Password
***** LOG-IN

Home • Subscriptions • Archives • Feedback • Authors • Help • Circulation Journals Home • AHA Journals Home

Search: Go Advanced Search

Articles

The Cardiac Arrhythmia Suppression Trial Casting Suppression in a Different Light

Craig M. Pratt, MD; Lemuel A. Moyé, MD, PhD

+ Author Affiliations

Correspondence to Craig M. Pratt, MD, Professor of Medicine, Baylor College of Medicine, 6535 Fannin, MS F1001, Houston, TX 77030.

Key Words:

- trials
- Editorials
- arrhythmia

The Cardiac Arrhythmia Suppression Trial (CAST) was instituted in 1986 by the National Heart, Lung, and Blood Institute (NHLBI) after the completion of a 502-patient pilot study (CAPS).^{1 2 3} The initial results of CAST I was published in 1989 and the CAST II results published in 1992.^{2 3} In both trials, antiarrhythmic drugs effectively suppressed asymptomatic ventricular arrhythmias but increased arrhythmic death. Because the suppression hypothesis was refuted, the common

« Prev Article | Next Article »
Table of Contents

This Article

Circulation.
1995;91:245-247
doi: 10.1161/01.CIR.91.1.245

- Extract Free
- Figures Only Free
- » Full Text Free

[-] Classifications

Articles

[-] Services

- E-mail this article to a friend
- Alert me when this article is cited
- Alert me if a correction is posted
- Similar articles in this journal

Current Issue

April 23, 2013

Alert me to new issues of Circulation »

About Circulation

Instructions for Authors

Online Submission/Peer Review

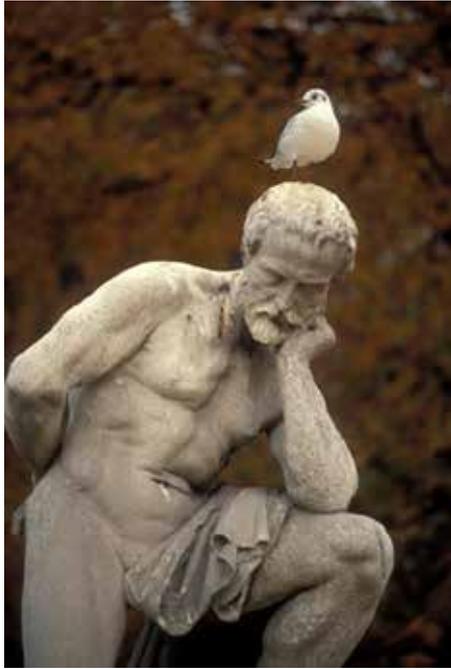
Editorial Office

Follow us on:

Advertiser Information

Heart Disease and Stroke Statistics
2013 UPDATE

“Thinking Hard” May Not Work



CAST Trial

- Twice the risk of death (all cause & coronary artery disease) if post-MI patient has ventricular premature beats (VPCs): Ruberman et al. N Engl J Med 297:750-757, 1977. PMID: 70750
- Flecainide suppresses >90% VPCs—“Highly effective and well tolerated...for long-term treatment of serious arrhythmias”: Meinertz et al. Am J Cardiol 1984;54:91-96. PMID: 6741844
- 50% treatment rate (cardiologists) for patients after MI with asymptomatic or mildly symptomatic VPCs: Morganroth et al. Am J Cardiol 1990;65(1)40-48. PMID: 1688481
- Finally a valid RCT: Encainide and Flecainide increased mortality in patients (after MI) with asymptomatic or mildly symptomatic VPCs—Active drug: cardiac deaths & arrests 8%; Placebo: cardiac deaths and arrests 3%, ($p < 0.001$): Echt et al. CAST. N Engl J Med. 1991;324:781-8. PMID: 1900101

**Bottom Line:
Out of 100
people, 3 died
with placebo,
but 8 died with
new drugs**

Quality Science & Critical Appraisal Skills Make A Difference



- Roughly 58,000 US lives lost in Vietnam war

<http://www.digitalhistory.uh.edu/modules/vietnam/index.cfm>

- We estimate ~63,000 preventable deaths due to encainide/flecainide for premature ventricular contractions (PVCs) after acute myocardial infarction

various sources available on request

Wake-up Call 1998 Study Design

www.nhlbi.nih.gov/news/press-releases/1998/the-hers-study-results-and-ongoing-studies-of-women-and-heart-disease.html

U.S. Department of Health & Human Services NIH National Institutes of Health Contact Us Get Email Alerts Font Size

NIH National Heart, Lung, and Blood Institute

Accessible Search Form Advanced Search

NHLBI Entire Site SEARCH

Public Health Professionals Networks Funding & Research Clinical Trials Training & Careers Researchers Educational Campaigns News & Resources About NHLBI

Home » News & Resources » Press Releases

News & Resources

PRESS RELEASES

NHLBI IN THE NEWS

SPOTLIGHT ON RESEARCH

EVENTS

SPOKESPEOPLE

EMAIL ALERTS

E-NEWSLETTERS

Embargoed for Release: August 18, 1998, 4:00 PM EDT [E-MAIL](#) | [PRINT](#) | [SHARE](#) +

The HERS Study Results and Ongoing Studies of Women and Heart Disease

The results of the first large randomized clinical trial to examine the effect of hormone replacement therapy (HRT) on women with heart disease appear in the August 19 issue of the *Journal of the American Medical Association (JAMA)*.

The Heart and Estrogen/Progestin Replacement Study (HERS) found that the use of estrogen plus progestin in postmenopausal women with heart disease did not prevent further heart attacks or death from coronary heart disease (CHD). This occurred despite the positive effect of treatment on lipoproteins: LDL (bad) cholesterol was reduced by 11 percent and HDL (good) cholesterol was increased by 10 percent.

The hormone replacement regimen also increased the risk of clots in the veins (deep vein thrombosis) and lungs (pulmonary embolism).

HERS involved 2,763 postmenopausal women, average age 67, who were treated for approximately 4 years. Women were randomly assigned to an estrogen/progestin combination or to a placebo. The study was funded by Wyeth-Ayerst Laboratories.

The results of HERS are surprising in light of previous observational studies, which found lower rates of CHD in women who take postmenopausal estrogen. Although observational studies cannot provide conclusive answers about a particular treatment since the groups studied can differ in significant ways, they provide important clues and the unexpected null finding in HERS is worthy of analysis. The

For the Media

NHLBI Communications Office
nhlbi_news@nhlbi.nih.gov
301-496-4236
Ask for press officer on duty

Related Health Topics

[Coronary Heart Disease](#)

[Deep Vein Thrombosis](#)

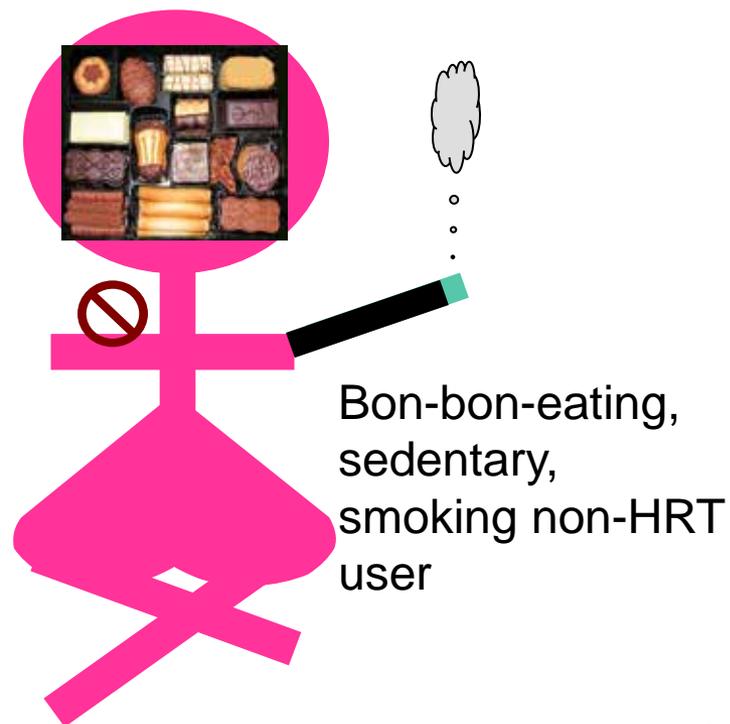
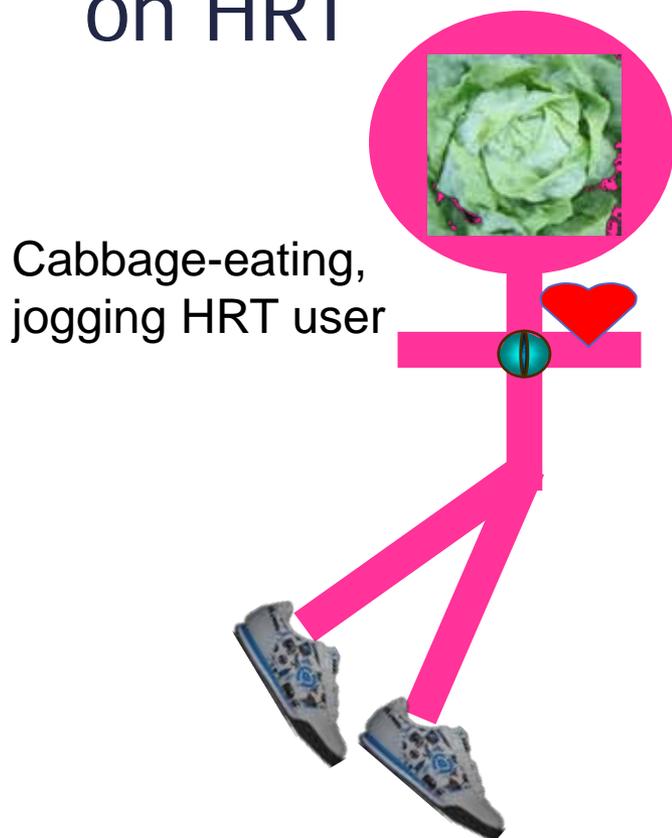
[Heart Attack](#)

[Heart Disease in Women](#)

[Pulmonary Embolism](#)

But If We Looked More Closely At The Women, Maybe We'd See Something Like This...

- Women choosing to take HRT may be more health-conscious than women not on HRT



Salpeter

NCBI Resources ▾ How To ▾

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed ▾ 15209595[uid]

RSS Save search Limits Advanced

Display Settings: ▾ Abstract Send to: ▾

J Gen Intern Med. 2004 Jul;19(7):791-804.

Mortality associated with hormone replacement therapy in younger and older women: a meta-analysis.

Salpeter SR, Walsh JM, Greyber E, Ormiston TM, Salpeter EE.
Department of Medicine, Santa Clara Valley Medical Center, San Jose, CA 95128, USA. salpeter@stanford.edu

Abstract

OBJECTIVE: To assess mortality associated with hormone replacement in younger and older postmenopausal women.

DESIGN: A comprehensive search of MEDLINE, CINAHL, and EMBASE databases was performed to identify randomized controlled trials of hormone replacement therapy from 1966 to September 2002. The search was augmented by scanning selected journals through April 2003 and references of identified articles. Randomized trials of greater than 6 months' duration were included if they compared hormone replacement with placebo or no treatment, and reported at least 1 death.

MEASUREMENTS: Outcomes measured were total deaths and deaths due to cardiovascular disease, cancer, or other causes. Odds ratios (OR) for total and cause-specific mortality were reported separately for trials with mean age of participants less than and greater than 60 years at baseline.

MAIN RESULTS: Pooled data from 30 trials with 26,708 participants showed that the OR for total mortality associated with hormone replacement was 0.98 (95% confidence interval [CI], 0.87 to 1.12). Hormone replacement reduced mortality in the younger age group (OR, 0.61; CI, 0.39 to 0.95), but not in the older age group (OR, 1.03; CI, 0.90 to 1.18). For all ages combined, treatment did not significantly affect the risk for cardiovascular or cancer mortality, but reduced mortality from other causes (OR, 0.67; CI, 0.51 to 0.88).

CONCLUSIONS: Hormone replacement therapy reduced total mortality in trials with mean age of participants under 60 years. No change in mortality was seen in trials with mean age over 60 years.



Lack of Effective Training

- The majority of healthcare professionals lack critical appraisal skills—
 - This includes academicians, “experts,” researchers, editors, peer-reviewers, guideline developers, colleagues...
- What can we be likely to count on?
 - In one survey, of 60,352 studies, 7% passed criteria of high quality methods and clinical relevancy (PMID 15350200), and fewer than 5% passed a validity screening for an evidence-based journal (PMID 17213115)
 - Informal estimates of ours and highly trained groups are similar

§ ~70%
Healthcare
Professionals
Fail Pre-test

§ ~90% Articles
Fail Appraisal





Rofecoxib

The New England Journal of Medicine

COMPARISON OF UPPER GASTROINTESTINAL TOXICITY OF ROFECOXIB AND NAPROXEN IN PATIENTS WITH RHEUMATOID ARTHRITIS

CLAIRE BOMBARDIER, M.D., LOREN LAINE, M.D., ALISE REICIN, M.D., DEBORAH SHAPIRO, DR.P.H.,
RUBEN BURGOS-VARGAS, M.D., BARRY DAVIS, M.D., PH.D., RICHARD DAY, M.D., MARCOS BOSI FERRAZ, M.D., PH.D.,
CHRISTOPHER J. HAWKEY, M.D., MARC C. HOCHBERG, M.D., TORE K. KVIEN, M.D.,
AND THOMAS J. SCHNITZER, M.D., PH.D., FOR THE VIGOR STUDY GROUP

PMID 11087881



The Abstract

RESULTS: Rofecoxib and naproxen had similar efficacy against rheumatoid arthritis. During a median follow-up of 9.0 months, 2.1 confirmed gastrointestinal events per 100 patient-years occurred with rofecoxib, as compared with 4.5 per 100 patient-years with naproxen (relative risk, 0.5; 95 percent confidence interval, 0.3 to 0.6; $P < 0.001$). The respective rates of complicated confirmed events (perforation, obstruction, and severe upper gastrointestinal bleeding) were 0.6 per 100 patient-years and 1.4 per 100 patient-years (relative risk, 0.4; 95 percent confidence interval, 0.2 to 0.8; $P = 0.005$). The incidence of myocardial infarction was lower among patients in the naproxen group than among those in the rofecoxib group (0.1 percent vs. 0.4 percent; relative risk, 0.2; 95 percent confidence interval, 0.1 to 0.7); the overall mortality rate and the rate of death from cardiovascular causes were similar in the two groups.

CONCLUSIONS: In patients with rheumatoid arthritis, treatment with rofecoxib, a selective inhibitor of cyclooxygenase-2, is associated with significantly fewer clinically important upper gastrointestinal events than treatment with naproxen, a nonselective inhibitor.

The Abstract

RESULTS: Rofecoxib and naproxen had similar efficacy against rheumatoid arthritis. During a median follow-up of 9.0 months, 2.1 confirmed gastrointestinal events per 100 patient-years occurred with rofecoxib, as compared with 4.5 per 100 patient-years with naproxen (relative risk, 0.5; 95 percent confidence interval, 0.3 to 0.6; $P < 0.001$). The respective rates of complicated confirmed events (perforation, obstruction, and severe upper gastrointestinal bleeding) were 0.6 per 100 patient-years and 1.4 per 100 patient-years (relative risk, 0.4; 95 percent confidence interval, 0.2 to 0.8; $P = 0.005$). The incidence of myocardial infarction was lower among patients in the naproxen group than among those in the rofecoxib group (0.1 percent vs. 0.4 percent; relative risk, 0.2; 95 percent confidence interval, 0.1 to 0.7); the overall mortality rate and the rate of death from cardiovascular causes were similar in the two groups.

CONCLUSIONS: In patients with rheumatoid arthritis, treatment with rofecoxib, a selective inhibitor of cyclooxygenase-2, is associated with significantly fewer clinically important upper gastrointestinal events than treatment with naproxen, a nonselective inhibitor.

Results from FDA Review

FDA Cardiovascular Safety Review, NDA 21-042, S-007

www.fda.gov/ohrms/dockets/ac/01/briefing/3677b2_06_cardio.doc

9 Month Study Duration

	Complicated Upper GI Events	Myocardial Infarction	All Thrombotic Events
Rofecoxib	0.6 per 100 pt-yr N=4047	0.4 per 100 pt-yr	1.67 per 100 pt-yr
Naproxen	1.4 per 100 pt-yr N=4029 RR=0.4 95%CI (0.2 to 0.8) P=0.005	0.1 per 100 pt-yr Relative Risk=0.2 95%CI (0.1 to 0.7)	0.7 per 100 pt-yr Relative Risk=0.42 95%CI (0.25 to 0.72)
ARR (100 pt-yr calculation)	0.8%	(0.3%)	(0.97%)
NNT (NNH)	125	(333)	(103)

Quality Science & Critical Appraisal Skills Make A Difference



- Roughly 58,000 US lives lost in Vietnam war
<http://www.digitalhistory.uh.edu/modules/vietnam/index.cfm>
- It is estimated that rofecoxib “may have contributed to 27,785 heart attacks and sudden cardiac deaths between 1999 and 2003.”

http://www.consumeraffairs.com/news04/vioxx_estimates.html#ixzz0cqOwLu3m

Vol. 281 No. 12, March 24, 1999

Brief Report

Accuracy of Data in Abstracts of Published Research Articles

Roy M. Pitkin, MD; Mary Ann Branagan; Leon F. Burmeister, PhD

- **Main Outcome Measure** Abstracts were considered deficient if they contained data that were either inconsistent with corresponding data in the article's body (including tables and figures) or not found in the body at all.
- **Results** The proportion of deficient abstracts varied widely (18%-68%) and to a statistically significant degree ($P < .001$) among the 6 journals studied: **JAMA, NEJM, BMJ, Annals, Lancet, CMAJ**.
- **Conclusions** Data in the abstract that are inconsistent with or absent from the article's body are common, even in large-circulation general medical journals.



Censoring Case Study

Steven E. Nissen, M.D.
Cleveland Clinic Foundation
Cleveland, OH 44195

TO THE EDITOR:

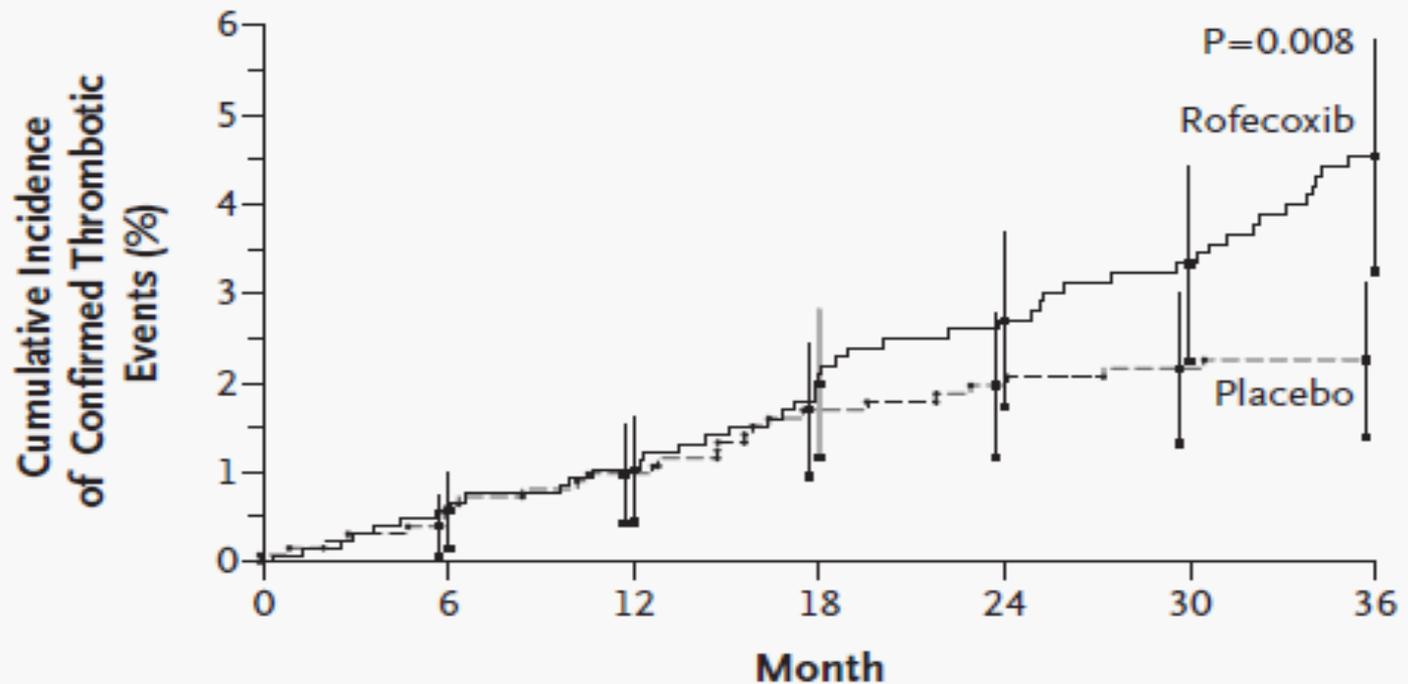
N ENGL J MED 355;2 WWW.NEJM.ORG JULY 13, 2006

In the original article, the APPROVe investigators reported event rates using an unusual censoring rule in which events were excluded if they occurred more than 14 days after the study drug was stopped. All data in the new report are assessed by a conventional intention-to-treat analysis.

Original Report

Bresalier RS, Sandler RS, Quan H, et al. Cardiovascular

events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med* 2005;352:1092-102. PMID: 15713943



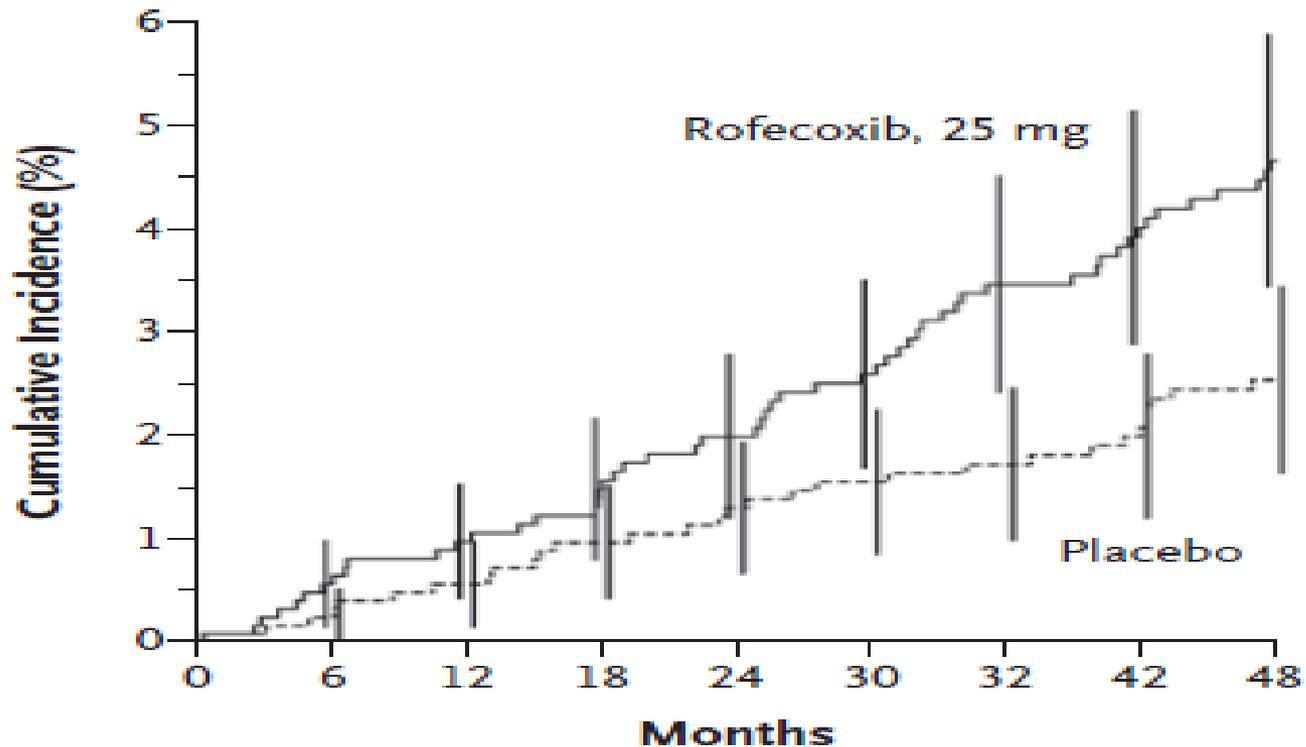
No. at Risk

Rofecoxib	1287	1129	1057	989	938	896	727
Placebo	1299	1195	1156	1079	1042	1001	835

Figure 2. Kaplan–Meier Estimates of the Cumulative Incidence of Confirmed Serious Thrombotic Events.

Vertical lines indicate 95 percent confidence intervals.

Nissen's Recalculation Curve With ITT Calculation



No. at Risk

Rofecoxib	1287	1220	1188	1158	1140	1125	1102	1042	1002
Placebo	1300	1249	1228	1196	1181	1165	1140	1079	1036

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.

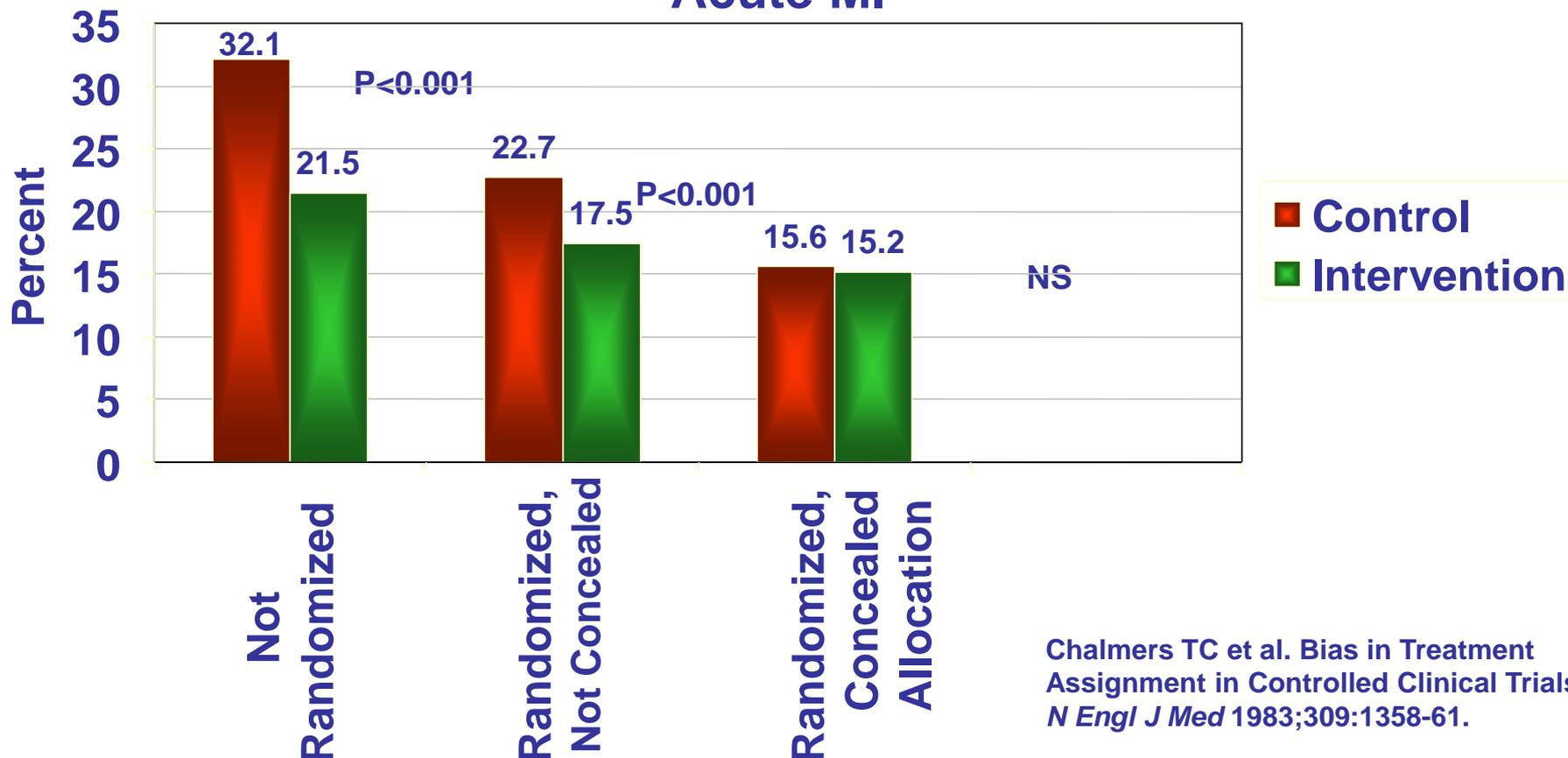
I bars represent 95 percent confidence intervals.

Does Critical Appraisal Make A Difference? Yes!

- Low quality clinical trials compared to high quality are **likely to overestimate benefit** by up to a relative 30-50% or more
- This would result in interventions tending to more effective than they are—or effective when they are not
- The upcoming numbers are **not for application**—they are for awareness

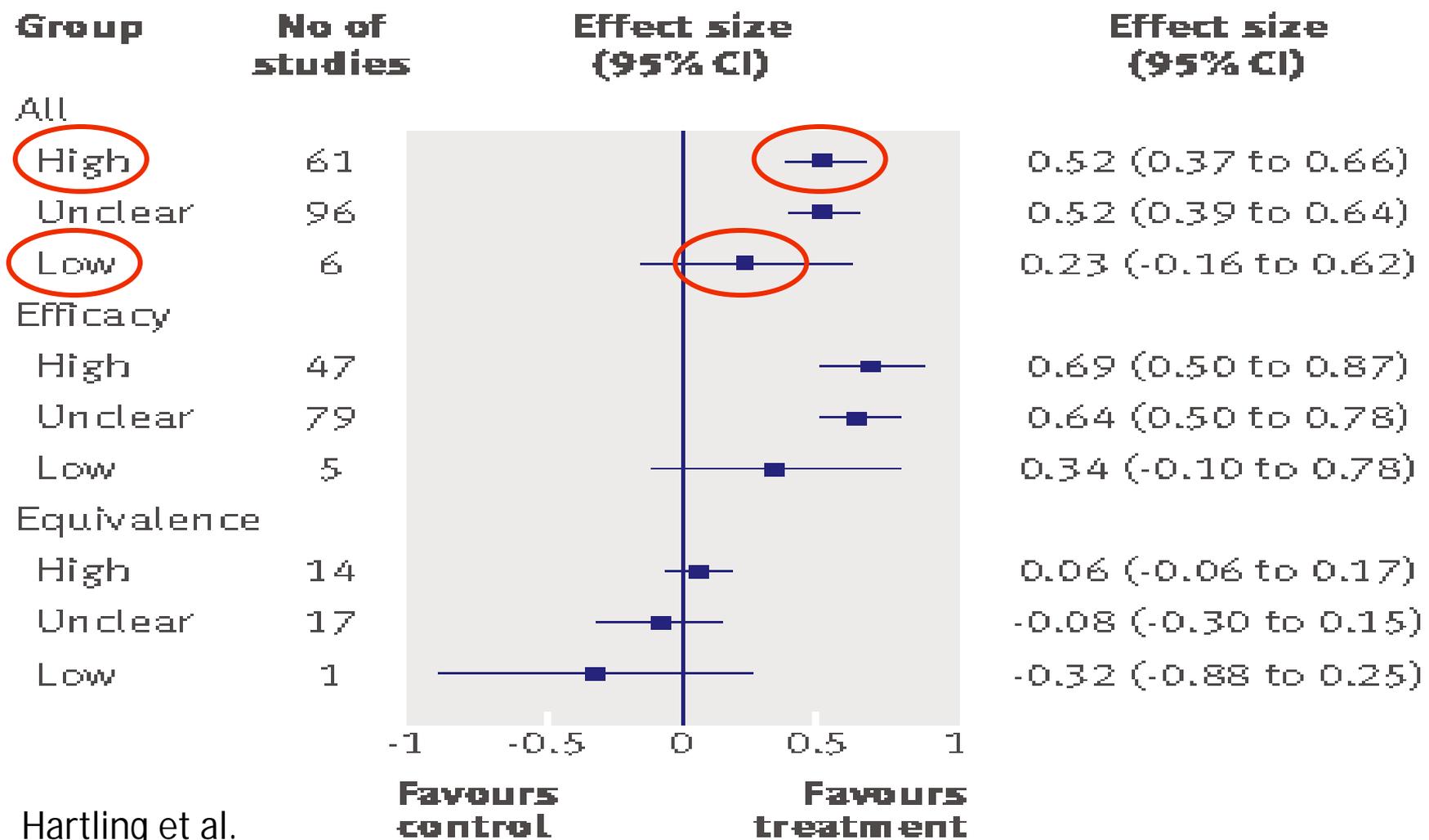
Randomization, Concealment of Allocation & Impact on Results

Case Fatality Rates Acute MI





Effect Size in Systematic Reviews: Comparison of Reviews That Included High Risk of Bias Studies vs Reviews Including Low Risk of Bias Trials



Hartling et al.
PMID: 19841007

Stage of Trial & Area of Concern	Estimated Range of Relative Distortion of Study Results	References
Stage I: Establishing Comparable Groups (Selection Bias)		
<ul style="list-style-type: none"> Inadequate Generation of Sequence 	17% to 75%	4. Juni 01 PMID: 11440947 5. Kjaergard 01 PMID: 11730399 6. van Tulder 09 PMID: 19770609
<ul style="list-style-type: none"> Inadequate Concealment of Allocation of the Randomization Sequence 	14% to 73%	1. Schulz 95 PMID: 7823387 2. Moher 98 PMID: 9746022 4. Juni 01, PMID: 11440947 5. Kjaergard 01 PMID: 11730399 7. Chalmers 83 PMID: 6633598
Stage II: Intervention and Context (Performance Bias)		
<ul style="list-style-type: none"> Inadequate Double Blinding 	4% to 72%	1. Schulz 95 PMID: 7823387 2. Moher 98 PMID: 9746022 4. Juni 01, PMID: 11440947 5. Kjaergard 01 PMID: 11730399
Stage III: Loss of Data (Attrition Bias)		
<ul style="list-style-type: none"> Loss of Data (Up to 38%) 	2% to 35%	6. van Tulder 09 PMID: 19770609 8. Tierney 04 PMID: 15561753 9. Nuesch PMID: 19736281 24. Canadian Orthopaedic Trauma Society 07 PMID: 17200303
Stage IV: Assessment (Assessment Bias)		
<ul style="list-style-type: none"> Inadequate Blinding of Assessors 	35% to 69%	11. Poolman 07 PMID: 17332104 25. Juni 99 PMID: 10493204
<ul style="list-style-type: none"> Completer Analysis 	56% with 44% early withdrawal	26. Shih 02 PMID: 11985778
<ul style="list-style-type: none"> Assessment Models For Missing Data 	With loss of 20% risk of type I error* is approximately 10%; With loss of 40% risk of type I is approximately 50% *Type 1 - or alpha error - A difference is reported, but there is no difference.	12. Lachin 00 PMID: 11018568

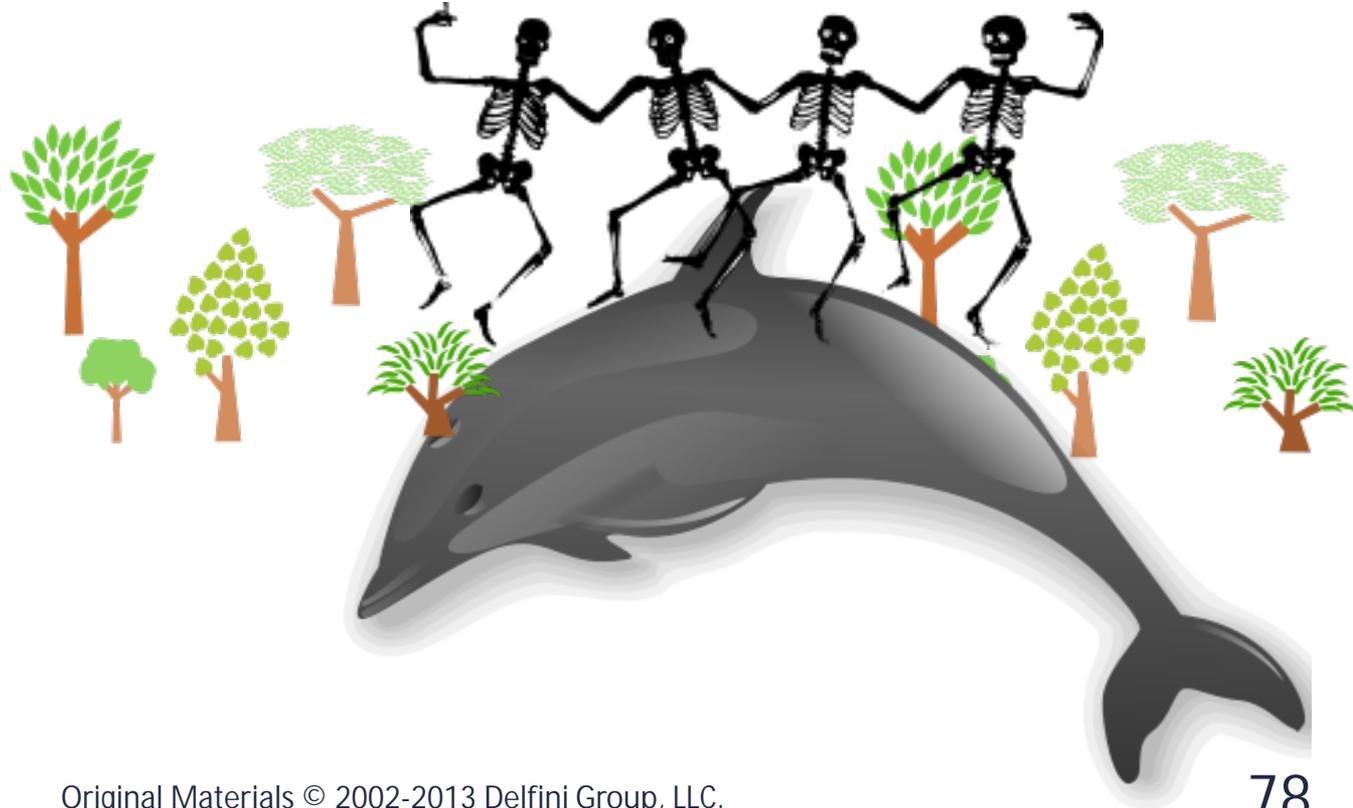


The CAPPP Trial



Research Is Hard to Do!

- Design
- Methods
- Execution
- Study Performance Outcomes
- Reporting
- Topic Area



1982 Volvo Award in Clinical Science

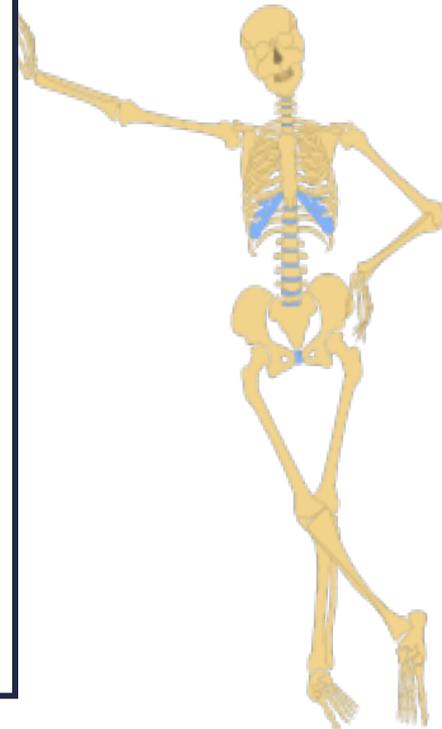
Lumbar Disc Herniation A Controlled, Prospective Study with Ten Years of Observation

HENRIK WEBER, MD

Two hundred eighty patients with herniated lumbar discs, verified by radiculography, were divided into three groups. One group, which mainly will be dealt with in this paper, consisted of 126 patients with uncertain indication for surgical treatment, who had their therapy decided by randomization which permitted comparison between the results of surgical and conservative treatment. Another group comprising 67 patients had symptoms and signs that beyond doubt, required surgical therapy. The third group of 87 patients was treated conservatively because there was no indication for operative intervention. Follow-up examinations in the first group were performed after one, four, and ten years.

The controlled trial showed a statistically significant better result in the surgically treated group at the one-year follow-up examination. After four years the operated patients still showed better results, but the difference was no longer statistically significant. Only minor changes took place during the last six years of observation. [Key words: lumbar disc herniation, surgery, prognostic factors, epidemiology]

NOT FOR REPRODUCTION WITHOUT THE
PERMISSION OF THE PUBLISHERS



Washington's Prescription Drug Program

Siri Childs,
Pharm D

Pharmacy
Administrator,
(Formerly)
Health and
Recovery
Services
Administration,
Washington
State

AHRQ Annual
Meeting

Bethesda MD

Sept 20, 2011



**Using systematic reviews to make
policy decisions
in the effort to contain prescription
drug expenditures**



Key Factors to Success

- Enabling legislation:
 - State Senate Bill 6088 was passed by the 2003 Legislature to mandate an evidence-based Preferred Drug List for state agencies purchasing drugs for Washington residents and the formation of a P&T Committee
- Credible research: OHSU EPC and the Drug Effectiveness Project
- Continuing education: Delfini Group LLC conduct annual training to state agency decision makers and P&T Committee members



Successes

- **Preferred drugs:** “Cost avoidance” estimated at \$46 million for the 2008 state fiscal year, based on market shift to the preferred drug(s) in 24 drug classes; average of 90% compliance to preferred drug
- **Supplemental rebates:** Estimated savings from supplemental rebates grew each year with new drug class implementation and intensified competition:
 - 2008 state fiscal year: \$6.6 Million
- **Pharmacy spend as a percent of Medicaid spend** dropped 3.4% from FY 2005 to FY 2008
- **Drug Utilization Review:** “Cost avoidance” estimated at \$24 million per year from targeted drug initiatives in the prior authorization program

But Let's Talk Cost for a Second

The cost-effectiveness of cyclooxygenase-2 selective inhibitors in the management of chronic arthritis. Spiegel BM, Targownik L, Dulai GS, Gralnek IM. PMID 12755551



Critical Appraisal Matters

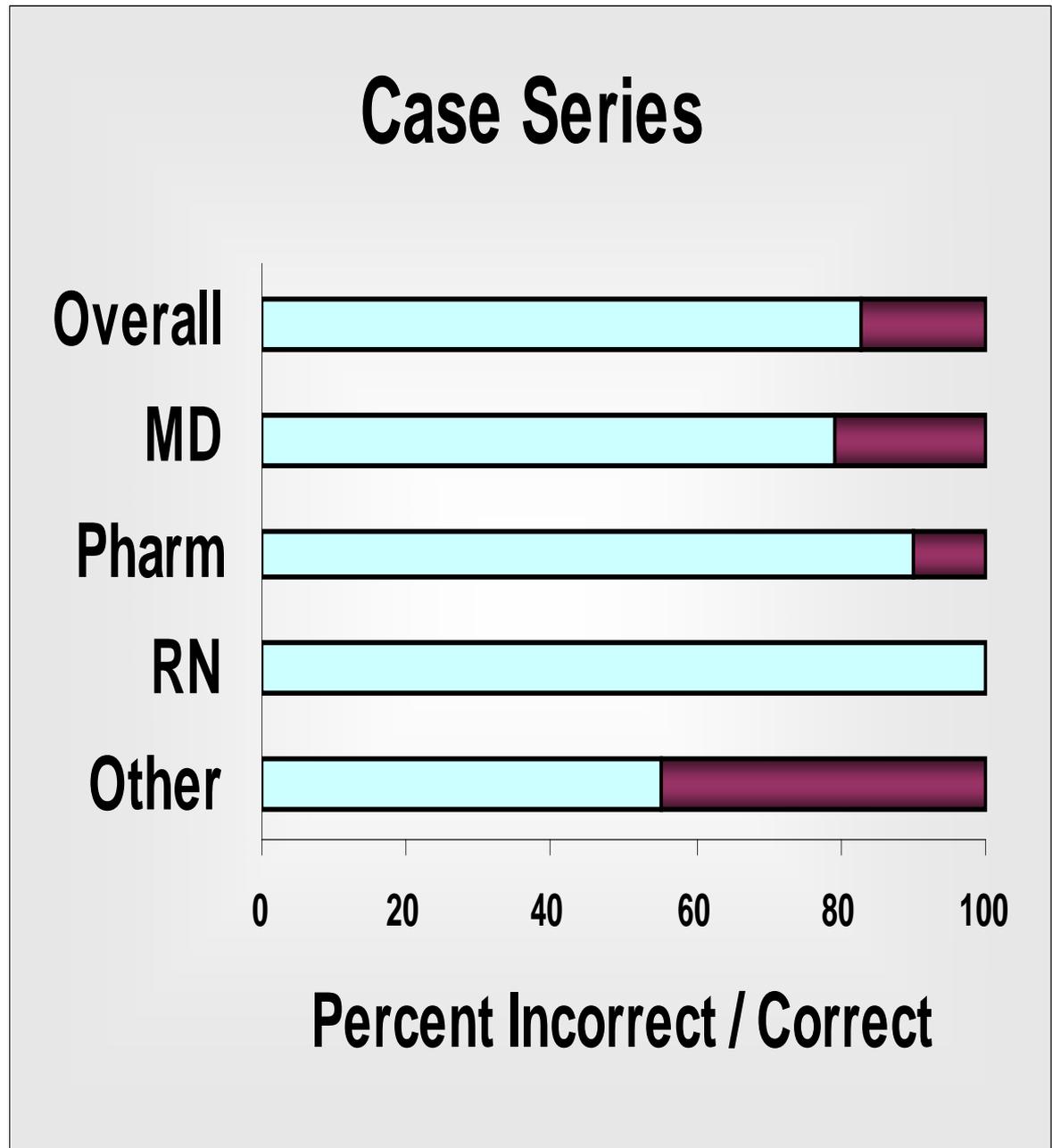


Nuevo Magico

- 240 patients presenting with a number of symptoms of Condition X treated with Nuevo-Magico
- People may be ill for weeks. The disease is highly contagious and can lead to significant complications.
- Side effects of Nuevo-Magico are documented in numerous well-done studies to be very rare.
- Of the 240 patients treated, 232 patients are asymptomatic within 3-5 days of coming into the doctor's office.
- Will most doctors prescribe this drug based on this information?

Yes, they will!

Should they?





How Medicine is Practiced: *“It Works!”*





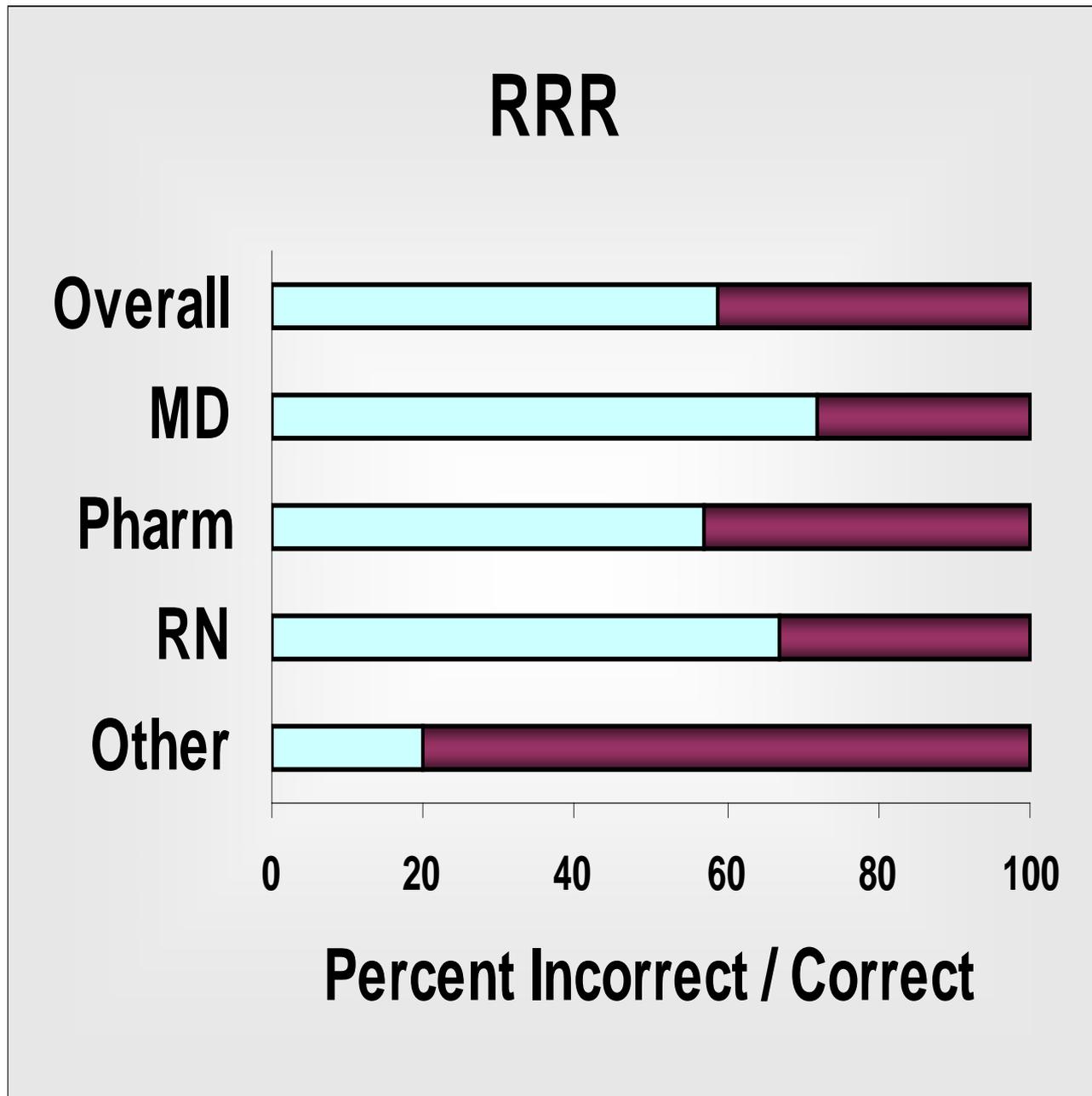
Pre-test Question

8. A well done study reports a statistically significant relative risk reduction of 60% for patients in the intervention group. Is this a result that may persuade you to use the intervention with your patients?

Yes: _____ (Reason)

No: _____ (Reason)

Simple
Math &
A Simple
Concept
We Can All
Understand
(As
Consumers)



Informed Consent is Not Possible Without Critically Appraised Information

Critical appraisal is the best method known for assessing if a research study is likely to be reliable—and the purpose of reliable research is to help improve the ability to predict what will happen to a patient.

- Only after the evidence has been critically appraised for validity can we conclude that beneficial outcomes reported in clinical trials were not caused or distorted by bias or chance.
- Patients deserve to know the benefits and risks of interventions and the likelihood of experiencing various outcomes. Without critical appraisal of the evidence, this is not possible.
- Patient preferences are very likely to differ if patients are provided with information about the quality of the evidence and the amount of benefit and risk.

Understanding Research Bias

A Quick Tour

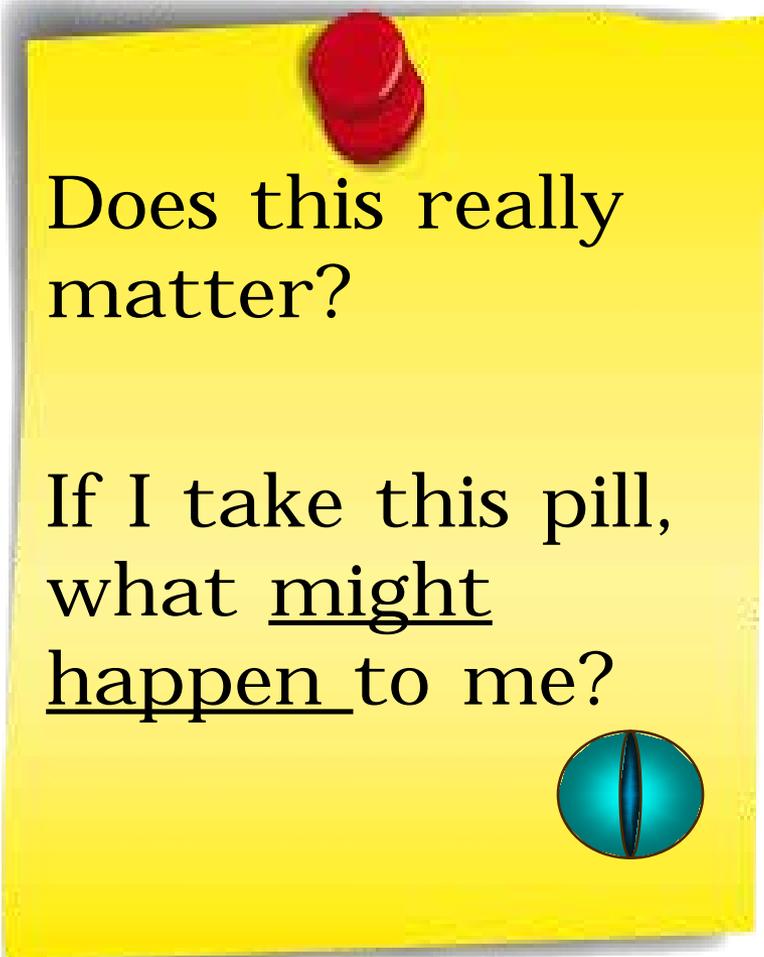


Critical Appraisal Sampler

- Caution: not complete



What's the Goal of Medical Research for Interventions?



Does this really matter?

If I take this pill, what might happen to me?

Q: Why do we want good science for medical decision-making?

A1: Determine cause & effect

A2: Assess likelihood of effects (probability)

Is it True? Is it Useful?

Are the reported results—

1. Reliable and
2. Clinically useful?

- Reliability = assess validity (“closeness to truth”) by assessing bias, confounding and chance
- Clinical usefulness = assess meaningful clinical benefit

If True, Are The Results Useful?

Meaningful clinical benefit means benefit to patients in 5 areas—

1. Morbidity
2. Mortality
3. Symptom relief
4. Emotional, mental or physical functioning
5. Health-related quality of life

+

Size of the outcomes



You Need At Least Two (Concurrent) Groups to Compare for Differences



And the groups need to be the same...

