An Introduction to Clinical Trials: Design Issues

Edgar R Miller III PhD, MD
Welch Center for Prevention, Epidemiology and Clinical Research
Johns Hopkins University
School of Medicine and Bloomberg School of Public Health
Type of Studies

• Non-experimental (Observational)
  – Case report
  – Case series
  – Cross-sectional (survey)
  – Case-control
  – Prospective, observational (cohort)

• Experimental
  – Randomized, clinical trial (RCT)
Study designs

• Observational studies:
  – **Observe** both exposures and outcomes

• Experimental studies (clinical trials)
  – **Assign** exposures
  – **Observe** outcomes
Epidemiology in a box: The 2x2 table

- The EXPOSURE (E)
  - Example: obesity
- The OUTCOME (D)
  - Example: Hypertension
- Applicable to most study designs
Cross-Sectional Study

Begin with: Defined Population

Then: Gather Data on Exposure and Disease

- Exposed, with Disease
- Exposed, No Disease
- Not Exposed, with Disease
- Not Exposed, No Disease
Case-Control Study

Begin with:

- Disease

Then:

- Exposed
- Not Exposed

- No Disease
- Exposed
- Not Exposed
Prospective Observational Studies

- Defined Population
- NON-RANDOMIZED
  - Exposed
    - Disease
    - No Disease
  - Non-Exposed
    - Disease
    - No Disease
Limitations of Observational, Non-Experimental Studies

- **Selection bias**: bias in selection of participants
- **Information bias**: bias in ascertainment of exposure or outcome status
- **Confounding** (possibly a bias as well): The association is real, but the inference is wrong.
Confounding

• Confounding describes a relationship between TWO exposures and ONE outcome.

Guilt by association: In this example, smoking is a confounder in the relationship between coffee and lung cancer.

\[
\text{Smoking} \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{Coffee drinking} \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad ? \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{Lung cancer}
\]
Confounding by Indication

- Those who receive a therapy are placed on the therapy because it is clinically “indicated”, and are therefore more or less likely to develop the outcome on that basis alone.

- This is one of the most important limitations of evaluating treatments using cohort studies.
Randomized Clinical Trial

Target Population

Study Population

RANDOMIZED

Standard Treatment

Disease

New Treatment

Disease
Advantages of Clinical Trials

• Often provides the strongest evidence in support of cause-effect relationships

• Basis for clinical and public health policy

• Minimize/eliminate bias and confounding
# Comparison of Study Designs

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Type of Study Design</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cross-Sectional</td>
</tr>
<tr>
<td>Estimate Prevalence</td>
<td>A</td>
</tr>
<tr>
<td>Estimate Incidence</td>
<td>-</td>
</tr>
<tr>
<td>Prove Causality</td>
<td>C</td>
</tr>
<tr>
<td>Generalizability</td>
<td>A</td>
</tr>
<tr>
<td>Feasability</td>
<td>A</td>
</tr>
</tbody>
</table>
Core Elements of a Clinical Trial

- Research Question
- Hypotheses
- Core Design
- Study Participants
- Recruitment
- Allocation
- Masking (Blinding)
- Treatment Groups

- Data
- Analytical Issues
- Interpretation of Results
The Research Question

• Critical in the design of a trial

• Types of questions:
  – Assessing efficacy of an intervention
  – Assessing the effectiveness of an intervention
Types of Hypotheses

• Comparative Trial (a.k.a. Superiority Trial)
  – Objective: to demonstrate that a new therapy (n) is superior to standard therapy (s) in terms of incident outcome (I)

  \[ H_0: I_n = I_s \]

  \[ H_A: I_n < I_s \] (one tailed) or \[ H_A: I_n \neq I_s \] (two tailed) at some minimally detectable \( \Delta \) judged to have clinical significance
Types of Hypotheses

• Equivalence (non-inferiority trial)
  – Objective: to demonstrate that a new therapy (n) is no worse than standard therapy (s) in terms of incident outcome (I)
    
    \[ H_0: I_n > I_s \]
    
    \[ H_A: I_n = I_s \text{ at some } \Delta, \text{ the maximum tolerable difference considered to be clinically acceptable} \]
Basic Types of Design

Parallel

Cross-Over
Parallel Study Design (PREMIEER)

Randomization

ADVICE ONLY

EST

EST + DASH

↑↑↑↑↑↑

↑ = Data Visit

Primary Outcomes (6 months)

End of Intervention (18 months)
Conlin et al., Am J Hypertens, 2002
Cross-Over Study Design (OmniHeart)

Randomization to 1 of 6 sequences

Data:

- **Screening/Baseline**
- **Run-In** 6 days
- **Period 1** 6 weeks
- **Period 2** 6 weeks
- **Period 3** 6 weeks

Washout Period 2–4 wk

Washout Period 2-4 wk

Participants Ate Their Own Food

Participants Ate Study Food
## Blood Pressure Results (mmHg)

<table>
<thead>
<tr>
<th>Systolic BP</th>
<th>Baseline</th>
<th>CARB</th>
<th>PROT</th>
<th>UNSAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>131.2</td>
<td>-8.2</td>
<td>-9.5</td>
<td>-9.3</td>
</tr>
<tr>
<td>HTN Only</td>
<td>146.5</td>
<td>-12.9</td>
<td>-16.1</td>
<td>-15.8</td>
</tr>
<tr>
<td>PreHTN Only</td>
<td>127.5</td>
<td>-7.0</td>
<td>-8.0</td>
<td>-7.7</td>
</tr>
<tr>
<td>Diastolic BP</td>
<td>77.0</td>
<td>-4.1</td>
<td>-5.2</td>
<td>-4.8</td>
</tr>
</tbody>
</table>

Appel et al. 2005
Mixed Study Design (DASH-Sodium)

Randomization to Diet

Usual Diet

DASH Diet

Lower Sodium
Intermediate Sodium
Higher Sodium

Randomized Sequence

Run-in (11-14 days) Intervention (three 30-d periods in random order)
Effect of Increased Sodium Intake on Systolic Blood Pressure in Two Diets: Results of the DASH-Sodium Trial*

*Sacks et al, 2001
Factorial Design

- Type of trial in which individuals are randomized to two or more therapies (example: Physician’s Health Study: tested aspirin (ASA) and β-carotene)

<table>
<thead>
<tr>
<th></th>
<th>No β-carotene</th>
<th>β-carotene only</th>
<th>β-carotene only</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ASA</td>
<td>Neither</td>
<td>β-carotene only</td>
<td>10,000</td>
</tr>
<tr>
<td></td>
<td>ASA only</td>
<td>Both</td>
<td>10,000</td>
</tr>
</tbody>
</table>

| 10,000         | 10,000        | 20,000          |
The African American Study of Kidney Disease and Hypertension (AASK)
Among African-Americans with early evidence of hypertension-related kidney disease:

- Does aggressive blood pressure control to a target blood pressure below current recommendations retard the progression of kidney disease?

- Do specific classes of anti-hypertensive medications retard the progression of kidney disease?
Design of AASK

• Randomized, active controlled trial with a 2 x 3 factorial design
• Participants: 1,094 African-Americans with hypertension-related renal insufficiency
• Planned follow-up of 2.5 to 5 years
## Treatment Assignments
(2:2:1 ratio of drug assignment)
3 X 2 Factorial Design

<table>
<thead>
<tr>
<th></th>
<th>Metoprolol*</th>
<th>Ramipril</th>
<th>Amlodipine</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAP &lt;92</td>
<td>20%</td>
<td>20%</td>
<td>10%</td>
</tr>
<tr>
<td>MAP 102-107</td>
<td>20%</td>
<td>20%</td>
<td>10%</td>
</tr>
</tbody>
</table>

| N             | 441         | 436      | 217        |

MAP = Mean Arterial Pressure; * = referent group
Mean Arterial Pressure During Follow-up

- Lower BP Goal (Achieved: 128/78)
- Usual BP Goal (Achieved: 141/85)
**Composite Clinical Outcome**

**Declining GFR Event, ESRD or Death**

- **Lower BP (Achieved: 128/78)**
- **Usual BP (Achieved: 141/85)**

Low vs. Usual:
RR=2%, (p=0.85)

Follow-Up Time (Months):

% with Events

RR=Risk Reduction, adjusted for baseline covariates
Main Clinical Composite Outcome
Declining GFR Event, ESRD, or Death

Metoprolol vs. Amlodipine:
RR = 20%, p = 0.17

Ramipril vs. Amlodipine:
RR = 38%, p = 0.004

Ramipril vs. Metoprolol:
RR = 22%, p = 0.042

RR = Risk Reduction, Adjusting for Baseline Covariates
Cumulative % patients reaching GFR event, ESRD, or Death*

* to 36 months for Amiodipine, 48 months for Rempiril and Moloprolol
Study Participants

- Target Population
- Accessible Population
- Study Samples
Study Participants: Example

- Target Population -> Healthy Elderly
- Accessible Population -> Retired Teachers
- Study Sample -> Volunteer Teachers who respond to mass mailing
Study Participants

• Ideal ‘Accessible’ Population
  – high risk for disease
  – candidates for treatment
  – representative of target population
  – feasibility considerations
    • recruitment
    • follow-up
    • high quality data
Enrollment Criteria

• Inclusion Criteria
  – characteristics of accessible population

• Exclusion Criteria
  – considerations related to:
    • adherence to therapy
    • follow-up
    • safety
    • ethics
Common Recruitment Strategies

• General mailings
  – Licensed drivers
  – Voters
  – Employee paychecks
• Targeted mailings
  – HMO enrollees
  – AARP members
• Mass media
  – Radio
  – TV ads
  – Newspapers
  – Posters/flyers
• Screenings
  – Worksite
  – Community
• Physician Referral
• Medical Record Review
• Internet / WWW
  – Clinical trial registries
  – Banner ads
  – Social networks
Comments on Recruitment

- Recruitment begins with design
- Response rate is always lower than expected
- Required resources are more than expected
- Dedicated personnel are necessary
More Comments on Recruitment

• Recruitment period is often longer than expected
• Implement several strategies to identify best source
• Prepare back-up strategies
• Monitor recruitment
  – Early
  – Often
  – Locally
Recruitment “Funnel”
(Example: VITAL Pilot Study)

4,774 Mailed Invitations
  ↓ 43%
2,034 Questionnaires Returned
  ↓ 38%
765 Interested After Initial Mailing
  ↓ 41%
323 Randomizable after Second Mailing (7% cumulative)
  ↓
297 Randomized
Allocation

• Random
  – stratified
  – blocked

• Non-Random
  – haphazard
  – systematic
Why randomize?

- Two critical reasons:
  - to eliminate selection **BIAS**
  - to reduce/avoid **CONFOUNDING** from known and, more importantly, unknown confounders
Masking (Blinding)

- **Single Blind**
  - Observers (persons who collect outcome variable) do not know treatment assignment
- **Double Blind**
  - Study participants AND observers do not know treatment assignments
- **Triple Blind**
  - Data interpreters, study participants, and observers do not know treatment assignments
# Masking (Blinding)

<table>
<thead>
<tr>
<th></th>
<th>Single Masked</th>
<th>Double Masked</th>
<th>Triple Masked</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome Assessor(s)</strong></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Participant</strong></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Data Interpreter</strong></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Selection of Groups

• Active Treatment Group

• Comparison Group
  – Placebo (no active therapy)
  – Usual care (referral back to personal MD)
  – Active control group (provision of standard therapy)
Problems with selecting active treatment group

• Many Candidate treatments
  – observation studies, animal models, or theoretically based

• Strong evidence rarely exists to guide selection of intervention

• Dose/intensity are uncertain
Comparison Group

• Placebo – used in setting of:
  – No standard therapy OR
  – Standard therapy but risk of not providing it is minimal

• Usual care OR active control – common
Problems with standard of care approach

- Efficacy of ‘Usual care’ often not tested

- Variations in standard of care are common:
  - across providers
  - between experts and providers
  - secular trends occur
Data

• Baseline data
  – Determine eligibility
  – Describe study participants
  – Define subgroups
  – Address confounding

• Measures of Adherence

• Outcome Variables
Outcome Variables

• Principal outcome
  – most important variable after randomization code
  – specified in hypothesis
  – determinant of sample size

• Secondary Outcomes
  – relevant to research question
Desirable Features of Outcome Variable

• clinically relevant
• easy to measure
• little measurement error
  – random error – leads to imprecision
  – systematic error – leads to bias
• masked (blinded) ascertainment
Surrogate Outcomes

• Definition: a laboratory measurement or physical sign used as a substitute for a clinically meaningful outcome

• Types: physiologic variable, clinical risk factor, or sub-clinical disease
Advantages of Surrogate Outcomes

• Surrogate outcomes typically increase statistical power compared to clinical outcomes
  – Surrogate outcomes
    • often continuous
    • measured repeatedly
  – Clinical outcomes
    • often categorical
    • surveillance till outcome occurs
Advantages of Surrogate Outcomes (continued)

• Enhanced power means
  – shorter duration of follow-up and/or reduced sample size
  – less cost

• Less contamination by competing comorbidities if the study duration is short

• Useful in studies of mechanisms
Surrogate and clinical outcomes: a continuum

Antecedent of the Risk Factor  Established Risk Factor  Morbid Events  Cause-Specific Mortality  Total Mortality
Surrogate and clinical outcomes: an example

Weight  Blood Pressure  Angina  MI  CVD Mortality  Total Mortality
Relationship between Surrogate and Clinical Outcomes

Probability of Clinical Outcome

Surrogate Outcome
Relationship between Change in Surrogate Outcome and Change in Clinical Outcomes

Change in Probability of Clinical Outcome

Change in Surrogate Outcome
## Clinical and Surrogate Outcomes: Cardiovascular

<table>
<thead>
<tr>
<th>Clinical</th>
<th>Surrogate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke</td>
<td>Ultrasound measurement of intimal medial thickness of the carotid artery</td>
</tr>
<tr>
<td></td>
<td>Blood pressure</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>Quantitative coronary angiography</td>
</tr>
<tr>
<td></td>
<td>Electron beam computerized tomography</td>
</tr>
<tr>
<td>Sudden death</td>
<td>Ventricular arrhythmia</td>
</tr>
<tr>
<td>Heart failure</td>
<td>Ejection fraction</td>
</tr>
</tbody>
</table>
Schematic Overview of Carotid Artery B-Mode Ultrasound Measurements

Internal
Bifurcation

U-S Beam

Common

Interfaces

Near Wall Far Wall Far Wall

A M I I M A

A = Adventitia, M = Media, I = Intima
Weaknesses
Disadvantages of Surrogate Outcomes

• Measurement of surrogate outcomes can involve complex, technical procedures
  – procedures sometimes new (therefore, longitudinal data is scant)
  – procedures become obsolete
  – many technical and analytic issues, often unapparent
Disadvantages of Surrogate Outcomes (continued)

• Missing values are commonplace
• Missing values result from loss to follow-up and poor quality of data
• Potential for bias
  – missing values occur in the sickest people, sometimes because of the clinical outcome of interest
  – informative censoring, that is, loss of follow-up data potentially related to treatment assignment
Models for success and failure of surrogate outcomes*

Model for potential success: Surrogate outcome in the casual pathway

- Disease
- Intervention
- Surrogate Outcome
- Clinical Outcome

Time
Model for potential success: Surrogate outcome in the casual pathway

- Hypertension
- Diuretics
- Blood Pressure
- Stroke

Time
Model for failure: the surrogate is not in the causal pathway of the disease process

Diagram:
- Intervention
- Disease
- Surrogate Outcome
- Clinical Outcome
Model for failure: the surrogate is not in the causal pathway of the disease process

Fluoride

Osteoporosis \[\rightarrow\] ↑ Bone Density \[\rightarrow\] fractures
Model for failure: the intervention affects only the pathway mediated through the surrogate
Model for failure: the intervention affects only the pathway mediated through the surrogate.
Model for failure: The intervention has several mechanisms of action.
Example: Dihydropyridine calcium channel blockers

ASCVD + Calcium Channel Blockers → Blood Pressure → Myocardial Infarction
The Cardiac Arrhythmia Suppression Trial (CAST*): Background

- Ventricular arrhythmias are a risk factor for sudden death after MI
- Four fold higher risk of cardiac mortality among persons with frequent premature ventricular contractions (PVCs)
- In the CAST pilot study, the antiarrhythmic drugs (encainide, flecainide) suppressed PVCs

CAST Research Question

Does suppression of ventricular ectopy after a MI reduce the incidence of sudden death?
CAST Design

• Design: randomized trials of
  – encainide vs placebo
  – flecainide vs placebo
• Participants (n=1498)
  – recent MI (6 days to 2 years ago)
  – ventricular ectopy (6 or more PVCs /hr)
  – at least 80% suppression of PVCs by active drug during open label titration period prior to randomization
Figure 1. Actuarial Probabilities of Freedom from Death or Cardiac Arrest Due to Arrhythmia in 1498 Patients Receiving Encainide or Flecainide or Corresponding Placebo.

The number of patients at risk of an event is shown along the bottom of the figure.

CAST results: number of deaths and cardiac arrests by group

- Active treatment: 63 events / 755
- Placebo: 26 events / 743
  \[ p = 0.0001 \]

- same pattern of results for
  - death from arrhythmia
  - death from any cardiac cause
  - death from any cause
Lessons from CAST

• Active treatments can be harmful (one of several recent trials in which placebo was superior to active treatment)
• Reliance on surrogate outcomes can be misleading
• The scientific community should encourage researchers and sponsors to conduct studies with ‘hard’ clinical outcomes
Examples from the Field

• Surrogate that go in that go the right direction (easy to explain – fit your hypothesis)

• Surrogates that go in unexpected directions (create a greater need for hand-waving and but can still be made to fit your hypothesis)

• Surrogates that behave badly
Model for potential success: Surrogate outcome in the casual pathway

Diet Change

↑ oxidative stress → ↓ oxidative stress → ASCVD

Time
LDL Cholesterol

Oxidized LDL

Fatty Streak Formation

Atherosclerosis

Dietary Antioxidants
Vitamin C
Vitamin E
beta-carotene

Dietary Patterns

Free Radical Activity

Oxidative stress Markers

Inflammatory Markers

Figure 2b
Fig. 1. Oxidation of plasma components (*left axis*; scale, 0–
Figure 2. Correlation between age-specific ischemic heart disease (IHD) mortality and median lipid-standardized $\alpha$-tocopherol levels among men in 16 European countries.\textsuperscript{22}
Nurses Health Study

• Design: Prospective Cohort Study
• Participants: 121,700 female nurses free of diagnosed cardiovascular disease
• Exposure Assessment: Dietary questionnaire at baseline Vitamin E and Multivitamin Use
• Follow-up: 8 years
• End Points: 1) Major Coronary Disease 2) Non-fatal MI 3) Deaths Due to Coronary Disease

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>QUINTILE GROUP FOR VITAMIN E INTAKE</th>
<th>P VALUE FOR TREND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total intake (including supplements)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (IU/day)</td>
<td>2.8</td>
<td>—</td>
</tr>
<tr>
<td>Range (IU/day)</td>
<td>1.2–3.5</td>
<td>—</td>
</tr>
<tr>
<td>Age-adjusted relative risk</td>
<td>1.0</td>
<td>—</td>
</tr>
<tr>
<td>95% Confidence interval</td>
<td>0.70–1.16</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Relative risk adjusted for age and smoking</td>
<td>1.0</td>
<td>—</td>
</tr>
<tr>
<td>95% Confidence interval</td>
<td>0.78–1.28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Dietary intake (without supplements)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (IU/day)</td>
<td>2.6</td>
<td>—</td>
</tr>
<tr>
<td>Range (IU/day)</td>
<td>0.3–3.1</td>
<td>—</td>
</tr>
<tr>
<td>Age-adjusted relative risk</td>
<td>1.0</td>
<td>—</td>
</tr>
<tr>
<td>95% Confidence interval</td>
<td>0.75–1.26</td>
<td>0.12</td>
</tr>
<tr>
<td>Relative risk adjusted for age and smoking</td>
<td>1.0</td>
<td>—</td>
</tr>
<tr>
<td>95% Confidence interval</td>
<td>0.80–1.35</td>
<td>0.99</td>
</tr>
</tbody>
</table>

*Major heart disease includes nonfatal myocardial infarction and death due to coronary disease.
Prospective observational studies of vitamin E: Effects on cardiovascular end points

<table>
<thead>
<tr>
<th>Study (Ref. No.)</th>
<th>Outcomes</th>
<th>No. Events/Sample Size</th>
<th>Minimum Dose Ratio*</th>
<th>Risk Reduction and 95% CI†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nurses (23)</td>
<td>MI/CHD Mortality</td>
<td>552/87 245</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td>Health Professionals (24)</td>
<td>Revascularization MI/CHD Mortality</td>
<td>667/39 910</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>Finland Men (25)</td>
<td>CHD Mortality</td>
<td>186/2748</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>Finland Women (25)</td>
<td>CHD Mortality</td>
<td>58/2348</td>
<td>1.3</td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 1. Mean changes (and 95% CIs) in urinary 8-iso-prostaglandin F$_{2\alpha}$ (PGF$_{2\alpha}$): 9.0 (−125.1, 143.1) in the placebo group, −150.0 (−275.4, −24.6) in the vitamin C group, −141.3 (−230.5, −52.1) in the vitamin E group, and −112.5 (−234.8, 9.8) in the vitamins C + E group. There was no synergistic interactive effect of vitamins C and E ($P = 0.12$).
Summary of Biological Evidence

- Antioxidants are necessary

- Oxidized lipids are associated with CVD

- Oxidation of lipids is reduced by antioxidant supplementation

- Does supplementation lower risk of CVD?
  - Observational studies
  - trials
Do Vitamin E supplements reduce risk?

• Observational studies are confounded – vitamin E takers exercise more, have a lower BMI, eat healthier diets and smoke less often than non-vitamin users.

• Observational studies are hypothesis generating.

• Surrogate markers are only indirectly related to clinical events.

• Benefits can only be assessed in randomized controlled clinical trials.
Clinical Trials – Clinical Outcomes

• Cardiovascular Events
  – Fatal and Non-fatal MI
  – Stroke
  – Peripheral artery disease

• Mortality
ATBC Study

• Design: Randomized, double-blind, placebo-controlled primary prevention trial

• Participants: 29,133 male Finnish smokers, age 50-69

• Intervention: 1) Vitamin E 50 IU/day  
2) B-carotene 20 mg/day  
3) Combination  
4) Placebo

• Follow-up: 5-8 years

• End Points: Incident lung cancer & deaths

ATBC, 1993 NEJM
Figure 3: Kaplan-Meier estimates of mortality
ATBC Trial Results

- Beta-carotene group (20 mg/day)
  - increase in total mortality (9%)
  - increased incidence of angina (13%)*
  - increased CVD mortality (11%)*
  - increased incidence of lung cancer (18%)

- Vitamin E Group (50 mg/day)
  - reduction in total coronary events (3%)
  - reduction in incident angina (9%)
  - reduction in non-fatal MI (11%)

ATBC, 1994 NEJM
CARET Study

- **Design:** Randomized, double blind, placebo-controlled primary prevention trial
- **Participants:** 18,314 smokers, former smokers, and workers exposed to asbestos
- **Intervention:**
  1) B-carotene (30 mg/day) and vitamin A (25,000 IU/day)
  2) Placebo
- **Follow-up:** 4 years
- **End Points:** Incident lung cancer, Cardiovascular Disease

Omenn, 1996 NEJM
Meta-Analysis: High-Dosage Vitamin E Supplementation May Increase All-Cause Mortality

Edgar R. Miller III, MD, PhD; Roberto Pastor-Barriuso, PhD; Darshan Dalal, MD, MPH; Rudolph A. Riemersma, PhD, FRCPE; Lawrence J. Appel, MD, MPH; and Eliseo Guallar, MD, DrPH

Background: Experimental models and observational studies suggest that vitamin E supplementation may prevent cardiovascular disease and cancer. However, several trials of high-dosage vitamin E supplementation showed non-statistically significant increases in total mortality.

Purpose: To perform a meta-analysis of the dose–response relationship between vitamin E supplementation and total mortality by using data from randomized, controlled trials.

Patients: 135,967 participants in 19 clinical trials. Of these trials, 9 tested vitamin E alone and 10 tested vitamin E combined with other vitamins or minerals. The dosages of vitamin E ranged from 16.5 to 2000 IU/d (median, 400 IU/d).

Data Sources: PubMed search from 1966 through August 2004, complemented by a search of the Cochrane Clinical Trials Database and review of citations of published reviews and meta-analyses. No language restrictions were applied.

Data Extraction: 3 investigators independently abstracted study reports. The investigators of the original publications were contacted if required information was not available.

Data Synthesis: 9 of 11 trials testing high-dosage vitamin E (≥400 IU/d) showed increased risk (risk difference > 0) for all-cause mortality in comparisons of vitamin E versus control. The pooled all-cause mortality risk difference in high-dosage vitamin E trials was 39 per 10,000 persons (95% CI, 3 to 74 per 10,000 persons; P = 0.035). For low-dosage vitamin E trials, the risk difference was −16 per 10,000 persons (CI, −41 to 10 per 10,000 persons; P > 0.2). A dose–response analysis showed a statistically significant relationship between vitamin E dosage and all-cause mortality, with increased risk of dosages greater than 150 IU/d.

Limitations: High-dosage (≥400 IU/d) trials were often small and were performed in patients with chronic diseases. The generalizability of the findings to healthy adults is uncertain. Precise estimation of the threshold at which risk increases is difficult.

Conclusion: High-dosage (≥400 IU/d) vitamin E supplements may increase all-cause mortality and should be avoided.

For author affiliations, see end of text.
Failed surrogate marker: example

β-carotene supplements

Smoking
↓β-carotene

↓β-carotene
↑Lung Cancer

Need for reliable surrogate markers
Disadvantages of Surrogate Outcomes (continued)

• The relationship between a surrogate outcome and a clinical outcome has face validity but is often uncertain

• Relationship between change in surrogate and risk of clinical outcomes is rarely known
The Bottom Line

“Trust but verify”

Ronald Reagan
Analytical Issues

• Sample Size (Power Calculations)

• Analytical Approach (a priori)

• Intention-to-treat (vs ‘as treated’)

Analytic Techniques: Crude analyses

• Analysis depends on the type of outcome data
• Basic tests
  – Continuous outcome variable: \texttt{t-test}
    • Examples: Blood pressure, serum cholesterol
  – Dichotomous or categorical data: \texttt{chi-squared}, \texttt{logistic regression}, \texttt{cox modeling for time to event}
    • Example: Incident HIV, MI, cancer, renal failure, death
Analytic Techniques: Adjusted (Regression) Analyses

• Regression determines association between exposure and outcome
• Procedures depends on outcome variable:
  – Continuous outcome: linear regression
  – Dichotomous outcome: logistic regression
  – Time-to-event: Cox proportional hazards
Interpretation of Results

• Internal Validity
  – conclusions correctly describe what happened in the study

• External Validity (‘generalizability’)
  – the degree to which the conclusions apply to the study population and other populations
Why RCTs Can Be Difficult

• Hard to find and recruit the right people
  – Many don’t want to be “guinea pigs”
• Greater responsibility, documentation
• May take years for outcomes to develop
• People are free to do as they please
  – Some assigned to treatment don’t adhere
  – Some assigned to control seek treatment
  – Some drop out of the trial completely
Adherence (compliance)

• Difficult to measure

• Difficult to promote

• Must be promoted and measured, at least in efficacy or explanatory trials
Why be worried about adherence?

Active

Control

Drop-In’s

Drop-Out’s

Intention-to-Treat: analysis by randomized group, not by final groupings
Adherence (compliance)

• Measurement
  – self report
  – pill count
  – blood levels of drug
  – biological changes (urine or blood)

• Promotion
  – exclude poor candidates before randomization
  – keep intervention simple
  – respond to evidence of inadequate adherence
How To Handle Participants Who Don’t Adhere to Trial Assignment

• Intention-to-Treat Approach
  – Least optimistic
  – Maintains initial balance from randomization
  – Highlights problems from adverse effects

• On-Treatment Approach
  – Most optimistic
  – Upsets initial balance from randomization
  – Downplays problems from adverse effects

Because of its conservatism, the Intention-to-Treat approach is strongly preferred.
Cardiac Event-Free Survival in 192 Adults with Refractory Angina by Random Assignment and Cross-Over (from Medical Treatment to TMR) Status

- Randomized to TMR, no crossing over to Medical Rx
- Randomized to Medical Rx, did poorly, needed TMR as last ditch effort
- Randomized to Medical Rx, did OK, no need for TMR

TMR = transmyocardial laser revascularization
Cardiac Event-Free Survival in 192 Adults with Refractory Angina by Random Assignment and Cross-Over (from Medical Treatment to TMR) Status

Were X-overs reclassified as “TMR”, it would tend to make TMR look worse.
Cardiac Event-Free Survival in 192 Adults with Refractory Angina by Random Assignment and Cross-Over (from Medical Treatment to TMR) Status

Were X-overs classified as “Medical Rx”, it would tend to make Medical Rx look better
Clinical Trials: Design and interpretation Considerations
When Trials Are Impossible (or Nearly Impossible)

- Adverse Exposures (e.g. Cigarettes)
- Rare Outcomes (e.g. Reye’s Syndrome)
- Intervention Already in Wide Use

In these circumstances, one must rely on observational studies—i.e. prospective cohort studies and case-control studies. When interventions are already in wide use, “outcomes research” is a good option. In outcomes research, medical interventions (e.g. drugs, surgical procedures) are considered as exposures. Data on these interventions, and on relevant clinical outcomes, are available from medical records and often from large-scale electronic databases.
Statistical vs Clinical Significance

• Statistical significance pertains to whether or not the observed results could occur from chance alone.

• Clinical significance pertains to whether or not the observed results have “important” clinical, research or public health relevance.
How To Interpret Negative Results

• Treatment is worthless
• Treatment is worthwhile, **BUT** study had…
  – Bias against the treatment (e.g. crossing in)
  – Inadequate contrast between groups
    • Suboptimal treatment (e.g. unskilled surgeons)
    • Low adherence (e.g. drug causes GI distress)
    • Controls sought treatment despite assignment
  – Insufficient statistical power
    • Very common cause of negative findings
    • Meta-analysis a potential remedy
Efficacy (Explanatory) Trial vs Effectiveness (Pragmatic) Trial

• Theory
  – Efficacy: What is the effect of the therapy under ideal conditions
  – Effectiveness: What is the effect of therapy under ‘real world’ conditions

• Reality
  – The dichotomy between efficacy and effectiveness is artificial
  – Broad continuum
Typical Implementation Units

• Clinical Centers
  – recruit participants
  – collect data
  – administer intervention/therapy

• Laboratory or Reading Centers
  – perform assays or readings of procedures

• Data Coordinating Center*
  – receive/assemble data
  – coordinate activities
  – perform data analyses

* similar to Contract Research Organization (CRO)
Oversight Units

• Internal
  – Sponsor
  – Data Coordinating Center or Contract Research Organization

• External
  – Institutional Review Board
  – Data and Safety Monitoring Board
Organizational Structure of a Multi-Center Trial
(Weight Loss Maintenance Trial)

- Steering Committee
- NIH Project Office
- DSMB
- Subcommittees
- Design & Analysis
- Publications
- Measurement & Quality Control
- Clinic Coordinators
- Enrollment and Retention
- Intervention
- Minority Implementation
- Clinical Centers
- Center for Health Research
- Johns Hopkins University
- Pennington LSU
- Duke University
- Coordinating Center
- Data Management
- Intervention Development