Welcome!
Clinical and Translational Science Awards (CTSA) Program

- National Institutes of Health (NIH) program
- Launched in 2006
- A national consortium of 64 institutions
- **Mission:** to develop innovative solutions that will improve the efficiency, quality and impact of the process for turning observation in the laboratory, clinic and community into interventions that improve the health of individuals and the public
Spectrum of Clinical and Translational Research

Bench to Bedside...

...to Widespread Clinical Practice...

...to Public Health...

...to Health Policy

Translation (T1)

Translation (T2)

Translation (T3)

Translation (T4)
Tufts CTSI’s Mission & Purpose

Established in 2008 to translate research into better health

- Stimulate and expedite innovative clinical and translational research, with the goal of improving the public’s health
- *Entire spectrum* of clinical and translational research is critical to meeting the promise and the public’s needs of biomedical science
39 Tufts CTSI Partners

13 Tufts Schools & Centers
Cummings School of Veterinary Medicine
Fletcher School of Law & Diplomacy
Friedman School of Nutrition Science & Policy
Graduate School of Arts & Sciences
Institute for Clinical Research & Health Policy Studies at Tufts Medical Center
Jean Mayer USDA Human Nutrition Research Center on Aging
Sackler School of Graduate Biomedical Sciences
School of Dental Medicine
School of Engineering
School of Medicine
Tisch College of Citizenship & Public Service
Tufts Center for the Study of Drug Development
Tufts Innovation Institute

7 Tufts-Affiliated Hospitals
Baystate Medical Center
Lahey Clinic
Maine Medical Center
New England Baptist Hospital
Newton-Wellesley Hospital
St. Elizabeth’s Medical Center
Tufts Medical Center

6 Industry/Non-Profit Partners
Blue Cross Blue Shield of Massachusetts
Eli Lilly and Company
Institute for Systems Biology and P4 Medicine Institute
Minuteman Health Network
Pfizer, Inc.
Tufts Health Plan

10 Community-Based Partners
Action for Boston Community Development (ABCD)
Asian Community Development Corporation
Asian Task Force Against Domestic Violence
Asian Women for Health
Boston Chinatown Neighborhood Center
Center for Information and Study on Clinical Research Participation
Greater Boston Chinese Golden Age Center
Health Resources in Action
Museum of Science, Boston
New England Quality Care Alliance

3 Academic Partners
Brandeis University
Northeastern University
RAND Corporation
How Can CTSI Help?

- **Connections** with other researchers, industry, the community, and policy-makers across the Tufts CTSI network and national CTSA consortium via our Navigators & Research Collaboration team.

- **Consultations** on **comparative effectiveness**, **one health**, **research process improvement** and **stakeholder and community engagement** projects and grants, as well as **regulatory issues** and other areas of translation.

- **Study design and data analysis** (pre- and post-award) through the **Biostatistics, Epidemiology, and Research Design (BERD) Center**, including drop-in sessions.
How Can CTSI Help?

• **24/7 clinical trial support** through our **Clinical and Translational Research Center (CTRC)**.

• **Informatics tools** for electronic data capture (**REDCap**), resource sharing, and collaboration.

• **Training & professional development** including MS and PhD degrees, certificate programs, seminars & workshops, and paid career development awards and fellowships.

• **Funding** through one-year interdisciplinary **pilot studies grants** that support the initial stages of research.
How to Request Tufts CTSI Services

• Visit www.tuftsctsi.org and submit a request
http://ilearn.tuftsctsi.org/

Live seminars are recorded for our I LEARN site. Seminar videos can be viewed at any time, and are free!
Get Connected: CTSI Happenings

- Weekly e-newsletter with news, professional development and funding opportunities, resources, and success stories.
- Issued every Monday at 8AM
- Sign up on our website or at http://eepurl.com/C4d9X
For more information: www.tuftsctsi.org
The Clinical Trial Protocol

How to develop a protocol including all the critical elements

Andreas Klein, MD

Director, Hematologic Malignancies Program
Assistant Director, Bone Marrow and Hematopoietic Cell Transplant Program
Chair, Tufts Health Sciences Campus Institutional Review Boards
Associate Professor, Tufts University School of Medicine
In your packets

- Agenda
- Evaluation
- Print out of slides
- Example protocol contents and objectives
- Example protocol template (Tufts IRB)
- Example protocol Table of Contents (CTEP)
- Series Pre-Work (on I LEARN)
Introduction

• General aspects of protocol development
• Components
• Tools
Research vs Clinical Practice

• Research
  – Definition

• Clinical Practice
  – Treatment plans
  – Standard operating procedures

• Research vs “Treatment”
  – Treatment implies a potential for beneficial outcome
  – Therapeutic misconception in research
Patients vs. Subjects

• Patients
  – clinical care managed according to physician best judgement

• Subjects
  – voluntary participants in clinical trial
  – managed according to defined plan
What is a Protocol

• Formal description of the planned work
• Guide to follow explicitly
  – Work out every detail before hand
  – Anticipate every *reasonable* contingency
• Anyone should be able to pick up protocol and know what to do
What Kind of Research Requires a Protocol?

- ALL research needs a plan
  - Research = systematic investigation designed to develop generalizable knowledge (45 CFR 46.101)

- All research involving human subjects needs a protocol
  - Abbreviated protocols may be ok for some research
  - The greater the complexity or risk, the more defined the procedures need to be
What is a Protocol Not

• Not *absolute*: subject safety ALWAYS comes first
• Not malleable
  – Revisions require careful consideration
  – E9704
• Not open to interpretation
Protocol Review Responsibilities

• Disease Groups
  – General enthusiasm of colleagues
  – Commitment from co-investigators

• Resources & priorities
  – PRMC

• Scientific
  – SRC, delegated from IRB

• Ethical
  – IRB
Protocol Review Criteria

• Is the plan adequate to address the stated objectives?
• Is the selection of participants equitable?
  – Language restrictions
• Are all reasonable risks minimized?
  – Are assessments frequent enough to catch problems in time?
  – Is the starting dose reasonably expected to be safe?
• Are the expected risks justified by anticipated benefits?
  – Answer may depend on phase of study, underlying conditions
Where Do Protocols Come From?

- Industry, cooperative group, individual investigator
- Recycling
- Online tools (CTEP)
Birth of a Protocol

• Starts with an idea
  – Clinical observation
  – Compelling biological hypothesis
  – Promising early phase data
• Defining objectives
  – Demonstrate that Drug X is safe and effective
• Design an intervention to address the objectives
Elements of a Protocol

- Schema
- Introduction
  - Rationale
- Statement of Objectives
- Eligibility / Subject Selection
- Treatment Plan
- Study Evaluations
- Dosing Delays / Modifications
- Safety Monitoring Plan
- Description of study article / device
- Correlative Studies
- Measurement of Effect
- Study Oversight / Administrative Details
- Statistical Plan
- References
- Appendix 1: Study Calendar
- Additional Appendices
Schema

• Outline in graphical form
  – Highlight overall format of the study
    • Selection
    • Treatment assignment(s)
    • Randomization
    • Follow up
Introduction

• ALL necessary background
  – Disease / condition targeted
    • Epidemiology
    • Natural history / expected outcomes
  – Current standard(s) of care
  – Description of study article / device / intervention
  – Rationale for proposed research
Statement of Objectives

• Objectives vs. Outcomes
  – Objectives – what you want the research to show
  – Outcomes – measures you will use to support the objectives
  – Example

• Primary vs Secondary

• Exploratory
Eligibility / Subject Selection

• Inclusion
  – Define population for intervention
    • Diagnosis & stage
    • Allowed prior therapy (or not)
    • Age range
    • ECOG
    • Organ and marrow function requirements

• Exclusion
  – From included, define who should not take part
    • Features associated with increased risk

  – Allergies to study article
  • Features associated with limited benefit
    – Exclusion for CNS disease
  • Features associated with interference with study
    – Recent prior therapy

• Inclusion of Women and Minorities
  – Recruitment plan required for federally funded research

• Risk / Benefit Assessment
Treatment Plan

• Study article administration details
  – Dose, Route, Schedule
• Device use instructions
• Concomitant drugs
  – Premeds
  – Required supportive care
Study Evaluations

- Assessments required to support Objectives = outcomes
- Baseline assessments
- On study assessments
  - Toxicity
  - Response
- Follow up assessments
  - Duration of response
  - Long term / delayed toxicities
Dosing Delays / Modifications

• Modifications for toxicity
  – Toxicity triggers
  – Dose reductions
  – Delay schedules
  – Discontinuation criteria
Safety Monitoring Plan

• Adverse event definition
  – Untoward event or lab abnormality occurring during study
  – Severity (AE vs SAE)
  – Grade (CTCAE)
  – Attribution

• Unanticipated problems (UP)
  – Not expected based on condition or known effects of study article
  – Possibly or probably related to study participation
  – Put(s) others at greater risk than anticipated
Reporting Procedures

- Local / Institutional
  - IRB
- HIPAA Privacy Office
- Federal
  - FDA
    - Required for covered research (IND/IDE)
    - Voluntary for non-covered research
  - OHRP
    - Required for ALL unanticipated or recurring problems (45 CFR 46.103)

- Other investigators
  - Critical in multi-institutional studies
    - Ad hoc vs scheduled
Description of Study Article or Device

• Basic information
  – Source
  – Form
  – Storage and preparation instructions

• Can reference official sources
  – Investigator brochure – unapproved drug / device
  – Package insert – FDA approved
Correlative Studies

- Typically support secondary objectives
  - Exploratory biomarkers
  - Immunologic assays
  - Translational collaborations
Measurement of Effect

- Outcomes
  - Response criteria definitions
    - RECIST
    - NCI Working group
- Define time-dependent parameters
  - OS, PFS, RFS, TTTF, TTNT
- Other outcome measures
  - How does a lab value translate to a measurement of effect?
Study Oversight / Administrative Details

- Task delegation
- Procedures for registering subject
- Procedures and timelines for collecting data
- Recruitment and retention strategies
- Extras
  - Tissue banking
  - Central data review
Statistical Plan

• How will you analyze your data
• How do you justify your design
• Number of subjects / power calculations
• Covered in detail in following sessions
References

- Literature citations in support of Background
- Links to specific tools / techniques cited in procedures
Appendix A: Study Calendar

- Lists all planned assessments and dates/timing
- Quick reference to make sure everything done at the right time
- Must match Study Evaluations section

<table>
<thead>
<tr>
<th></th>
<th>Pre-Study</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
<th>8th</th>
<th>9th</th>
<th>10th</th>
<th>11th</th>
<th>12th</th>
<th>Diff Stands</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Study Drug A]</td>
<td>A</td>
<td>a</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Other aspect]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infusion consent</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical history</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consent mark</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical exam</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vital signs</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance state</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBC w/diff</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B...

- Protocol specific definitions
  - Expected standard of care interventions
  - Allowable conditioning regimens
- General references
  - ECOG scale
  - Registration / data submission forms
Tools and Resources

• CTSI
  – BERD (Wednesday am drop-in) and Regulatory Affairs (consultation)

• Tufts IRB Templates
  – http://viceprovost.tufts.edu/HSCIRB/templates/protocol-templates/

• CTEP
  – Protocol templates for multiple phases of study
  – Common Terminology Criteria for Adverse Events (CTCAE)
Pet Peeves

• Version control is *critical*
  – Choose a scheme and stick to it
    • Embed in file name
    • Include “Version date” in header or footer
  – Update with each iteration
• Only work on one version at a time
• Get someone(s) to proofread your protocol
• Recycle / Reuse / Borrow
  – But don’t make it obvious!
Final Tips

• Keep it simple
  – Don’t try to cure *all cancers all at once*
  – Don’t impose too many measurements / procedures
    • You are responsible for non-compliance if you miss anything…
  – Be mindful of the costs and who is going to pay

• Talk to IRB and/or SRC *early*
  – Simple changes in design can have large impacts on regulatory burden
  – Expeditable vs Full Board review
#Seriously?

- Yes, you can do it!
- Practice makes perfect
- Reading lots of protocols helps
- Join a protocol review committee (IRB or SRC)
Thank You
Successfully Collaborating with Statisticians

Lori Lyn Price, MAS
Assistant Professor
Tufts University School of Medicine
Statistician
BERD Center, Tufts CTSI
Collaboration

Collaboration is the process of two or more people working together to realize or achieve something successfully (Wikipedia).

Objective:
• Know what you can do to make the first meeting and collaboration more successful
• Understand the roles and responsibilities of collaborative researchers and statisticians
Overview

- Why collaborate with a statistician or epidemiologist
  - Single project
  - Multiple projects
- Preparation for first meeting
- First meeting
- After first meeting
Collaboration with Statistician Maximizes Probability of Successful Study

• Design study in a way that will adequately answer your study questions

• Calculate how many participants are needed

• Design feasible study aims & hypotheses

• Define outcomes

• Develop appropriate statistical analysis plan
What Might Happen if You Don’t Meet with Statistician Prior to Starting Study

• Data not collected at correct time points
  - Want to compare change in outcome from day 1 to day 5
  - Only 25% of participants have data collected at days 1 and 5
  - The rest have data collected somewhere between days 3 & 7

What to do?
When to Meet With Statistician

• As early as possible
  Grant—at least 2 months before it is due
  IRB protocol—at least 1 month before it is due
  Abstract—at least 1 month before it is due

• The closer to the deadline, the less help the statistician may be able to provide
Prior to 1st Meeting with Statistician

• Meet with research mentor (if applicable) & team to discuss
  – Research questions
  – Hypotheses
  – Outcomes
  – Risk factors

• Ask research mentor (if applicable) to attend meeting with statistician
Prior to 1\textsuperscript{st} Meeting with Statistician

- Identify key articles in literature relevant to your study
  - May be used for sample size calculations
  - May want to replicate methods
Prior to 1st Meeting with Statistician

• Send statistician the following a few days before 1st meeting
  - Background
  - Research aims and hypotheses
  - Primary outcomes
  - Primary predictor variables/risk factors
Example of What We Would Like to See

Instead of “I need help with study design”:

• I am implementing a hospital wide intervention and am unsure whether I need a control arm to answer my research question.
  - If I need a control arm, how should it be selected?

• I can realistically enroll only 20 participants in this study. Given that constraint, what can I reasonably hope to conclude about my research question?
Example of What We Would Like to See

Instead of “I need help with study design”:

• Which of the following definitions of the outcome variable will answer my research question? I prefer definition A, but other groups have used definition B. What are the pros and cons of using each definition?
Example of What We Would Like to See

Instead of “I need help with an analysis plan”:

• I want to replicate the methods in the attached paper. Would following the methods in this paper be appropriate for my research question?

• How do I know which variables I should consider adjusting for in a multivariable regression analysis?
Statistician’s Role Before First Meeting

• Read the material you have sent

• Prepare a list of questions she has

• Prepare some suggestions
Discussion

• You are meeting with a statistician to discuss a new research project

• What questions will you have for her?

• Will you be able to provide a rough draft of research questions, hypotheses and outcomes?
  - If not, what will you do prior to the meeting to further develop your ideas?
At First Meeting

- Discuss your research questions and possible study designs and analyses
  - Pros and cons of each
  - Budgetary and sample size constraints
  - Data management (not necessarily discussed)
  - Ask questions if you don’t understand

- Not unusual to leave meeting with a much modified research question
At first meeting

- Clarify deadlines and timelines
- Confirm with statistician expectations about
  - Fees
  - Authorship
  - What she will provide to you and what you will provide to her
After the 1st Meeting

• Generally you will leave with questions that you need to answer re:
  – Sample size calculations
  – Definition of variables
  – Whether a particular study design or analysis is preferred by other members of your research team
After the 1\textsuperscript{st} meeting

- The statistician will also often have a to-do list
  - Think/research further about the possibilities for study design and analysis
  - Run sample size calculations
  - Draft analysis plan

- There is iterative communication via email or subsequent meetings to finalize the details
Communication

• Good communication from both parties is key to a successful collaboration
• Both statistician and researcher should be able to explain key points in lay language
• Ask questions throughout process if you don’t understand anything
• Be prepared for the statistician to also ask many questions to better understand your research
Take Home Message

• Preparation is key to a successful first meeting and collaboration with a statistician

• Both researchers and statisticians have roles and responsibilities in a successful collaboration effort
Thank You
Break!
Weighing the Merits of Observational and Experimental Clinical Research

Jessica Paulus, ScD

Assistant Professor of Medicine
Tufts University School of Medicine

Associate Director
Tufts Clinical and Translational Science Program
Outline and goals

1. Evidence based medicine – a definition
2. Discuss the strengths and limitations of the major observational study designs
3. Does sunscreen prevent melanoma? – integrating the evidence across study designs
4. Metformin as a treatment strategy for colorectal cancer – a multipronged approach
No one study can answer a research question definitively.

Need to look at the current status of knowledge, i.e., the totality of evidence.
Evidence Based Medicine: A definition

“. . the conscientious, explicit and judicious use of current best evidence in making decisions about the care of the individual patient. It means integrating individual clinical expertise with the best available external clinical evidence from systematic research.” (Sackett D, 1996)
Does sunscreen prevent melanoma?
Group 1: Human Carcinogen

SOLAR AND ULTRAVIOLET RADIATION
Solar radiation (Group 1)
Ultraviolet A radiation (Group 2A)
Ultraviolet B radiation (Group 2A)
Ultraviolet C radiation (Group 2A)
Use of sunlamps and sunbeds (Group 2A)
Exposure to fluorescent lighting (Group 3)
Age-standardized death rates from Melanoma and other skin cancers by country (per 100,000 inhabitants)

Data from: Death and DALY estimates for 2004 by cause for WHO Member States (Persons, all ages)
Melanoma of the Skin Incidence Rates* by State, 2007: an Ecologic study

- *Age-adjusted to the 2000 U.S. standard population.
Prevention: Skin should be protected from intense sun exposure by wearing tightly woven clothing and a wide-brimmed hat, applying sunscreen that has a sun protection factor (SPF) of 30 or higher to unprotected skin, seeking shade (especially at midday, when the sun’s rays are strongest), and avoiding sun-bathing and indoor tanning. Sunglasses should be worn to protect the skin around the eyes. Children should be especially protected from the sun because severe sunburns in childhood
Solar Radiation and Skin Cancer

Six lines of evidence:
1. Skin cancer occurs more frequently in residents of areas of high solar radiation
2. Skin cancer occurs more frequently in sun-sensitive people
3. Skin cancer occurs more frequently in sun-exposed body sites
4. Skin cancer occurs more frequently in people with a history of sunburn
5. Skin cancer occurs more frequently in people who have a benign sun-related skin condition
6. Skin cancer is reduced by protection of skin against the sun
Sunscreen is controversial?

“Sunscam: Think sunscreen protects against cancer? Think again” –Mother Jones, 1998

The sunscreen and melanoma controversy. –Archives of Dermatology, 1999

Sunscreens as a preventative measure in melanoma: an evidence-based approach or the precautionary principle? - British J Dermatology, 2009

“Sunscreens and melanoma: an on-going controversy” –Melanoma Research, 2010

As Summer Nears, Sunscreen Controversy Reignites - CBS News, 2010
Sunscreen does not meet principles of evidence-based medicine?

“Applying the principles of evidence-based medicine, there is not the strength of evidence to use sunscreens as a preventative measure in melanoma as would be expected before a new drug was introduced as a therapeutic intervention.

And to acquire the evidence... would take a decade or more. Because of this lack of evidence, it has been argued that the focus of recommendations for melanoma prevention in public health campaigns should be more emphasized to sun avoidance, shade and clothing.

On the other hand, the precautionary principle, which states that if an action or policy might cause severe or irreversible harm to the public, then in the absence of a scientific consensus that harm would not ensue, the burden of proof falls on those who would advocate taking the action. In other words, those who advocate that sunscreens should not be used as a preventative measure in melanoma because of lack of evidence for their efficacy must demonstrate this lack of efficacy for their advice to be followed. Logic would suggest that this demonstration of lack of efficacy would be difficult as exposure to UV radiation is widely recognised as a risk factor in melanoma and modern sunscreens attenuate the intensity of solar UV entering the skin, the magnitude of attenuation depending more on compliance and application technique than technical performance (i.e. SPF and UVA rating) of the product.”

BL Diffey, British Journal of Dermatology 2009 161 (Suppl. 3), pp25–27
Totality of Evidence

**Basic Research** – lab, animals, mechanism (precision, ? relevance)

**Epidemiologic Studies** – direct evidence in humans (relevance, ? precision)

- Observational studies
  - Case-control
  - Cohort

- Interventional or experimental studies
  - Randomized clinical trials
Observational vs. Experimental Studies

1. Observational Studies
   Investigator *observes* the relationship between a risk factor or treatment and health outcome
   ✓ Case-control, Cohort, Cross-sectional, Ecologic studies

2. Interventional, or Experimental, Studies
   Investigator *assigns* treatment status for the primary reason of assessing the scientific question at hand
   ✓ Randomized clinical trials
Hierarchy of Study Designs?

Meta-analysis/Evidence Synthesis

Randomized Controlled Trials

Cohort studies  Case-control studies

Cross-sectional studies

Ecologic studies

Case reports
Observational studies of sunscreen use and risk of melanoma
Prospective Cohort Study

November 2016

- **Sunscreen user**
- **Non-user**

Melanoma?

Basis on which groups are selected at beginning of study
Retrospective Cohort Study

November 2016

Sunscreen user

Melanoma?

Non-user

Basis on which groups are selected at beginning of study
Case-Control Study

- **Cases:** Melanoma
- **Controls:** No melanoma

Basis on which groups are selected at beginning of study

November 2016

Used sunscreen?
“Originally developed to protect against sunburn, sunscreen has been assumed to prevent skin cancer. However, conflicting reports include claims that sunscreen increases risk for melanoma.”

Meta-Analysis of 18 Case-Control Studies

Odds Ratio = odds of melanoma in sunscreen users / odds of melanoma in non-users
Use of Topical Sunscreens and the Risk of Malignant Melanoma: A Meta-Analysis of 9067 Patients From 11 Case-Control Studies

Michael Huncharek, MD, MPH, and Bruce Kapelman, BA

Objectives. This study examined the methodology of epidemiological studies that suggest use of topical sunscreen preparations is associated with increased risk of malignant melanoma.

Methods. We pooled data from observational studies using a general variance-based meta-analytic method that employed confidence intervals (previously described). The outcome of interest was a summary relative risk (RR) reflecting the risk of melanoma associated with sunscreen use versus nonuse. Sensitivity analyses were performed when necessary to explain any observed statistical heterogeneity.

Results. Combining studies that used nonheterogeneous data yielded a summary RR of 1.01, indicating no association between sunscreen use and development of malignant melanoma.

Conclusions. The available epidemiological data do not support the existence of a relationship between topical sunscreen use and an increased risk of cutaneous malignant melanoma. (Am J Public Health. 2002;92:1173-1177)
Today's Random Medical News

According to a report released today...

Jim Borgman
What are some possible explanations for this surprising finding?
**Conclusions:** No association was seen between melanoma and sunscreen use. Failure to control for confounding factors may explain previous reports of positive associations linking melanoma to sunscreen use. In addition, it may take decades to detect a protective association between melanoma and use of the newer formulations of sunscreens.
Confounding

• A confounder is a third variable (one other than the exposure and the outcome) that creates a spurious association between the exposure and outcome or can mask a true relationship that does exist.

• A “mixing of effects”
3 Properties of a Confounder

1. Is associated with the exposure or treatment

2. Is an independent predictor of the disease or health outcome

3. It is not a consequence of the exposure or treatment
A confounder is a "common cause" of exposure and outcome.

Exposure: Sunscreen

Outcome: Melanoma

Confounder

Property 1
Property 2
Property 3
Higher UV Exposure Linked to Sunscreen Use

Fig 2. UV doses [in standard erythemal dose (SED) per day] received by volunteers wearing personal UV dosimeters, Denmark (Ref. 14).

Sun Sensitivity Linked to Sunscreen Use

<table>
<thead>
<tr>
<th>Phenotypic Factor</th>
<th>%</th>
<th>OR (95% CI)</th>
<th>OR (95% CI)(^+)</th>
<th>OR (95% CI)(^++)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin reaction to sunlight</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blistering sunburn</td>
<td>23.9</td>
<td>7.71 (3.88–15.30)</td>
<td>6.04 (2.82–12.95)</td>
<td>4.85 (2.56–9.21)</td>
</tr>
<tr>
<td>Sunburn without blisters</td>
<td>22.5</td>
<td>6.95 (4.09–11.81)</td>
<td>5.73 (3.21–10.23)</td>
<td>4.23 (2.61–6.86)</td>
</tr>
<tr>
<td>Mild sunburn that becomes a tan</td>
<td>10.1</td>
<td>2.44 (1.43–4.18)</td>
<td>2.35 (1.32–4.19)</td>
<td>2.25 (1.40–3.64)</td>
</tr>
<tr>
<td>Tan or darken with no sunburn</td>
<td>4.3</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td>No change in skin color</td>
<td>11.3</td>
<td>2.81 (0.98–8.04)</td>
<td>2.61 (0.81–8.45)</td>
<td>2.27 (0.83–6.19)</td>
</tr>
</tbody>
</table>
Positive Confounding

Bias is in the *upwards* direction

Sunscreen → ? → Melanoma

Sun sensitivity,
Sun exposure,
Etc…

*Tufts CTSI*

Tufts Clinical and Translational Science Institute
Adjusted for Sun-Sensitivity
What is a Randomized Clinical Trial?

In a randomized controlled clinical trial:

Participants who are eligible are *randomly* assigned to

- Sunscreen
- No sunscreen

→ ?

Melanoma → ?

*Tufts CTSI*
Equipoise

✓ To justify random treatment assignment, principle of equipoise must hold

✓ Must be adequate uncertainty about benefit/risk of investigational agent

✓ Equipoise may exist if:

  Previous trials were conducted in animals or cell culture

  Previous trials were not definitive with respect to benefit and/or risk

  Previous trials were conducted in a different population and application to a new population might be unclear
RCT design minimizes bias

✓ RCT's are optimal to detect small to moderate, but clinically worthwhile, treatment effects because they can minimize sources of bias through randomization, blinding, placebos, etc.

✓ When treatment is assigned by a “coin flip,” and assuming a large enough sample size, the active and comparison groups will have an equal distribution of other risk factors.

✓ Freedom from confounding by all known and unknown factors
Why are RCTs free from confounding bias (if properly conducted)?
Randomization as a Tactic to Limit Confounding Bias

Sunscreen → Melanoma

Age, sex, socio-economic status, genetics, sun-sensitivity …

… And every covariate you hadn’t anticipated or measured
An experimental study of sunscreen and melanoma
Randomized Trial Design

1986: Queensland township residents were randomly sampled from the electoral roll for a skin cancer prevalence survey.

1992: Surveyed were invited to participate in a RCT about sunscreen.
A (Pragmatic) Trial Design

Sunscreen group (n=812):

- Given a free, unlimited supply of broad-spectrum sunscreen with a sun protection factor (SPF) of 16
- Asked to apply to head, neck, arms, and hands every morning
- Reapplication advised after heavy sweating, bathing, or long sun exposure

Comparison group (n=809):

- Continued using sunscreen of any SPF at their usual, discretionary frequency, which included no use
- Allocation of a placebo sunscreen to the control group was deemed unethical, given the subtropical location
Compliance assessed via participant diaries, weighing sunscreen bottles

- Sunscreen group: 75%
- Control group: did not use sunscreen (38%), used 1-2x/week (35%), used it at non-intervention sites (8%)

Dermatologists blinded to treatment assignment examined participants for melanoma in 1992, 1994, 1996 (scheduled trial completion)

- After 1996, participants completed questionnaires about new skin cancers
- Queensland Cancer Registry
Balanced Baseline Characteristics

Table 1. Demographic and Clinical Characteristics of Participants at Baseline in 1992 According to Sunscreen Allocation

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Intervention*</th>
<th></th>
<th></th>
<th></th>
<th>P+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sunscreen 812</td>
<td>No Sunscreen 809</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>356</td>
<td>354</td>
<td>44</td>
<td>44</td>
<td>.97</td>
</tr>
<tr>
<td>Female</td>
<td>460</td>
<td>455</td>
<td>56</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 50</td>
<td>446</td>
<td>442</td>
<td>55</td>
<td></td>
<td>.97</td>
</tr>
<tr>
<td>50-59</td>
<td>166</td>
<td>164</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 60</td>
<td>200</td>
<td>203</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin color</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair</td>
<td>463</td>
<td>442</td>
<td>56</td>
<td></td>
<td>.98</td>
</tr>
<tr>
<td>Medium</td>
<td>299</td>
<td>315</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olive/brown</td>
<td>88</td>
<td>91</td>
<td>7</td>
<td></td>
<td>.67</td>
</tr>
<tr>
<td>Skin reaction to acute sun</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burn, never tan</td>
<td>171</td>
<td>169</td>
<td>21</td>
<td></td>
<td>.94</td>
</tr>
<tr>
<td>Burn, then tan</td>
<td>552</td>
<td>547</td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only tan</td>
<td>88</td>
<td>92</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous occupations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mainly outdoors</td>
<td>165</td>
<td>138</td>
<td>20</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Indoors and outdoors</td>
<td>283</td>
<td>318</td>
<td>35</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>Mainly indoors</td>
<td>363</td>
<td>352</td>
<td>48</td>
<td>44</td>
<td>.10</td>
</tr>
<tr>
<td>No. of sunburns</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nona</td>
<td>97</td>
<td>88</td>
<td>12</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Once</td>
<td>131</td>
<td>144</td>
<td>16</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>2-6</td>
<td>362</td>
<td>354</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 5</td>
<td>220</td>
<td>222</td>
<td>27</td>
<td></td>
<td>.77</td>
</tr>
<tr>
<td>New on back</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>134</td>
<td>127</td>
<td>17</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1-10</td>
<td>537</td>
<td>526</td>
<td>68</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>≥ 11</td>
<td>123</td>
<td>135</td>
<td>15</td>
<td>17</td>
<td>.66</td>
</tr>
<tr>
<td>Previous history skin cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>207</td>
<td>211</td>
<td>25</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>805</td>
<td>598</td>
<td>75</td>
<td>74</td>
<td>.79</td>
</tr>
</tbody>
</table>
Results

<table>
<thead>
<tr>
<th>Melanoma by Level</th>
<th>No. of Participants Affected</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sunscreen (n = 812)</td>
<td>No Sunscreen (n = 809)</td>
</tr>
<tr>
<td>All</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>I: in situ</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Invasive</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>II: in papillary dermis</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>III: filling papillary dermis</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>IV: reticular dermis</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>
Results: Kaplan-Meier curve
Bottom Line

✓ RCTs are more *logistically difficult*, more *expensive*, and have more issues related to *ethical considerations* than any other epidemiologic design strategy.

✓ But if ethically appropriate, and if well designed and conducted, they provide a *level of assurance* about the effect of the intervention itself on the outcome that cannot be achieved by any other epidemiologic design strategy.
Limitations to the gold standard

1. Limited follow-up
2. Unrepresentative treatments
3. Unrepresentative patients
Metformin and CRC Survival:
A Multi-Pronged Approach

• *In vitro* and *in vivo* studies suggests that metformin may have anti-cancer activity
• Blood insulin and glucose, cancer cell proliferation and apoptosis, and cancer stem cell growth
• Imagine that to date there have been no human studies
• How would you test the potential anticancer effect of metformin on CRC survival?
Hierarchy of Study Designs?

Meta-analysis/Evidence Synthesis
↓
Randomized Controlled Trials
↓
Cohort studies  Case-control studies
↓
Cross-sectional studies
↓
Ecologic studies
↓
Case reports

Tufts CTSI
Metformin and CRC Survival

• What kind of database do you need?
  • Assessment of T2DM status and therapies used
  • Ascertainment of CRC and CRC outcomes
  • Sufficient numbers of patients with both T2DM and CRC, and metformin treated vs. untreated
  • Rich information on confounding variables (DM severity, CRC stage and treatment, PS)
  • Adequate follow-up time
The Search for a Database: Primary Data

Primary data is collected by the investigator directly from study participants to address a specific question or hypothesis

- Prospective observational studies
  - Subjects are selected on the basis of specific characteristics, and their progress is monitored.

- Registries
  - Registries use an observational study design to collect data and do not specify treatments or require therapies intended to change patient outcomes.
  - Used for public health surveillance, to generate descriptive statistics (incidence and mortality rates), risk assessment

- Repurposed trial data or data from completed observational studies
The Search for a Database: Secondary Data

Secondary data is data collected for other purposes that can be used to answer the research question.

- Electronic medical record data
- Administrative data
  - Typically generated as part of the process of obtaining insurance reimbursement
- Pharmacy data
  - Claims submitted to insurance companies for payments, as well as pharmacy dispensing records
- Regulatory data
  - FDA has data from regulatory approval submissions.
Metformin and CRC Survival

1. Women’s Health Initiative
   • 2,066 postmenopausal women with CRC (1850 without DM, 85 DM and metformin, 125 DM and no metformin)
   • National health study focused on heart disease, breast ca, CRC, and osteoporotic fractures in postmenopausal women

2. VA Central Cancer Registry and VA Corporate Data Warehouse
   • 21,300 patients diagnosed with CRC (16500 without DM, 2000 DM and metformin, 2100 DM and other DM drug, 800 DM and no anti-DM drug)
     • Registry (patient demographics, tumor characteristics, and primary treatment) linked to CDW (pharmacy, diagnostic, lab, and vital status data)
Diabetes, Metformin Use, and Colorectal Cancer Survival in Postmenopausal Women

Furha Iram Cossor\textsuperscript{a}, Lucile L. Adams-Campbell\textsuperscript{b}, Rowan T. Chlebowski\textsuperscript{c}, Marc J Gunter\textsuperscript{d}, Karen Johnson\textsuperscript{e}, Robert E. Martelli\textsuperscript{a}, Anne McTiernan\textsuperscript{f}, Michael S. Simon\textsuperscript{g}, Thomas Rohan\textsuperscript{h}, Robert B. Wallace\textsuperscript{i}, and Jessica K. Paulus\textsuperscript{j}
### A. Metformin and Survival in Diabetics with CRC (N=212)

<table>
<thead>
<tr>
<th></th>
<th>Unadjusted HR (95% CI)</th>
<th>Age-adjusted HR (95% CI)</th>
<th>Multivariate-Adjusted HR (95% CI)</th>
<th>Propensity Score-adjusted HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Colorectal Cancer Specific Survival</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Metformin</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Metformin</td>
<td>0.77 (0.42 – 1.39)</td>
<td>0.76 (0.41 – 1.41)</td>
<td>0.75 (0.40 – 1.38)</td>
<td>0.78 (0.38 – 1.55)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.38</td>
<td>0.39</td>
<td>0.67</td>
<td>0.47</td>
</tr>
<tr>
<td><strong>Overall Survival</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Metformin</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Metformin</td>
<td>0.82 (0.50 – 1.33)</td>
<td>0.85 (0.52 – 1.30)</td>
<td>0.84 (0.51 – 1.37)</td>
<td>0.86 (0.49 – 1.52)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.41</td>
<td>0.52</td>
<td>0.48</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Metformin, Diabetes, and Survival among U.S. Veterans with Colorectal Cancer

Jessica K. Paulus¹, Christina D. Williams²,³, Furha I. Cossor⁴, Michael J. Kelley²,³, and Robert E. Martell⁵
<table>
<thead>
<tr>
<th>Diabetes therapy</th>
<th>Unadj HR (95% CI)</th>
<th>P</th>
<th>AHR&lt;sup&gt;b&lt;/sup&gt; (95% CI)</th>
<th>P&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Metformin</td>
<td>1.00</td>
<td>Ref</td>
<td>1.00</td>
<td>Ref</td>
</tr>
<tr>
<td>Metformin</td>
<td>0.70 (0.65–0.76)</td>
<td>&lt;0.0001</td>
<td>0.87 (0.79–0.95)</td>
<td>0.003</td>
</tr>
<tr>
<td>None</td>
<td>0.94 (0.85–1.03)</td>
<td>0.17</td>
<td>1.02 (0.90–1.15)</td>
<td>0.76</td>
</tr>
</tbody>
</table>
Prospective evaluation of clinical safety of combining metformin with anticancer chemotherapy

PROTOCOL VERSION DATE

Version 3.0
November 3, 2015
Amendment 2

PI: Wasif M. Saif, MD
Professor of Medicine, Tufts School of Medicine
Program leader, Exp. Therapeutic, Tufts Medical Center – Tufts Cancer Center
800 Washington Street
7-South, Suit: 7099
Boston, MA 02111
Ph: 617-636-5627
Fax: 617-636-8538
Email: wsaif@tuftsmedicalcenter.org
A Delayed Start Phase 1 Randomized Trial (n=100)

Study goal: to obtain prospective safety and pharmacodynamic information in cancer patients

Run-in

- PD
- glucose

Documented cancer; Intended chemotherapy; No recent metformin

Chemotherapy

No DLT

DLT

If subject experiences DLT, chemotherapy alone should be continued in order to identify a tolerable dose/regimen if considered medically appropriate.

Stage 1

- PD
- glucose

Chemotherapy + Metformin 500mg bid

Stage 2

Chemotherapy + Metformin 500mg bid for up to 4 months

Metformin DLT

Chemotherapy
Thank you!

Jess Paulus, ScD
jpaulus@tuftsmedicalcenter.org
EXTRA SLIDES
Comparing RCTs and Observational Studies

RANDOMIZED, CONTROLLED TRIALS, OBSERVATIONAL STUDIES, AND THE HIERARCHY OF RESEARCH DESIGNS

John Concato, M.D., M.P.H., Nirav Shah, M.D., M.P.H., and Ralph I. Horwitz, M.D.

ABSTRACT

Background In the hierarchy of research designs, the results of randomized, controlled trials are considered to be evidence of the highest grade, whereas observational studies are viewed as having less validity because they reportedly overestimate treatment effects. We used published meta-analyses to identify randomized clinical trials and observational studies that examined the same clinical topics. We then compared the results of the original reports according to the type of research design.

Methods A search of the Medline data base for articles published in five major medical journals from 1981 to 1995 identified meta-analyses of randomized, controlled trials and meta-analyses of either cohort or case-control studies that assessed the same intervention. For each of five topics, summary estimates and 95 percent confidence intervals were calculated on the basis of data from the individual randomized, controlled trials and the individual observational studies.

The literature identified six different therapies evaluated in both randomized, controlled trials (50 studies) and trials with historical controls (56 studies). For each study, subjects in the treatment group were found to have similar rates of the outcome in question regardless of study design, but subjects in the control group in trials with historical controls had worse outcomes than control subjects in randomized, controlled trials. The agent being tested was considered effective in 44 of 56 trials with historical controls (79 percent), but in only 10 of 50 randomized, controlled trials (20 percent). The authors concluded that biases in patient selection may irrevocably weight the outcome of historical controlled trials in favor of new therapies.

Current criticisms of observational studies involve, in addition to trials with historical controls, cohort studies with concurrent selection of control subjects, as well as case-control designs. Advocates of "evidence-based medicine" classify studies according to

Systematic Review of RCTs and Obs. Studies

• Identified meta-analyses of RCTs or observational studies studying the same clinical question

• 5 research topics, 1991-1995

• Compared summary RR’s by research design
<table>
<thead>
<tr>
<th>Clinical Topic</th>
<th>Type of Study</th>
<th>Meta-Analysis*</th>
<th>Total No. of Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacille Calmette–Guérin vaccine and tuberculosis</td>
<td>Randomized, controlled</td>
<td>Colditz et al.</td>
<td>359,922</td>
</tr>
<tr>
<td>Mammography and mortality from breast cancer</td>
<td>Case–control</td>
<td>Colditz et al.</td>
<td>6,511</td>
</tr>
<tr>
<td>Cholesterol levels and death due to trauma</td>
<td>Randomized, controlled</td>
<td>Kerlikowske et al.</td>
<td>429,043</td>
</tr>
<tr>
<td></td>
<td>Case–control</td>
<td>Kerlikowske et al.</td>
<td>132,456</td>
</tr>
<tr>
<td></td>
<td>Cohort</td>
<td>Cummings and Psaty</td>
<td>36,910</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Jacobs et al.</td>
<td>9,377</td>
</tr>
<tr>
<td>Treatment of hypertension and stroke</td>
<td>Randomized, controlled</td>
<td>Collins et al.</td>
<td>36,894</td>
</tr>
<tr>
<td>Treatment of hypertension and coronary heart disease</td>
<td>Cohort</td>
<td>MacMahon et al.</td>
<td>405,511</td>
</tr>
<tr>
<td></td>
<td>Randomized, controlled</td>
<td>Collins et al.</td>
<td>36,894</td>
</tr>
<tr>
<td></td>
<td>Cohort</td>
<td>MacMahon et al.</td>
<td>418,343</td>
</tr>
</tbody>
</table>
Do the black circles represent:
A. Observational studies
B. Randomized trials
<table>
<thead>
<tr>
<th>Clinical Topic</th>
<th>Type of Study</th>
<th>Meta-Analysis*</th>
<th>Total No. of Subjects</th>
<th>Summary Estimate (95% CI)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacille Calmette–Guérin vaccine and tuberculosis</td>
<td>13 Randomized, controlled</td>
<td>Colditz et al.¹⁴</td>
<td>359,922</td>
<td>0.49 (0.34–0.70)</td>
</tr>
<tr>
<td></td>
<td>10 Case–control</td>
<td>Colditz et al.¹⁴</td>
<td>6,511</td>
<td>0.50 (0.39–0.65)</td>
</tr>
<tr>
<td>Mammography and mortality from breast cancer</td>
<td>8 Randomized, controlled</td>
<td>Kerlikowske et al.¹⁵</td>
<td>429,043</td>
<td>0.79 (0.71–0.88)</td>
</tr>
<tr>
<td></td>
<td>4 Case–control</td>
<td>Kerlikowske et al.¹⁵</td>
<td>132,456</td>
<td>0.61 (0.49–0.77)</td>
</tr>
<tr>
<td>Cholesterol levels and death due to trauma</td>
<td>6 Randomized, controlled</td>
<td>Cummings and Psaty¹⁶</td>
<td>36,910</td>
<td>1.42 (0.94–2.15)</td>
</tr>
<tr>
<td></td>
<td>14 Cohort</td>
<td>Jacobs et al.¹⁷</td>
<td>9,377</td>
<td>1.40 (1.14–1.66)</td>
</tr>
<tr>
<td>Treatment of hypertension and stroke</td>
<td>14 Randomized, controlled</td>
<td>Collins et al.¹⁸</td>
<td>36,894</td>
<td>0.58 (0.50–0.67)</td>
</tr>
<tr>
<td></td>
<td>7 Cohort</td>
<td>MacMahon et al.¹³</td>
<td>405,511</td>
<td>0.62 (0.60–0.65)</td>
</tr>
<tr>
<td>Treatment of hypertension and coronary heart disease</td>
<td>14 Randomized, controlled</td>
<td>Collins et al.¹⁸</td>
<td>36,894</td>
<td>0.86 (0.78–0.96)</td>
</tr>
<tr>
<td></td>
<td>9 Cohort</td>
<td>MacMahon et al.¹³</td>
<td>418,343</td>
<td>0.77 (0.75–0.80)</td>
</tr>
</tbody>
</table>
Concato’s Refutation of the Hierarchy

1. RCT results are substantially varied and contradictory

2. Observational studies yield similar results as RCTs as long as study populations and questions are similar

3. Non-medical scientific (psych, educational, behavioral) disciplines do not support a hierarchy of research designs
Thank You
Sample Size and Power I: Binary Outcomes

Farzad Noubary, PhD

Assistant Professor
Tufts University School of Medicine
Sample Size and Power

Principles:

Sample size calculations are an essential part of study design.

Consider sample size requirements early.

A well-designed trial is large enough to detect clinically important differences between groups with high probability.

To perform sample size calculations, we need well defined study endpoints, hypotheses, and statistical tests.
Specify the Null Hypothesis

Study hypotheses should be based on a clearly defined endpoint and period of study.

In most RCTs, known as \textit{superiority trials}, the study hypothesis is stated as a null hypothesis of no difference in the distribution of the primary endpoint between study groups.

In the CORONARY Trial, the short-term null hypothesis was

\[ H_0: \text{Patients receiving on-pump and off-pump coronary artery bypass surgery will have identical event rates at 30 days post-randomization} \]
Specify the Alternative Hypothesis

We have an alternative hypothesis in mind, for example,

\[ H_A: \text{ The frequency of events at 30 days will differ in the two treatment groups.} \]

In superiority trials, we test the null hypothesis against a two-sided alternative.

We have a directional alternative hypothesis in mind, for example, that fewer events will occur within 30 days in the off-pump group.
Outcomes of Hypothesis Testing

When we test the null hypothesis, there are two possible states of nature and two decisions:

<table>
<thead>
<tr>
<th>Test Result</th>
<th>$H_0$ True</th>
<th>$H_a$ True</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td>Type I Error</td>
<td>No Error</td>
</tr>
<tr>
<td>Do Not Rej $H_0$</td>
<td>No Error</td>
<td>Type II Error</td>
</tr>
</tbody>
</table>
Power

We will perform a test that has a small probability of a Type 1 error, usually 0.05.

The power of the study is the probability that we will reject the null hypothesis when the alternative hypothesis is actually true.

We would like this probability to be large, typically at least 0.8.
Express the Hypotheses in Terms of Probabilities

The 30-day outcome is a binary event, occurrence or non-occurrence of death or complications within 30 days of surgery.

\[ p_T = \text{Probability that an off-pump patient will have an event} \]
\[ p_C = \text{Probability that an on-pump patient will have an event} \]

The study hypotheses are:

\[ H_0: \ p_T = p_C \ (T \text{ and } C \text{ are equally effective}) \]
\[ H_A: \ p_T \neq p_C \ (T \text{ and } C \text{ are not equally effective}) \]

We specify the direction of the alternative for the sample size calculation.
The Test Statistic

To test the null hypothesis, we calculate a test statistic, $T$, and a critical value, $C$, and reject the null hypothesis if $|T| > C$, that is, if $T > C$ or $T < -C$.

To calculate power or sample size, we will focus on significance in one direction, $T < -C$, implying that $p_T < p_C$.

For the CORONARY Trial, define $T$ as the difference between the observed proportions divided by the standard error of the difference.
The Test Statistic

The observed difference in proportions is \[ D = \hat{p}_T - \hat{p}_C \]

Assuming equal sample sizes in the two groups,

\[ \text{Var}(D) = \frac{p_T(1-p_T)}{n} + \frac{p_C(1-p_C)}{n} \]

Under the null hypothesis, \( p_T = p_C \). Define the test statistic as \( D \) divided by its standard deviation

\[ T = \frac{\hat{p}_T - \hat{p}_C}{\sqrt{2\bar{p}(1-\bar{p})/n}} = \frac{D}{SD(D)} \]

where \( \bar{p} \) is the average event rate
Choosing C

Choose C so that $P(T < -C \mid H_0) = \alpha/2$. Usually, $\alpha = 0.05$ (two-sided) so $\alpha/2 = 0.025$.

$T$ is approximately $N(0,1)$ if $H_0$ is true. Hence, if $\alpha/2 = 0.025, C = 1.96$.

Power is $P(T < -C \mid H_A) = 1 - P$(Type 2 error) = $1 - \beta$.

The investigator can control the power by choosing the sample size
Null and Alternative Hypothesis

In the CORONARY Trial, one possible scenario for the 30-day endpoint was

\[ p_C = 0.08, \text{ and, under the alternative hypothesis} \]
\[ p_T = (0.85) \times 0.08 = 0.068 \]

a 15% reduction in the event rate in the off-pump group. Under \( H_A \), the expected value of \( D \) would be

\[ \Delta = 0.068 - 0.08 = -0.012 \]

If \( H_0 \) is true, \( \text{var}(D) = \frac{2 \times 0.08 \times 0.92}{n} \)
Logic of Sample Size Calculations

Again consider the risk difference, \( D = \hat{p}_T - \hat{p}_C \)

\( D \) is approximately normally distributed with
- mean = 0 if \( H_0 \) is true and
- mean = -.012 if \( H_A \) is true

\[ \text{Var}(D) = p_T(1-p_T)/n_T + p_C(1-p_C)/n_C \]

The mean is independent of \( n \) but the variance decreases as \( n \) increases
The Logic of Hypothesis Testing

The distribution of \( D \) under the null and alternative hypotheses.

*pdf: probability density function
Type 1 Error

PDF

\( H_A \)

\( H_0 \)

\( \alpha/2 \)

\( D \)
Power

PDF

$1 - \beta$

$H_A$

$H_0$

$D$

-0.084 -0.042 0.0 0.042 0.084
Var($D$) Depends on Sample Size

\begin{align*}
\text{n} &= 500 & \text{Var} &= 0.00029 & \text{SD} &= 0.017 \\
\text{n} &= 1,000 & \text{Var} &= 0.00015 & \text{SD} &= 0.012 \\
\text{n} &= 2,000 & \text{Var} &= 0.000074 & \text{SD} &= 0.0086
\end{align*}
$n = 500$

$D$

PDF

$H_a$

$H_0$

Tufts CTSI

Tufts Clinical and Translational Science Institute
$n = 1,000$
$n=2,000$

PDF

$D$

$H_A$

$H_0$
For a fixed sample size, the power of the study will increase with the size of the true difference.
Sample Size Formula

To achieve the desired Type 1 and Type 2 error, we need to satisfy two conditions

\[-Z_{\alpha/2} \cdot SD(D) = -C \text{ and } \Delta + Z_{\beta} \cdot SD(D) = -C\]

Recall that we estimate the variance of D by

\[2\tilde{p} \cdot (1 - \tilde{p})/n\]

To determine n, set \(-Z_{\alpha/2} \cdot SD(D) = \Delta + Z_{\beta} \cdot SD(D)\) and solve for n.
PDF

n=1000

-H_A - H_0

-Z_{\alpha/2}SD(D) = -C
PDF

\[ \Delta + Z_0 \sigma(D) = -C \]
The Sample Size Formula

\[ n = \frac{2\bar{p} \times (1 - \bar{p}) (Z_{\alpha/2} + Z_{\beta})^2}{\Delta^2} \]

\( Z_{\alpha/2} \) and \( Z_{\beta} \) are the critical values of the normal distribution, \( \bar{p} \) is the average of the event rates under the alternative hypothesis, and \( \Delta \) is the true difference under \( H_A \). For the CORONARY Trial, with

\[ \alpha = 0.05, \beta = 0.20, n = 1,903 \text{ or } 2n = 3,806. \]

(Note: If \( p_c = 0.08 \), and \( p_t = 0.068 \) if \( H_A \) is true, the correct value for the sample size is \( n = 7,462. \))

The CORONARY investigators considered a range of scenarios and settled on a total sample size of 4,700 patients.
Thank You
Statistical Analysis: Monitoring & Interim Analyses

Angie Mae Rodday, PhD, MS
Assistant Professor
Institute for Clinical Research and Health Policy Studies
Outline

• Reasons for monitoring clinical trials
• Methods for interim analyses
  – Stopping rules
• Examples of interim analyses from the literature
Monitoring & Interim Analyses

• A trial should only be continued if
  – Remains ethical to randomly assign the study treatments
  – Potential to answer the proposed research question

• Ongoing clinical trials must be monitored to assess:
  – Ethics (risks and benefits)
  – Data quality
  – Precision of results
  – Treatment effects and side effects
  – Resource availability
  – Outside information

• Accomplished by interim analyses throughout the trial
Treatment Effects and Side Effects

- **Safety:** Unacceptable side effects or toxicity
- **Efficacy:** Experimental treatment is convincingly superior (or non-inferior) to the control
- **Harm:** Experimental treatment is convincingly worse than the control
- **Futility:** Experimental treatment is convincingly not superior (or inferior) to the control
Repeated Testing

• When conducting interim analyses, multiple comparisons are conducted
• If $H_0$ is true, but is tested repeatedly at the $\alpha$ level, the probability of one or more statistically significant tests will exceed $\alpha$ even if $H_0$ is true
• Must adjust for multiple comparisons
Methods for Interim Analyses

- Group Sequential Designs
  - Stopping boundaries
  - Alpha spending function
- Conditional power (futility)
- Continuous toxicity monitoring (safety)
- Bayesian approaches
Group Sequential Design

- Conduct interim analysis after certain number of patients have reached endpoint
- Divide total sample size into $K$ groups of equal size
- Analyze data after results from each group have been collected and compare to stopping boundaries
  - Continue or stop study based on stopping boundaries

![Diagram showing decision pattern in a sequential trial](image)

*Fig. 1 The pattern of decisions in a sequential trial*

Lew 2015
Stopping Rules

• Perform each test with symmetric boundaries at same $\alpha$ level (Pocock)

• Perform each test with symmetric boundaries at an $\alpha$ that changes over time. $\alpha$ starts small and preserves most of $\alpha$ for the final analysis (O’Brien-Fleming)

• $\alpha=0.001$ until the last analysis, then $\alpha=0.05$ (Haybittle-Peto)
Stopping Rules for Group Sequential Design

Fig. 15-4 Three group sequential stopping boundaries for the standardized normal statistic (Z) for up to five sequential groups with two-sided significance level of 0.05.

Friedman 1998
Alpha Spending Designs

• Stopping rules can be modified to allow interim analyses at uneven intervals
  – Number or timing of interim analyses need not be specified in advance, but spending function does

• Describes rate at which total alpha is spent as a continuous function of the information fraction
  – Information fraction for survival=number of observed deaths/expected number of deaths
  – Information fraction for comparison of means=number of patients observed/target sample size

DeMets 1994
Stopping for Futility

- What happens when the interim result is unlikely to change after accruing more subjects?
- Assessing for futility
  - Group sequential methods
  - Conditional power
    - Calculate the power of the study to reject the null given the current results
Stopping Rules for Efficacy, Harm, Futility

DeMets 2006

Figure 3 Vesnarinone trial (VEST) group sequential bounds

Group Sequential Bounds
\(1\alpha = 0.025\)

Information Fraction

OBF = O’Brien–Fleming Bounds [12].
E-F = Emerson–Fleming Bounds [26].
Conditional Power

Figure 5  Conditional power boundaries: outer boundaries represent symmetric O'Brien-Fleming type sequential boundaries ($\alpha = 0.05$). Three lower boundaries represent boundaries for 10%, 20% and 30% conditional power to achieve a significant ($P < 0.05$) result of the trial conclusion.
Stopping for Safety

• What happens when the interim results indicate unacceptable side effects or toxicity?

• Assessing for safety
  – Group sequential design
  – Continuous toxicity monitoring
    • Patients enrolled one at a time
    • If toxicity is acceptable, new patient enrolled. If not, study is stopped.
Continuous Toxicity Monitoring

Figure 3. Adjusted Pocock (dashed line) and O’Brien-Fleming boundaries (solid line) with $K = 30$, $H_0: \theta = 0.2$, $\alpha \leq 0.05$. The vertical axis is the number of patients experiencing toxicity.

Nguyen 2009
Examples of Interim Analyses

• What was the purpose of the interim analysis?
  – Safety? Efficacy? Harm? Futility?
• What type of interim analysis?
  – Group sequential? Conditional power?
• What stopping rules were used?
• What were the results of the interim analysis?
Example 1: Stupp 2016

- **Objective:** To evaluate the efficacy and safety of Tumor-Treating Fields (TTFields) used in combination with temozolomide maintenance treatment after chemoradiation therapy for patients with glioblastoma.

- **Statistics:** This pre-specified interim analysis was to be performed after the first 315 randomized patients reached a minimum 18-month follow-up. The final type I error rate of 0.05 was split between the interim and final analyses based on a standard $\alpha$ spending function.
Results

Data for the interim analysis included 210 patients randomized to TTFields plus temozolomide and 105 to temozolomide alone. The independent data and safety monitoring committee met in October 2014 to review the interim analysis; the trial met the predefined boundaries for success (improvement of both progression-free and overall survival) and the committee recommended study termination, thus allowing patients in the control group to crossover and receive TTFields.
Example 2: WHI 2002

- **OBJECTIVE:** To assess the major health benefits and risks of the most commonly used combined hormone preparation in the United States.

- **MAIN OUTCOMES MEASURES:** The primary outcome was coronary heart disease (CHD), with invasive breast cancer as the primary adverse outcome. A global index summarizing the balance of risks and benefits included the 2 primary outcomes plus stroke, pulmonary embolism (PE), endometrial cancer, colorectal cancer, hip fracture, and death due to other causes.
Monitoring Methods

Trial monitoring guidelines for early stopping considerations were based on O'Brien-Fleming boundaries using asymmetric upper and lower boundaries: a 1-sided, .025-level upper boundary for benefit and 1-sided, .05-level lower boundaries for adverse effects. Trial monitoring for early stopping considerations was conducted semiannually by an independent DSMB.
Monitoring & Early Stopping

Formal monitoring began in the fall of 1997 with the expectation of final analysis in 2005 after an average of approximately 8.5 years of follow-up. Late in 1999, with 5 interim analyses completed, the DSMB observed small but consistent early adverse effects in cardiovascular outcomes and in the global index. None of the disease-specific boundaries had been crossed. In the spring of 2000 and again in the spring of 2001, at the direction of the DSMB, hormone trial participants were given information indicating that increases in MI, stroke, and PE/DVT had been observed and that the trial continued because the balance of risks and benefits remained uncertain.
Monitoring & Early Stopping

In reviewing the data for the 10th interim analyses, the DSMB found that the adverse effects in cardiovascular diseases persisted, although these results were still within the monitoring boundaries. However, the test statistic for breast cancer ($z = -3.19$) crossed the designated boundary ($z = -2.32$) and the global index was supportive of a finding of overall harm ($z = -1.62$). Updated analyses including 2 months of additional data, available by the time of the meeting, did not appreciably change the overall results. On the basis of these data, the DSMB concluded that the evidence for breast cancer harm, along with evidence for some increase in CHD, stroke, and PE, outweighed the evidence of benefit for fractures and possible benefit for colon cancer over the average 5.2-year follow-up period. Therefore, the DSMB recommended early stopping of the estrogen plus progestin component of the trial. Because the balance of risks and benefits in the unopposed-estrogen component remains uncertain, the DSMB recommended continuation of that component of the WHI. Individual trial participants have been informed.
Example 3: Piperno-Neumann 2016

- **OBJECTIVE:** We assessed whether zoledronate combined with chemotherapy and surgery improved event-free survival in children and adults with osteosarcoma.

- **STATISTICAL ANALYSIS:** Three interim analyses were done with a Lan and DeMets $\alpha$-spending function based on the O’Brien-Fleming group sequential boundary function. These analyses were to be disclosed to the IDMC… For each toxicity term, the proportion of patients who had a severe toxicity was compared between randomised groups using a $\chi^2$ test.
Results

• After the second interim analysis, accrual was prematurely stopped for futility because the estimated likelihood of showing an event-free survival benefit if the trial had continued was practically null (conditional power to show a benefit <0.0001).

• No significant increase in acute toxicity was noted during treatment in the zoledronate group than in the control group, except for a large excess of hypocalcaemia and hypophosphataemia
Discussion

• The toxicity noted with these chemotherapy regimens was as expected, with no major increase of toxicity in patients receiving zoledronate, except for reversible hypocalcaemia and hypophosphataemia.

• The conditional power was practically null even under our optimistic hypothesis that the 3-year event-free survival would increase from 55% to 68%.
Do studies that stop early for benefit have different results?
Systematic review & meta-analysis: Bassler 2010

• **Objective:** Compare the treatment effect from truncated RCTs with that from meta-analyses of RCTs addressing the same question but not stopped early (non-truncated RCTs)

• **Methods:** Selected studies where RCTs stopped early for benefit and matching non-truncated RCTs answered similar research questions
Figure 2. Pooled Ratio of Relative Risks (RRs) and 95% Confidence Intervals (CIs) for Truncated vs Nontruncated Randomized Controlled Trials (RCTs)

<table>
<thead>
<tr>
<th>Matching Question No.</th>
<th>Truncated RCTs</th>
<th>Nontruncated RCTs</th>
<th>Ratio of RRs (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Events</td>
<td>Sample Size</td>
<td>RR (95% CI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>15</td>
<td>0.36 (0.14-0.94)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1477</td>
<td>0.79 (0.72-0.88)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>158</td>
<td>0.46 (0.36-0.60)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2122</td>
<td>0.72 (0.66-0.78)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>162</td>
<td>0.47 (0.22-0.79)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>152</td>
<td>0.62 (0.46-0.78)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>152</td>
<td>0.63 (0.50-0.80)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>49</td>
<td>0.27 (0.14-0.54)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>41</td>
<td>0.25 (0.12-0.68)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>183</td>
<td>0.70 (0.64-0.82)</td>
<td></td>
</tr>
</tbody>
</table>

|                       | Total Events  | Sample Size      | RR (95% CI)           | P Value |
|                       | No.           |                  |                       |         |
| 11                    | 1             | 456              | 0.39 (0.34-0.46)      |         |
| 12                    | 1             | 471              | 0.70 (0.60-0.82)      |         |
| 13                    | 1             | 43               | 0.33 (0.18-0.64)      |         |
| 14                    | 1             | 32               | 0.02 (0.01-0.13)      |         |
| 15                    | 3             | 1558             | 0.90 (0.82-1.00)      |         |
| 16                    | 3             | 120              | 0.50 (0.30-0.88)      |         |
| 17                    | 1             | 5                | 0.42 (0.12-1.64)      |         |
| 18                    | 1             | 9                | 0.13 (0.02-0.82)      |         |
| 19                    | 1             | 26               | 0.16 (0.06-0.52)      |         |
| 20                    | 1             | 23               | 0.62 (0.42-0.94)      |         |
| 21                    | 2             | 40               | 0.61 (0.46-0.84)      |         |
| 22                    | 2             | 47               | 0.47 (0.30-0.78)      |         |
| 23                    | 1             | 12               | 0.88 (0.16-0.98)      |         |
| 24                    | 1             | 67               | 0.43 (0.30-0.66)      |         |

Random effects: P<.001 for heterogeneity, I² = 57%.
Test for overall effect: z = 9.55 (P<.001).

First column indicates number associated with the question addressed by each review that included 1 or more truncated and matching nontruncated RCTs. Results ordered by P values associated with results of nontruncated RCTs; size of the data markers indicates weight of review questions in meta-analysis.
References

• DeMets et al. The agonising negative trend in monitoring of clinical trials. The Lancet, 1999;354.
• DeMets. Futility approaches to interim monitoring by data monitoring committees. Clinical Trials, 2006;3.
• Stupp et al. Maintenance Therapy With Tumor-Treating Fields Plus Temozolomide vs Temozolomide Alone for Glioblastoma: A Randomized Clinical Trial. JAMA 2015;314(23)
• WHI. Risk and benefits of estrogen plus progestin in healthy postmenopausal women: Principal results from the Women’s Health Initiative Randomized Trial. JAMA 2002;288(3).
Thank You
Overview of Next Session

Andreas Klein, MD
<table>
<thead>
<tr>
<th>Agenda Item</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration</td>
<td>12:30 – 1:00 PM</td>
</tr>
<tr>
<td>Institutional Regulatory Review</td>
<td>1:00 – 2:00 PM</td>
</tr>
<tr>
<td>Andreas Klein &amp; Jack Erban</td>
<td></td>
</tr>
<tr>
<td>Investigator Responsibilities</td>
<td>2:00 – 2:45 PM</td>
</tr>
<tr>
<td>Susan Parsons</td>
<td></td>
</tr>
<tr>
<td>Break &amp; Snack</td>
<td>2:45 – 3:00 PM</td>
</tr>
<tr>
<td>Biorepository: Use of tissue specimens for research</td>
<td>3:00 – 3:15 PM</td>
</tr>
<tr>
<td>Sandra Gaston</td>
<td></td>
</tr>
<tr>
<td>Data and Safety Monitoring Board (DSMB)</td>
<td>3:15 – 4:00 PM</td>
</tr>
<tr>
<td>Tamsin Knox</td>
<td></td>
</tr>
<tr>
<td>Study Budgets</td>
<td>4:00 – 4:30 PM</td>
</tr>
<tr>
<td>Doug Reichgott</td>
<td></td>
</tr>
<tr>
<td>Closing Remarks/Happy Holidays!</td>
<td>4:45 – 5:00 PM</td>
</tr>
</tbody>
</table>
Pre-Work

III. Clinical Trials (December 16, 2016)

i. Principal Investigator IRB Responsibilities (Under Regulatory Affairs course)

ii. Working with the IRB: Common Myths and Successful Strategies Best Practices in Clinical Trial (Under Regulatory Affairs course)

iii. Data Safety and Monitoring Boards (Under Regulatory Affairs course)

iv. What you need to know about DSMBs (But were afraid to ask), (Under Regulatory Affairs course)

v. Research Ethics of Clinical Investigation (Under Clinical Research course)

vi. Research Data Management (Under Clinical Research course)