Stat.BD

APPLIED BIOSTATISTICS & RESEARCH METHODOLOGY • SESSION 1

The Story of Research

Where a research question comes from, and what statistics is actually for

Muhtasim Munif Fahim, MSc
Mentor, Stat.BD
Research Associate, Data Science Research Lab
Department of Statistics & Data Science, University of Rajshahi

Session 1 of 30

Before anything else

What this course is

A course about *asking and answering clinical questions properly*. Statistics is the tool it uses, not the subject itself.

What you already bring

Ten years at the bedside. That is not a gap in your training. It is the one thing in this room that cannot be taught.

Three ground rules

- No derivations. Every idea arrives through a paper or a patient.
- Stop me. A question mid-sentence is worth more than a nod at the end.
- You will read real papers from the first session, not simplified extracts.

Where this goes: 15 weeks, 30 sessions, 14 SPSS labs

Phase	What you build	Sessions
I	Statistical foundations: the language, summaries, distributions, sampling variability, confidence intervals, p values	1–8
II	Research methodology: reading papers, questions, study designs, measurement, bias, sampling, sample size, protocol & ethics	9–18
III	Applied biostatistics: choosing a method, group comparisons, effect measures, regression, multivariable thinking	19–26
IV	Research literacy & communication: diagnostic and prediction studies, survival, systematic reviews, manuscript	27–30

The shape of it

Phase I gives you the vocabulary. Phase II teaches you to read and design. Phase III lets you analyse. Phase IV makes you dangerous.

The exit competency

By Session 30

You will be handed a medical paper you have never seen, in a field that is not yours, and you will be able to say, out loud and in front of people, what its question was, whether its design could answer it, what its numbers mean, where it is vulnerable, and whether you believe it.

Everything between here and there exists to make that one afternoon possible.

Today, in ninety minutes

I Why does research exist at all?
Three ways a doctor comes to believe something, and why only one of them can be checked.

II The anatomy of a paper
We take one real trial apart, section by section.

III Where statistics actually enters
And the one distinction this session exists to teach.

IV Your turn
Your clinical irritation, converted into a research question.

Notice what is missing

There is no formula in today's session. Not one. If you leave able to *calculate* something, I have taught the wrong class.

Act I

Why does research exist at all?

Stat.BD

A question for a doctor with ten years behind him

You have treated thousands of patients.

How do you know what works?

Three ways a doctor comes to believe something

Authority

A professor said so. A textbook says so. A guideline says so.

Fails when: the authority was repeating *their* teacher. Someone, somewhere, has to have actually looked.

Experience

"In my hands, this works."

Fails because: you only ever see the patients you treated. You never meet the version of them you did not treat. Improvement and recovery look identical from the bedside.

Evidence

"Here is what happened when we compared."

Costs: time, money, effort, and the willingness to be proved wrong.

Research is not a different *kind* of knowing. It is the third one, done deliberately enough to be checked by someone else.

Act II

The anatomy of a paper

Stat.BD

The paper we will live inside today

COBRA-BPS

Jafar TH, Gandhi M, de Silva HA, Jehan I, **Naheed A**, et al. *A Community-Based Intervention for Managing Hypertension in Rural South Asia.* N Engl J Med 2020;382(8):717–726.

- 30 rural communities in **Bangladesh**, Pakistan and Sri Lanka
- 2,645 adults with hypertension
- Bangladesh arm run from **icddr,b**, Dhaka
- Follow-up completed for more than 90% at 24 months

Why this one

Because it is *yours*. The patients are Bangladeshi, the health workers are government health workers, the problem is one you see every clinic day.

And because it was published in the *New England Journal of Medicine*. Both of those facts are true at the same time. Remember that.

Every paper has the same skeleton

Title & Abstract

Introduction

Methods

Results

and **D**iscussion

References

Every paper has the same skeleton

Title & Abstract

Introduction

Methods

Results

and **D**iscussion

References

The whole study in 250 words.
Read this first, and read it last.

Every paper has the same skeleton

Title & Abstract

Introduction

Methods

Results

and **D**iscussion

References

The whole study in 250 words.
Read this first, and read it last.

Why this question?
Ends with the objective. If you cannot find it, that is the paper's fault.

Every paper has the same skeleton

Title & Abstract

Introduction

Methods

Results

and **D**iscussion

References

The whole study in 250 words.
Read this first, and read it last.

Why this question?
Ends with the objective. If you cannot find it, that is the paper's fault.

What they did, exactly.
Design, population, variables, analysis plan. The section that decides whether to believe the rest.

Every paper has the same skeleton

Title & Abstract

Introduction

Methods

Results

and **D**iscussion

References

The whole study in 250 words.
Read this first, and read it last.

Why this question?
Ends with the objective. If you cannot find it, that is the paper's fault.

What they did, exactly.
Design, population, variables, analysis plan. The section that decides whether to believe the rest.

What happened. No opinions.
Numbers, tables, figures.

Every paper has the same skeleton

Title & Abstract

Introduction

Methods

Results

and **D**iscussion

References

The whole study in 250 words.
Read this first, and read it last.

Why this question?
Ends with the objective. If you cannot find it, that is the paper's fault.

What they did, exactly.
Design, population, variables, analysis plan. The section that decides whether to believe the rest.

What happened. No opinions.
Numbers, tables, figures.

What the authors think it means.
Their interpretation, not evidence.

The title has already told you the design

"A **Community-Based** **Intervention** for **Managing Hypertension** in **Rural South Asia**"

Community-Based

Not a hospital study. The unit being acted on is a *community*, which will turn out to matter a great deal for the analysis.

Intervention

Somebody *did* something. This is not observation. Expect a control group and expect randomisation.

Rural South Asia

The population. Also the honest limit on who these results apply to.

Reading an abstract in five moves

1. **Who?** "2,645 adults with hypertension" in "rural districts in Bangladesh, Pakistan, and Sri Lanka"
2. **What was done, and instead of what?** "a multicomponent intervention" versus "usual care"
3. **What was measured?** "The primary outcome was reduction in systolic blood pressure at 24 months"
4. **What happened?** "the mean reduction was 5.2 mm Hg greater with the intervention (95% CI, 3.2 to 7.1; P<0.001)"
5. **Do the conclusions stay inside the data?** Read the last sentence, then check it against move 4

Common trap

Move 5 is where papers most often misbehave, and where reviewers catch them. The Results say 5.2 mm Hg. If the Conclusion says "transforms cardiovascular outcomes", the paper has left its own evidence behind.

The same five moves have a name: PICO

P	**Population**: adults with hypertension in rural Bangladesh, Pakistan, Sri Lanka
I	**Intervention**: home visits by trained government community health workers, physician training, care coordination
C	**Comparison**: usual care
O	**Outcome**: reduction in systolic blood pressure at 24 months

A question written in PICO form is already halfway to a design, which is exactly why Session 10 makes you write your own.

Where the numbers actually live

Table 1

Who was in the study.
Age, sex, baseline blood pressure and medication, for each group separately.

You read Table 1 to ask: *were these two groups actually comparable to start with?*

Table 2 onwards

What happened.
The outcomes. Effect estimates, confidence intervals, *p* values.

This is the table the abstract was summarising.

Figures

The story.
Flow diagrams, curves over time, forest plots.

The flow diagram tells you who was lost, and is often the most revealing figure in the paper.

A useful reflex

In COBRA-BPS, baseline mean systolic pressure was 146.7 mm Hg in the intervention group and 144.7 in the control group. Nearly the same, which is what randomisation is supposed to buy you.

What they found

Outcome at 24 months	Intervention	Usual care	Difference (95% CI)
Fall in systolic BP	9.0 mm Hg	3.9 mm Hg	**5.2 greater** (3.2 to 7.1), $P < 0.001$
Fall in diastolic BP	group means not reported		2.8 greater (1.7 to 3.9)
BP controlled (<140/90)	53.2%	43.7%	RR **1.22** (1.10 to 1.35)
All-cause mortality	2.9%	4.3%	*reported; not the primary outcome*

In plain language

Four rows. Two different *kinds* of number: a difference in millimetres of mercury, and a ratio of two percentages. They are answering the same question in two different currencies.

Values as reported in Jafar TH *et al.*, N Engl J Med 2020;382:717–26. Table re-typeset here; no figure or table from the paper is reproduced.

One sentence, in slow motion

"the mean reduction was **5.2 mm Hg greater** with the intervention (**95% CI, 3.2 to 7.1**; **P<0.001**)"

Magnitude

How big? **5.2 mm Hg.**
A clinical question. Only you can judge it.

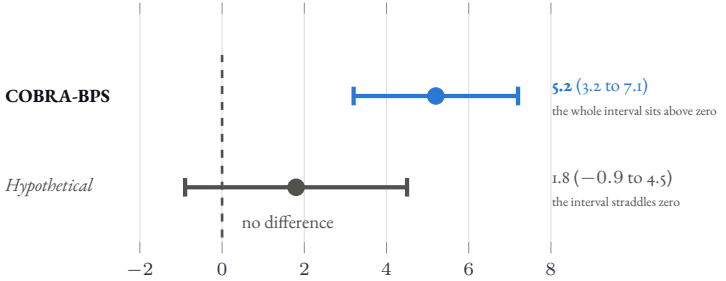
Precision

How sure of the size?
Somewhere between 3.2 and 7.1.
Not "between 3.2 and 7.1 people", but a range for the *true* difference.

Evidence

How badly does this fit "no difference at all"? **Very badly.**
That, and only that, is what the *p* value says.

Why the interval matters more than the verdict



Both studies found the intervention better on average. Only the first one found it *convincingly* better, because its entire plausible range sits on one side of "no difference".

The same result, in a different currency

Relative

Blood pressure was controlled in 53.2% versus 43.7%.

Relative risk 1.22. Control was 22% *more likely* with the intervention.

Sounds substantial. Relative numbers usually do.

Absolute

$53.2\% - 43.7\% = \mathbf{9.5}$ percentage points.

Roughly **11 people** need the programme for one extra person to reach control.

Computed here from the reported percentages; the paper reports the relative risk.

Common trap

Identical data. One framing sells a programme; the other lets a health ministry cost it. Whenever you are given only the relative number, ask for the absolute one.

Is 5.2 mm Hg worth it?

Statistical significance

A statement about *noise*.
"A difference this large would be unlikely if there were really no difference at all."

Given a big enough study, almost anything becomes statistically significant.

Clinical significance

A statement about *patients*.
"Is 5.2 mm Hg, sustained across a whole rural population, worth what this programme costs?"

No statistical test will ever answer this. It needs a clinician.

In the clinic

For one patient in your clinic, 5.2 mm Hg is almost nothing. Across millions of people, a population-wide shift of that size moves strokes and heart attacks. Both statements are true. Deciding which one applies is your job, not the statistician's.

Five questions, every paper, forever

1. **What was the question?** If you cannot state it in one sentence, either you have not found it or they never had one.
2. **Who was studied, and who was left out?** Generalisability is decided here, not in the Discussion.
3. **What was compared with what?** No comparison, no evidence.
4. **How big was the effect, and how precise?** The estimate *and* the interval, never the *p* value alone.
5. **What could explain this other than the intervention?** Bias, confounding, chance.

In plain language

Five questions. Twenty minutes. You now have a defensible first opinion about any clinical paper in the world.

Act III

So where does statistics actually enter?

Stat.BD

It was there the whole time

Title & Abstract

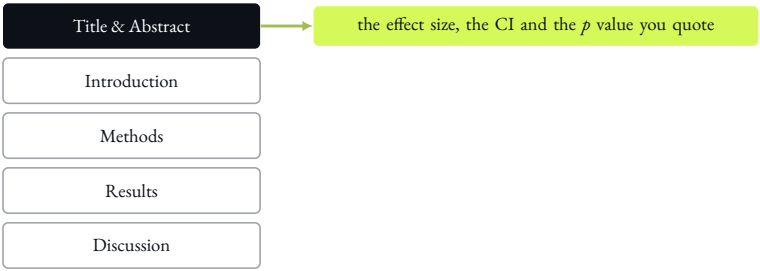
Introduction

Methods

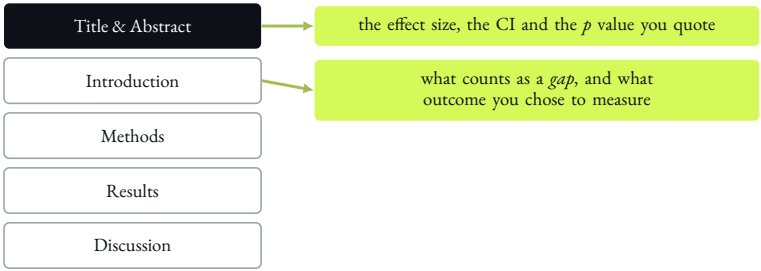
Results

Discussion

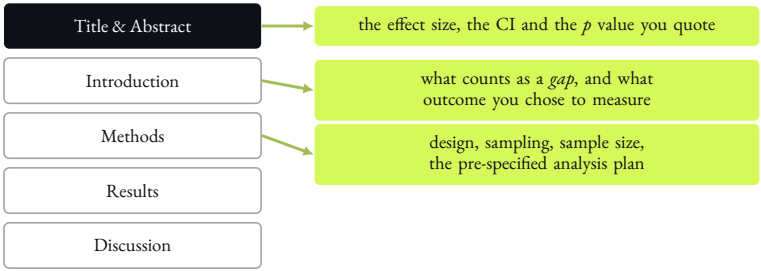
It was there the whole time



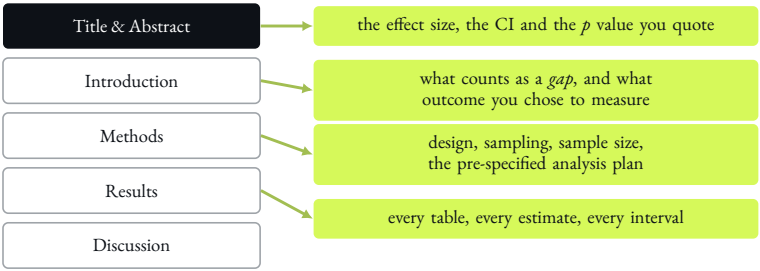
It was there the whole time



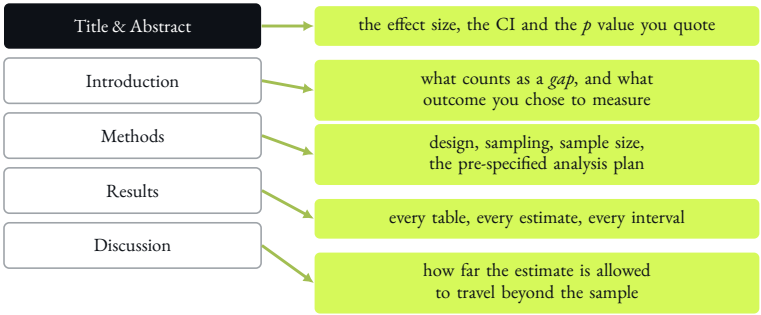
It was there the whole time



It was there the whole time



It was there the whole time



The one distinction this session exists to teach

Description

What happened in **these** people.

2,645 adults. Mean baseline 146.7 and 144.7 mm Hg. 53.2% controlled. A fall of 9.0 versus 3.9 mm Hg.

These are *facts*. Nobody can argue with them.

Inference

What that says about **everyone else**.

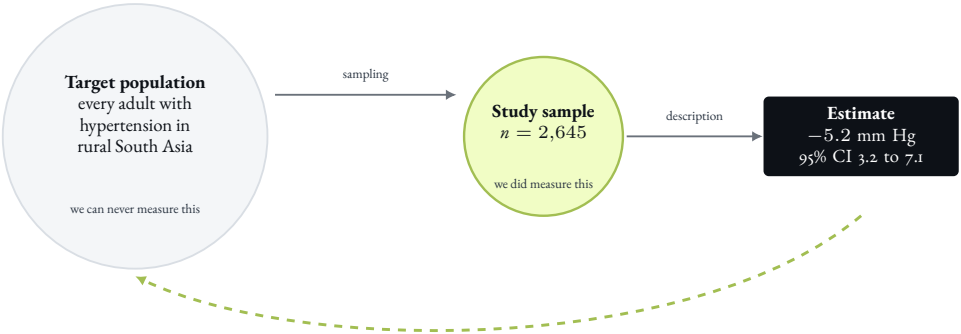
"3.2 to 7.1" is not a fact about the 2,645. It is a claim about the millions of rural South Asians who were never in the study.

This is a *bet*, made carefully, with its odds stated.

Common trap

Almost every misuse of statistics in medicine is a descriptive statement wearing the clothes of an inferential one, or the reverse.

The arrow that points backwards



inference: the only arrow that points back at the people we never met. This is the risky one.

Why we can never simply be certain

- To be certain, you would have to measure *everyone*. Nobody has ever done that.
- So you take a sample and a different sample would have given a slightly different answer.
- That wobble is not a mistake. It is not sloppy measurement. It is a permanent property of looking at part of something.
- Statistics does not remove the wobble. It *measures* it, and then reports it honestly.

The real definition

A confidence interval is a scientist admitting, in public and with a number attached, exactly how wrong they might be.

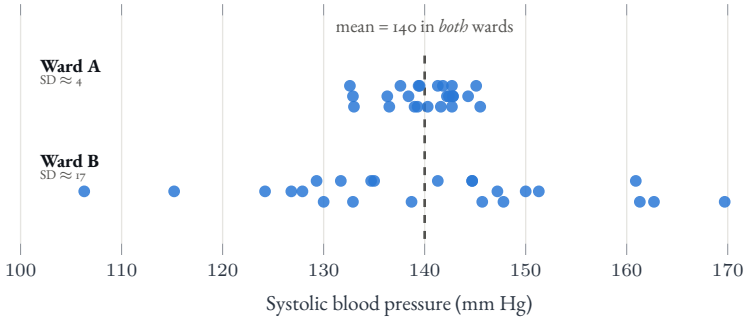
Session 7 gives this wobble its name the *standard error*, and Session 8 turns it into intervals and *p* values.

Two ways to be wrong, in a table you already know

YOUR STUDY SAYS	THE TRUTH (which nobody ever gets to see): there *is* an effect	THE TRUTH (which nobody ever gets to see): there is *no* effect
"significant"	**Correct** — you found a real effect	**Type I error** (α) — false positive: you treated noise as signal
"not significant"	**Type II error** (β) — false negative: you missed a real effect	**Correct** — nothing there, and you said so

You already own this table. It is sensitivity and specificity, except that the "patient" being tested is a *hypothesis*.

The mean is the same. The patients are not.



Two wards. Identical average blood pressure. Would you manage them the same way?

Illustrative data, constructed so that both means are exactly 140 mm Hg.

What statistics cannot fix

A bad question

Vague, unanswerable, or nobody's problem. No analysis rescues it.

→ Sessions 10–11

A biased sample

If the wrong people were studied, a bigger sample makes you *more confidently* wrong.

→ Sessions 15–16

A bad measurement

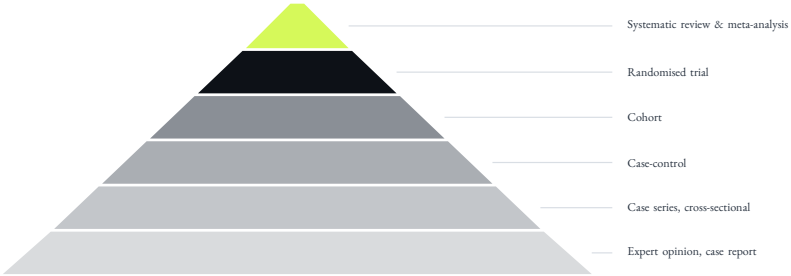
If the instrument does not measure what you claim, the numbers are tidy and meaningless.

→ Session 14

In plain language

Statistics is not a cleaning service you call after the data are collected. Almost everything that decides whether a study is worth anything is decided *before the first patient is enrolled*.

Not all evidence weighs the same



More studies sit at the bottom.
More *certainty* sits at the top.

In plain language

Higher on this pyramid means *harder to fool*, which is not the same as a better paper.

A well-run cohort study beats a badly-run trial every time. The pyramid ranks *designs*, not *quality*, and telling those two apart is most of what Phase II teaches you.

Act IV

Your turn

Stat.BD

Where research questions actually come from

Ten years of clinical practice is not a gap in your CV. It is the only part of this process that cannot be taught.

In the clinic

Nobody ever found a research question in a textbook. They are found in clinics, on ward rounds, and in the specific irritation of watching the same avoidable thing happen for the tenth time.

Now write one of your own

P Population .....

I Intervention or exposure .....

C Comparison .....

O Outcome .....

The two that beginners skip

C and **O**. If you cannot say what you are comparing against, you do not have a study. If you cannot say exactly what you will measure, you do not have an outcome. You have a hope.

What a clinical research career abroad actually asks for

What they look for	What it means in practice	Built in
Paper literacy	Appraise an unfamiliar study on the spot, in an interview	1, 9, 30
A real question	A focused, feasible question in PICO form, with a justified gap	10, 11
Design fluency	Name the right design for a question, and defend it	12, 13
Honest reporting	CONSORT / STROBE, registration, pre-specified analysis	18, 30
A named skill	Something you can actually run: SPSS, regression, survival	19–28, Labs
Output	One first-author paper beats any certificate	30

In plain language

None of that is talent. It is a checklist, and this course works through it in order.

Framework references: NIH IPPCR; EQUATOR Network reporting guidelines.

Before the next session

Do

1. Write **one** clinical question from your own practice, in PICO form. One sentence per letter.
2. Search PubMed and find **one** paper that answers part of it. Any paper. Do not worry about doing it properly yet.
3. Run the **five questions** over it and bring your answers.

Read

Doll R, Hill AB. *Smoking and carcinoma of the lung; preliminary report.* BMJ 1950;2:739–48.

Free full text on PMC (PMC2038856). Read it for the *design*, not the numbers, and watch how they built the control group.

Bring your PICO on paper. We will pull it apart together, kindly.

Next: Session 2, Population, Samples, Variables & Data

Where we started

You have treated thousands of patients.

Now you know why that is not yet an answer.

Everything we do for the next twenty-nine sessions is an argument about how to count well.

References & sources

Papers

- Jafar TH, Gandhi M, de Silva HA, Jehan I, Naheed A, et al. A community-based intervention for managing hypertension in rural South Asia. *N Engl J Med* 2020;382(8):717–726. doi:10.1056/NEJMoa1911965 (PMID 32074419)
- Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *Br Med J* 1950;2(4682):739–748. doi:10.1136/bmj.2.4682.739 (PMID 14772469; free full text PMC2038856)

Framework NIH *Introduction to the Principles and Practice of Clinical Research* (IPPCR); EQUATOR Network reporting guidelines. Literature identified via PubMed.