Module 1 Summary and Study Examples

Laura Lee Johnson, Ph.D.

Associate Director

Division of Biostatistics III

Center for Drug Evaluation and Research

US Food and Drug Administration

IPPCR Course Fall 2015

Disclaimer

This presentation reflects the views of the author and should not be construed to represent FDA's views or policies.

Overview

Many different people take the course

Some first introduction, some advanced

Tricks, tips, and concepts

General objectives

Better consumer of medical and scientific literature

Enhance conversations inside research team [with study statisticians and epidemiologists]

Better science

Asked What You Should Know

Here are their slides

A few edits

Ordered to tell a story

Easy to Write

The study will use a randomized, double-blind, controlled parallel arm design and an intent to treat analysis.

Subjects/participants will be consented.

Not Easy

To implement and maintain the integrity of

Randomization

Blinding/masking

Multiple study arms

Data collection

Transfer data to regulatory and other groups

Controls and Interventions

Specific question being addressed in the study directs the choice of the control group or groups for the study

Many times it is easier to decide what interventions should be in the intervention arm than the control arm

Effect Modification

Interaction

Synergy

Could be larger or smaller

Association between outcome and another variable (e.g. intervention) is modified by different levels of a third variable

Smoking, Asbestos, and
Lung Cancer

Smoking (alone) ↑ risk of lung cancer by A

Asbestos exposure (alone) ↑ risk of lung cancer by B

Smoking AND having asbestos exposure ↑ risk of lung cancer by MORE/LESS than A+B

Confounding

Two or more variables

Known or unknown to the researchers

Confounded when their effects on a common response variable or outcome are mixed together

Association between an exposure and outcome is misestimated due to the failure to account for a third factor (the confounder)

Consider

Association observed between carrying matches in your pocket and lung cancer

Carrying matches causes lung cancer

OR

Association between carrying matches and lung cancer is result of confounding by another unmeasured variable associated with both

(Pam Shaw, CTR Course 2013)

Two Types of Research Studies

Observational

Goal is to observe and collect data on characteristics of interest without influencing the participant, environment or disease course

Experimental

Researcher deliberately influences course of events and investigates effects of an intervention on a carefully selected population of subjects

Experimental studies done on human subjects are referred to as clinical trials or clinical studies

Observational Studies

Case Reports

Case Series

Cross-Sectional or Prevalence Surveys

Case-Control Study

Cohort Study (longitudinal)

Natural History Studies

Ecological Studies (data on population rather than individual level)

Epidemiology

Assumes disease has causal and preventative factors that can be identified through systematic investigation

Quasi Experimental, One/Single Arm, or
Non-Randomized Experimental Studies

No control group

Early in investigation

Concurrent control “group”

Treatment assignment not by randomization

Historically controlled

Missing data

Poor data

Non-comparability of groups

Intervention Based Research Spectrum

Quasi-experimental

Pre-clinical studies

Phase 0

Phase I

Early/Late Phase II

Phase III

Phase IV

Dissemination and Implementation

Comparative or Cost Effectiveness

Non-Randomized Randomized

Can ONLY show Association

You will never know all possible confounders!

Can show Association AND Causality

Well done non-adaptive randomization → unknown confounders should not create problems

Ideal Study - Gold Standard

Treatment / control

Parallel groups

Superiority

Prospective

Double blind / masked

Randomized

Types of Randomized Studies

Parallel Group – classic

Sequential Trials – physical sciences

Group Sequential trials – classic

Cross-over – intervention washout

Factorial Designs – independence

Adaptive Designs – gaining popularity

Enriched Enrollment – regression to the mean

Cluster Randomized Designs

Drill in a Bit

Intent-to-Treat versus Completers
"How Many in the Data Analysis"

ITT = Intent-To-Treat analysis

Include all randomized (if randomized study)

Assume all study participants

Adhered to study regime assigned

Completed the study

MITT = Modified ITT analysis

ITT, but only include people who start intervention they are assigned to

Completers or Adherers analysis

Only the well behaved

Masking/Blinding

Less common in non-randomized studies, but can mask outcome assessors as to hypothesis

Specify whom to be masked, why, how, and to what

Assess effectiveness of masking

Specify criteria for unmasking, whom to be unmasked

Mask determination of outcome so that reviewers are unaware of treatment assignment; provide information on "need to know" basis

What is being adapted?
(Types of adaptations)

Adaptive randomization

Adaptive dose-finding (dose-ranging)

Drop-the-loser (or pick-the-winner)

Adaptive seamless phase II/III

Biomarker-adaptive

Group sequential methods

Sample size re-calculation

Reproducible Measurements:
Regardless of Study Design

Well defined cohort

Exclusion and inclusion criteria

Study conduct

Outcomes

Study data, analyses

Biases in Clinical Research

Issues in Randomization

What was covered:

What is “random”?

What is randomization?

Why randomize?

Whom/what to randomize?

How to randomize?

Recommendations
(for common randomization methods)

Use computer program or online tools

Use permuted-block randomization with small random block size (<10)

For multi-site clinical trials, use site as a stratification variable and in the statistical model

Do not use too many stratification variables (≤4)

Unless necessary, avoid adaptive randomization methods

Recommendations
(for complex randomization methods)

Consult with a biostatistician

Implementation Recommendations

Make it possible to reproduce the string of treatment assignments

Document randomization method used

Put in place features that prevent treatment assignment until conditions for entry into the trial are fully satisfied

Mask (blind) assignments to everyone concerned

Make it difficult (impossible) to predict future assignments from past assignments

Put in place procedures for monitoring departures from established protocols

Shaw et al. 2012

Measurement

RELIABILITY: Problems

RELIABILITY: Improving

SENSITIVITY to CHANGE

Clinical Relevance

Self-Report Measures

Used in questionnaires

Some information not easily observable

Self-report measures are estimates of true scores

Susceptible to:

Respondent's mood, motivation, memory, understanding

Context of data collection

Social desirability

Rigorous methods will mitigate these pitfalls

Patient-Reported Outcome
(PRO)

"A measurement based on a report that comes directly from the patient (i.e., study subject) about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else."

December 2009 FDA Guidance

Put Together Draft Questionnaire

Use existing instruments when possible

HealthMeasures.net

PROMIS, Neuro-QoL, ASCQ-Me, NIH Toolbox

REDCap

National field surveys

NHIS, NHANES (US)

National Health Service Survey (UK)

Shared library of data collection instruments

Criteria for Good Survey Questions

Literacy < 9th grade U.S.

Specific better than broad

Culturally sensitive

Scales consistent

Terms well-defined

Instructions clear

Reference periods clear

Response options match question

Multiple concepts separated

Criteria for Good Survey Questions

Interpreted accurately by people with range of demographic characteristics

Capturing what researcher intended

Avoid

Social desirability effects

Negative wording

Double barreled

Jargon

Ambiguous

Leading

Development and Evaluation of PRO Measures

Determine what PRO concept we want to measure and why

Collect qualitative data to understand meaning of the PRO concept

Write items you think will measure the concept

Test items for understanding (cognitive interviews)

Administer items to a large sample of people

Use psychometric (statistical) analyses to see how well items are working and develop scoring method

Evaluate the reliability and validity of measure

5. Translation

PROMIS Approach

Harmonization: different words/ languages must mean the same

Universal approach: One language version for multiple countries

People from various countries/dialects involved

Item Bank

Large collection of items measuring a single domain

Any and all items can be used to provide a score for that domain

Dynamic, not fixed

Summary

Questionnaire development requires careful planning

Use existing validated instruments when possible

Rigorous methods will reduce response error

Participant Selection

Exclusion Criteria

Safety!

Reduce potential confounding

Secondary Data:
Using Claims and Similar Data

Where Does Data Come From?

Primary Data – generated for research purposes, including national surveys and disease registries

Secondary Data – Secondary data is administrative/billing/encounter data

Often generated with utilization in mind

Enough data to make meaningful population-base conclusions

Why Study Medicare Patients?

Largest purchaser of health care in the world

54 million enrollees

$613 billion in expenditures in 2014

The percentages

Almost 16% of U.S. budget outlays

22% of all health care dollars in US

26% of hospital spending

22% of nursing home spending

22% of physician billings

Single data system

Benefits of Medicare Data

Pre-existing data

less expensive

less time

Large numbers of cases

Generalizbility

Links to other data

Zip codes, SSN

Accurate measure of resource use

Can measure “effectiveness”

Limitations of Medicare data

Lots of limitations…

Limited data on severity of illness

Not generalizable to the US working population

Coding and billing errors/bias

Limited outcome measures of interest

No QOL, Patient Satisfaction, functional assessment, illness severity

Limitations

Studies limited to non-experimental design (observational studies)

Difficult to avoid selection bias

Impossible to control for all possible confounders (eg. severity of illness & functional status)

HMO patients are excluded

Cost of obtaining the data

Administrative overhead

How can I get data access?

Medicare administrative/billing/encounter data

http://www.resdac.org/about-resdac/our-services

Available through CCW Data Enclave

Different countries, different data, different access

Overview of
Hypothesis Testing

Basic Sample Size

Changes in the difference of interest have HUGE impacts on sample size

20 point difference → 25 patients/group

10 point difference → 100 patients/group

5 point difference → 400 patients/group

Changes in difference to be detected, α, β, σ, number of samples, if it is a 1- or 2-sided test can all have a large impact on your sample size calculation

Lots of Formulas are Not Basic

Depending on your study design you may need to do simulations

Power is Affected by…..

Variation in the outcome (σ^2)

$\downarrow \sigma^2$ → power ↑

Significance level (α)

↑ α → power ↑

Difference (effect) to be detected (δ)

↑ δ → power ↑

One-tailed vs. two-tailed tests

Power is greater in one-tailed tests than in comparable two-tailed tests

What to Do

Make a sample size or power table

Make a graph, or many graphs

Use a wide variety of possible standard deviations

Similarly, think about different possible relationships between the study arms

Protect with high sample size if possible

Survival Analysis

Making inference about EVENT RATES

Rate at t = Rate among those at risk at t

Look at Median survival (50%) not Mean survival

Mean: need everyone to have an event

Cox Regression is the most robust method

Kaplan Meier curves do not have sensible interpretations for competing risks

Independence is key

Truncation is about entering the study

Censoring is about leaving the study

Powers: Conclusions

Developing an efficient trial starts with planning and a good research question

Question comes first, sample size second

Various methods to increase effect sizes and decrease variability, when applied in the correct setting, can provide valid and reliable answers to important public health questions

For some diseases, developing the tools (better outcome measures, better data on natural history) is a good start to better trials

Analysis Follows Design

Questions → Hypotheses →

Experimental Design → Samples →

Data → Analyses →Conclusions

Take all of your design information to a statistician early and often

Guidance

Assumptions

Your Exam Questions

Will go beyond this, but this is a brief summary

Economist, 12 December 2015
Randomised Controlled Trials

In praise of human guinea pigs: Doctors use evidence when prescribing treatments. Policymakers should, too

http://www.economist.com/news/leaders/21679799-doctors-use-evidence-when-prescribing-treatments-policymakers-should-too-praise-human

Measure for measure: How to test everything from sluggish teenagers to corrupt bureaucrats

http://www.economist.com/news/international/21679811-how-test-everything-sluggish-teenagers-corrupt-bureaucrats-measure-measure

Norman (Norm) Breslow

February 21, 1941 - December 9, 2015

Wilms Tumor Study Group

Beginning 20th century 90% of children diagnosed with Wilms tumor died

Beginning of 21st century nearly 90% survived and lead relatively normal lives

Refined treatments

Long term effects of radiation and chemotherapy

Multi-modal cooperative group; full members from many disciplines

Obtaining Cure with Minimum Impact on Patient and Family

High priority on data collection, electronic data collection, careful follow up, registering patients

Pathologist identified small subgroup with unfavorable histology

Differential treatment strategies

Breslow (biostat) and Day (epi) developed and popularized case-control matched sample designs

Findings are not definitive but can inform design of slow and expensive longitudinal large-cohort studies that are (more) definitive

First Child Enrolled 1969

Over 40 years many studies

Some randomized, some not

Some interventional, some observational

Some small, some large

Some with genetic component, some not

Ethics of lowering dose of life saving therapy?

Patients and parents had to assent/consent

Late Effects Study

Adverse pregnancy outcomes, second tumors at relatively early ages, skeletal and muscle development

Questions?

Thanks!

Please fill out the course evaluations

Post specific examples

Send questions to the appropriate board for each lecture