Clinical Evidence: Asking the Question and Understanding the Answer

Keeping Up to Date

5,000? per day
1,500 per day
95 per day

Keeping Up to Date

5,000? per day
1,500 per day
Bias

- Bias refers to any systematic error in the design, conduct, analysis, publication, or review of a study leading to results that are misleading and conclusions that are wrong
- Interventions and treatments may appear more promising or less beneficial than they actually are
- Bias is used to describe an error that was not due to chance, and therefore it cannot be measured or controlled for statistically
- Researcher will try to reduce bias as much as possible

Confounding

- Confounding literally means “confusion of effects”
- Confounding is a problem if a variable is different between the two groups and affects outcomes being studied (e.g., coffee and smoking and CVD etiology)
- Confounders cannot be removed from a study but can be accounted for in statistical analysis

How Do We Assess the Information?
Evidence-Based Medicine (EBM) or Practice (EBP)

EBM (or EBP) describes the process of practicing medicine based on the combination of best research evidence with clinical expertise and patient values.

Evidence-Based Medicine – The Five Steps

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Question</td>
<td>Formulate an answerable question: “PICO” – Patient or Problem; Intervention; Comparison; Outcomes</td>
</tr>
<tr>
<td>2 Evidence</td>
<td>Search for the best evidence</td>
</tr>
<tr>
<td>3 Critical Appraisal</td>
<td>Evaluate the evidence – relevance, validity, impact (size of the benefit), generalizability</td>
</tr>
<tr>
<td>4 Application</td>
<td>Apply the results, combining with clinical expertise and patient circumstances</td>
</tr>
<tr>
<td>5 Implementation and Monitoring</td>
<td>Implement and monitor the process</td>
</tr>
</tbody>
</table>

Evidence-Based Medicine – Critical Appraisal

Why do we need critical appraisal?

- Could be conflicting results from studies
- Real-life medicine does not reflect restrictive environment of clinical studies
  - Efficacy describes the impact of interventions under optimal (trial) conditions
  - Effectiveness describes the impact of interventions under ordinary (clinical) conditions
- Critical appraisal seeks to answer:
  - Does the research have internal validity, i.e., reflect true results based on study design and methodology?
  - Does the research have external validity, i.e., to what extent can results be generalized to a wider population?
Evidence-Based Medicine –
Hierarchy of Evidence

Hierarchy of research methodologies is based on premise that different study* designs differ in ability to predict what will happen to patients in real life

- Systematic reviews/meta-analyses/RCTs
- Non-RCTs
- Cohort studies
- Case-controlled studies
- Case series and registries
- Case reports

*NB: Not all studies using same design are of equal quality

Evidence-Based Medicine –
Network of Evidence

Observational studies:
- Provide useful evidence on the effectiveness of an intervention
- Larger and wider number of patients can be studied and therefore more generalizable
- However, may have doubt over the actual treatment effect size due to bias and confounding
- Can be especially useful for side-effects (including rare ones)

RCTs:
- Pinnacle of clinical trial evidence; bias and confounding minimized
- However, RCTs are limited by their relatively small size and narrow scope
- Therefore less generalizable
- Observational studies and RCTs can therefore be regarded as complementary approaches

Reality Check

We already have great medical information from clinical studies but…

- Inertia
  - Physician (e.g., COURAGE study)
  - Patient (e.g., insulin – why dialog important)
- Financial – who pays for new treatments?
  - Cost-effectiveness
  - Healthcare reform
- Guidelines outdated (e.g., ATP-III; JNC 7)
- External validity (e.g., stent studies)

We will never have great medical information from clinical studies…

- Alternative medicine
- Older therapies
- Physical therapies
Translational Research – 
From Bench to Bedside

Clinical trial is any form of planned experimental study designed (in general) to evaluate a new Rx on a clinical outcome in humans.

- Phase 0: Microdosing* (10 subjects), biomarkers and PK/PD
- Phase I: First in human → Safety (20–100 subjects), PK
- Phase II: First in patient → POC, dose (100s patients)
- Phase III: Efficacy, ADRs (1000s patients)
- Phase IV: Evaluation in the "real-world" clinical setting
  - Monitoring safety and efficacy in larger groups with broad subject eligibility criteria

* Cancer J. 2008;14:133-137.

Evaluating Clinical Studies – 
Essential Elements

- What is the clinical study question?
- Appropriate study design (including duration) and treatment structure
- Appropriate and well-defined population
- An adequate sample size based on a clinically important Δ between treatment groups
- Concurrent control group
- Randomized/blinded assignment of subjects
- Standardization/optimization of treatment regimens, measures, procedures
- Appropriate statistical methods, defined a priori
- Clinically meaningful and well-characterized end-points

What Is the Clinical Study Question?

- Each clinical trial requires a primary question
- Hypothesis should be stated
  - Usually same as primary question
- In addition, there are usually secondary questions
- All should be carefully selected, clearly defined, and stated in advance
Primary Question

- Should be the one the investigator is most interested in
- Should be possible to answer
  - Usually demonstrating a difference in a clinical or surrogate endpoint, e.g.,
    - Lives saved
    - Decreasing symptoms
    - Changing an important surrogate marker (e.g., BP, A1C, LDL)
- Will help define the sample size

The Clinical Question

Determines which study designs can be used to answer that question

- One clinical question can be answered by a number of study designs
- No single study design can answer all clinical questions

<table>
<thead>
<tr>
<th>Clinical Question</th>
<th>Clinical Relevance and Possible Study Designs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis</td>
<td>How valid and reliable is a diagnostic test?</td>
</tr>
<tr>
<td></td>
<td>Example: Case-control study</td>
</tr>
<tr>
<td>Etiology/Causation</td>
<td>What caused the disorder?</td>
</tr>
<tr>
<td></td>
<td>Example: RCT, case-control study, cohort study</td>
</tr>
<tr>
<td>Therapy</td>
<td>Which treatments do more good than harm?</td>
</tr>
<tr>
<td></td>
<td>Example: RCT, systematic review, meta-analysis</td>
</tr>
<tr>
<td>Prognosis</td>
<td>What is the likely course of a patient’s illness?</td>
</tr>
<tr>
<td></td>
<td>Example: Cohort study</td>
</tr>
</tbody>
</table>

Clinical Study Design – Type of Clinical Question Determines Type of Study

- Observational studies (descriptive)
  - Report what has been observed in a patient study sample (e.g., survey, case report, case series and registries)
  - Hypothesis-generating
- Observational studies (analytical)
  - Report the similarities and differences observed between two or more patient study samples (e.g., case-control study, cohort study, cross-sectional survey)
- Experimental studies
  - Researcher intervenes in some way with the experimental group and reports any differences between this group and a control group in which no intervention or a different intervention was offered (e.g., crossover trial, RCT)
  - Hypothesis testing
Observational Studies –
General Features
- Groups are already defined (e.g., smoker/nonsmoker); study merely observes what happens—researcher does nothing to affect the outcome
- Investigator does not control allocation or assignment of factors or therapy
- Combinations self-selected or are “experiments of nature” (includes “usual care” of MD office practice)
- Can be prospective (e.g., cohort) or retrospective (e.g., case-control) observational studies (i.e., case-control studies past cross-sectional present; and cohort future)
- Weaker empirical evidence than experimental studies but can answer different questions (e.g., prognosis/outcome?) or generate new ideas (i.e., hypothesis-generating—e.g., case series or registries)

Observational Studies –
Case-Control Studies
- Compare cases with disease to controls without disease, looking for difference in previous exposures (e.g., smoker or nonsmoker)
- Good for answering clinical questions about diagnosis and etiology
- Quick and relatively cheap
- Particularly useful for rare conditions or for risk factors with long induction periods (e.g., smoking and lung cancer)
- Provides relatively weak empirical evidence
- Main problems are recall bias and access to and reliability of old medical records

Observational Studies –
Cohort Studies (e.g., Nurses Health Study)
- Identify/follow defined populations (exposed and unexposed)
- Determine association between exposure and outcome
- Good for etiology and incidence
- Establish temporal relationships and multiple outcomes
- Prospective, therefore leading to stronger evidence than case-control studies but more $$$
- Do not provide empirical evidence as strong as RCTs
- Problems include blinding difficulty (no randomization) and confounding factors can be a problem (e.g., smoking in coffee drinkers with CVD)
- Susceptible to bias by differential loss to follow-up (i.e., attrition bias)
Observational Studies – Cross-Sectional Studies (e.g., NHANES Study)

- Cross-sectional analysis studies relationship between different variables at one point in time
- Provides “snapshot” of disease frequency/characteristics at particular point in time, assumed typical of entire group
- Data can be used to assess prevalence of acute or chronic conditions
- Since exposure/disease status measured at same point in time, not always possible to distinguish if exposure preceded or followed disease
- Simple and relatively cheap
- Large numbers needed

Aspects of Experimental Clinical Trial Design – Modern Doctrine

- Concurrent control group be used against which the new intervention be compared
- Randomization is the preferred way of assigning participants to control and intervention groups and minimizes the effect of bias and confounding on results
- Blinding is important; in the best studies “the blind lead the blind” (i.e., double-blind)

Randomized Controlled Trials (RCTs)

- RCT = “gold standard” study design
  - Results lead to highest level of scientific credibility
  - However, other study designs may be more appropriate depending on clinical question

Copyright © 2011 by Joslin Diabetes Center, Inc. All rights reserved. These materials may be used for personal use only. Any distribution or reuse of this presentation or any part of it in any form for other than personal use without the express written permission of Joslin Diabetes Center is prohibited.
Aspects of Experimental Clinical Trial Design

- Open/blind (single, double, or triple blind)
- Controlled/uncontrolled (no comparator)
- Placebo/active comparator
  - Placebo control
    - Effect of test drug is compared to effect of placebo
    - Patients have a chance of receiving placebo during trial, and therefore not receiving effective treatment – “equipoise”
  - Active control
    - Effect of test drug is compared to effect of an active reference drug ("gold standard" or other)
    - Used to show superiority, equivalance, or noninferiority between new and current drugs

Aspects of Experimental Clinical Trial Design

- Parallel vs crossover studies
  - Parallel: Each patient receives one treatment
  - Crossover: Each patient receives every treatment – 2 or more in a specific order
- Crossover studies enroll fewer patients (requires about 1/2.4 sample size of standard RCT), assume no carry-over effect
- Parallel studies are simple, can be used for any study, and make few assumptions

Phase IIb Trial Design for Study 1202 – LAF237 Dose-Ranging Study

- Drug naive patients with T2DM and HbA1c 7.5-11%
- Primary objective: HbA1c reduction from baseline – LAF237 20 mg, 50 and 100 mg vs. placebo

Source: Study 1202 data on file, Novartis Pharmaceuticals Corporation.
Who Are the Study Population?

- Will often be the tip of a large disease population iceberg
- Trial population should be defined
  - In advance
  - Stating unambiguous eligibility inclusion and exclusion criteria
  - With consideration of the ability to generalize the trial to the wider population

Randomization

- Ensures that all individuals entering a study have an equal chance of being allocated to any group within the study
- Tends to produce study groups comparable with respect to known and unknown risk factors, removes investigator bias in allocation, and guarantees that statistical tests will have valid significance levels
- All trial publications should state randomization process used

What Response Variables Will Be Measured?

- One response variable per primary question; usually one per trial
- Examples include:
  - Death
  - Complication of disease
  - Symptomatic relief
  - Clinical/laboratory finding
- Can measure more variables for the secondary questions
Definitions of Surrogate Endpoint, Clinical Endpoint, and Composite Endpoint

- **Surrogate endpoint**: A biomarker that is intended to substitute for a clinical endpoint. A surrogate endpoint is expected to predict clinical benefit or harm, or lack of benefit or harm (e.g., A1C, BP)
- **Clinical endpoint**: Clinically meaningful measure such as morbidity, mortality, and survival
- **Composite endpoint**: Combine several measurements into a single composite (e.g., MACE)

Rosiglitazone Reduces Free Fatty Acids

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean Free Fatty Acids (mmol L⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>4</td>
<td>0.8</td>
</tr>
<tr>
<td>6</td>
<td>0.7</td>
</tr>
<tr>
<td>8</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Rosiglitazone 2 mg BID
Rosiglitazone 4 mg BID
Glyburide

Analysis Policies for Endpoints

- **Underlying objective** is to minimize bias in therapeutic comparisons
- **Intention to treat (ITT)**
  - All eligible patients, regardless of compliance with protocol, should be included in the analysis of results wherever possible
- **ITT methods**
  - Last observation carried forward (LOCF): Value at time of withdrawal is carried forward and substituted at all subsequent assessments
Analysis Policies for Endpoints

- Per-protocol patients (PP) or completers
  - When an analysis of a well-defined target group of patients is required
  - This group will include only those patients who satisfied inclusion and exclusion criteria (eligible patients) and who also adhered to all major requirements of the protocol during the trial (evaluable patients)

Types of Outcome Measures

- Outcome measures could involve:
  - “Counting people” (binary or categorical data)
    - Summarized by % or proportion (risk)
  - “Taking measurements on people” (continuous data)
    - Summarized by average and spread (e.g., mean and SD)
  - “Time-to-event” data
    - Summarized using Kaplan-Meier plot, or survival or event rate at specific time point

Analysis of Clinical Studies

- The type of effect, size, and how it is analyzed depends on the type of outcome measure:
  - “Counting people”
    - Risk difference, relative risk, odds ratio
  - “Taking measurements on people”
    - Difference between two means or medians
  - “Time-to-event” data
    - Hazard ratio, difference between two survival or event rates
Analysis of Clinical Studies

- Is there a difference?
  - Examine the effect size.
- How big is it?
  - What are the implications of conducting a trial on a sample of people (confidence interval)?
- Is the effect real?
  - Could the observed effect size be a chance finding (P value or statistical significance)?
- How good is the evidence?
  - Are the results clinically significant?

Superiority, Equivalence, Noninferiority Trials

- Superiority trials: Designed to demonstrate that one treatment is more effective (i.e., statistically significantly superior) to another
- Noninferiority trials: Designed to demonstrate that a treatment is no worse than another
- Equivalence trials: Designed to demonstrate that one treatment is as effective as another
- Equivalence and noninferiority trials are usually conducted when the new intervention is expected to have fewer side effects, be more cost-effective, or more convenient

JUPITER – Study Design

No history of CAD or CVD
Men aged ≥ 50 years
Women aged ≥ 60 years
LDL-C < 130 mg/dL
(CR P ≥ 2.0 mg/L)

Visit:
- Week: 1, 2, 3, 4, 6
- Randomization
- Lead-in eligibility
- Placebo run-in
- Rosuvastatin 20 mg (n = 8901)
- Placebo (n = 8901)
- 6-monthly
- Final
- Median follow-up 1.9 years; maximum follow-up 9 years

CR P: C reactive protein
A1C: glycated hemoglobin

Access to Clinical Trial Results

- “The whole of medicine depends on the transparent reporting of clinical trials.”
- Declaration of Helsinki (2000) stated that all trial results (negative or positive) should be published.
- However, journals have publication bias for positive or topical studies.
- Despite clinical trials being listed at various sites (e.g., clinicaltrials.gov), transparent and timely access to results is less easy.
- This can lead to difficulty in EBM, especially systematic reviews and meta-analysis generation.

Copyright © 2011 by Joslin Diabetes Center, Inc. All rights reserved. These materials may be used for personal use only. Any distribution or reuse of this presentation or any part of it in any form for other than personal use without the express written permission of Joslin Diabetes Center is prohibited.
Systematic Reviews and Meta-Analyses – General Features

- Systematic reviews attempt to access and review systematically all of the pertinent articles in the field
  - Obtain, analyze, and interpret all available reports on particular topic
- A meta-analysis is a formal statistical analysis combining the results of several studies that address a set of related research hypotheses
- Normally done by identification of a common measure or endpoint
- More powerful estimates of true effect size than derived in single study

Systematic Reviews and Meta-Analyses – General Features (cont’d)

- Goals
  - Identify important gaps in existing knowledge (e.g., are new RCTs needed?)
  - Obtain overall precise estimates of associations
  - Identify possible sources of heterogeneity in associations
- Methods
  - Studies selected on a well-specified subject (i.e., specific therapy in a defined population) – this depends on research question
  - Search literature, select studies (incorporation criteria) based on quality criteria – clinical study design, randomization, and blinding
  - Identify which dependent variables/summary measures are allowed – critically appraise information
  - Perform meta-analysis
  - Interpret and summarize findings

Systematic Reviews and Meta-Analyses – General Features (cont’d)

- Disadvantages
  - Study quality: A good meta-analysis of badly designed studies will still result in misleading findings
    - Meta-analysis findings were incorrect ~30% of the time when a definitive, large RCT was ultimately performed
  - Identifying studies and publication bias: Most meta-analyses depend heavily on published studies (i.e., studies that show positive as opposed to negative results are usually published—termed publication bias); this can result in exaggerated or false outcomes
Meta-analysis of Intensive Statin Trials: Coronary Death or Myocardial Infarction


LDL 75 vs 101 mg/dL