Introduction to the Division of Biostatistics and other PHS services

Yueren Zhou
Henry Ford Health, YZHOU1@hfhs.org

Xuan Ma
Henry Ford Health, xma1@hfhs.org

HENRY FORD HEALTH®

Introduction to the Division of Biostatistics and other PHS services

Nursing Research Conference

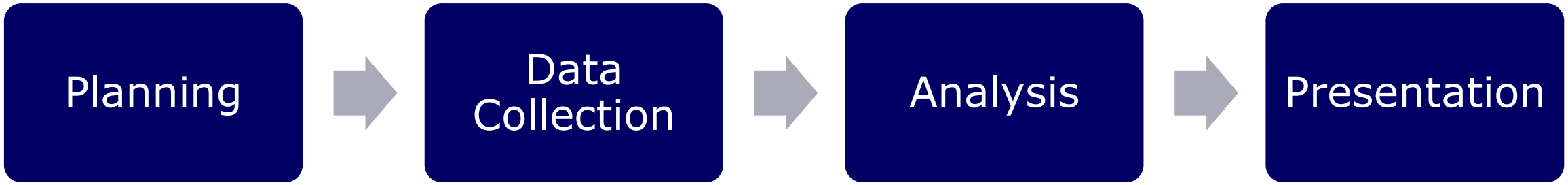
September 24, 2024

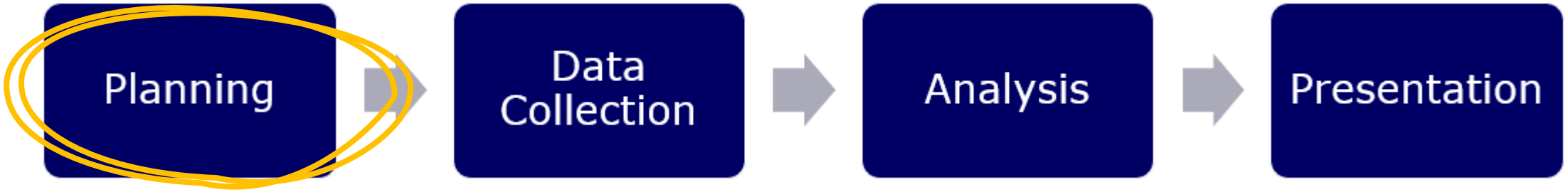
Yueren Zhou, Xuan Ma

Research Process and PHS services

- Planning
- Data Collection
- Analysis
- Presentation

Research process



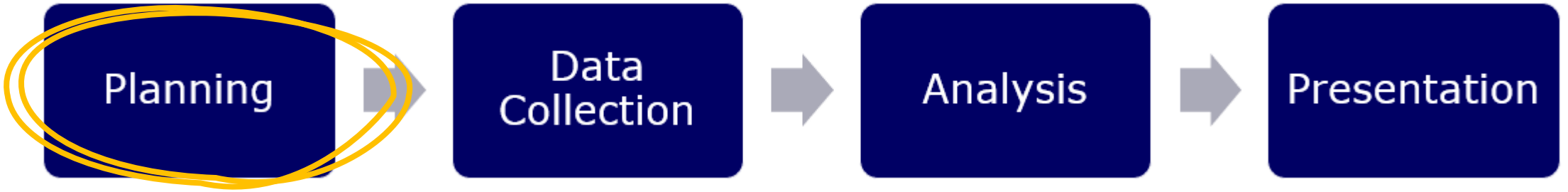


Step 1 – Identify the project objectives, consider:

- What is the gap in understanding that this research will fill?
- How will this project take the **next step** in a bigger goal?
- What do we want to learn?

Step 2 – Determine if the research will **test** or **measure** something

- Test – set up a hypothesis (e.g., comparing outcomes from two clinical practices)
- Measure – identify the information that you want to understand better (e.g., the rate of falls, the level of health literacy of current employees)

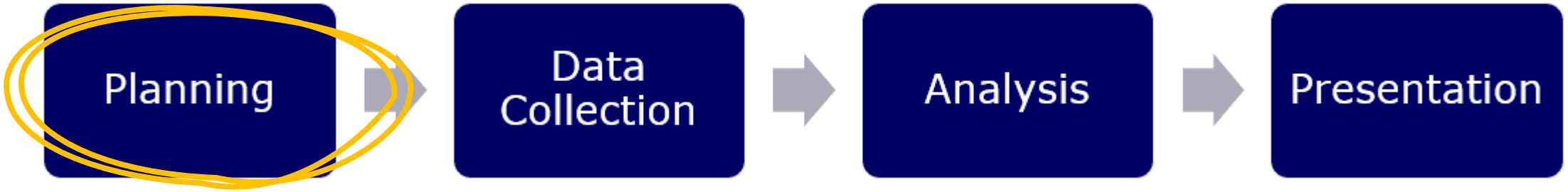


Step 3 – Identify the boundaries of the study

- Who will be selected to participate?
- How long will you recruit participants and follow them?

Step 4 – prepare a project protocol

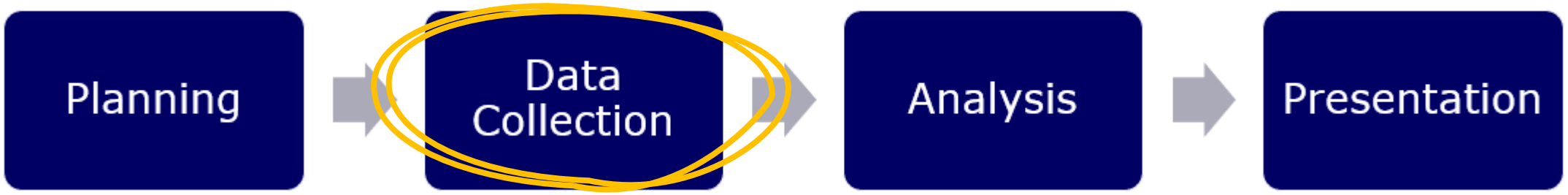
- Submit the protocol to the IRB for risk assessment
- Submit to leadership for funding/time support



How PHS & Biostatistics can help with planning

- The Division of Biostatistics will meet with researchers to discuss their study goals
- We are experienced at defining testable hypotheses and setting data-driven objectives
- We will help to write an analysis plan for the protocol
- We can help to identify a reasonable sample size so to avoid inconclusive results from too few people – or excess expense from too many

- PHS also has faculty epidemiologists and health services researchers who may be available for consultation about project goals



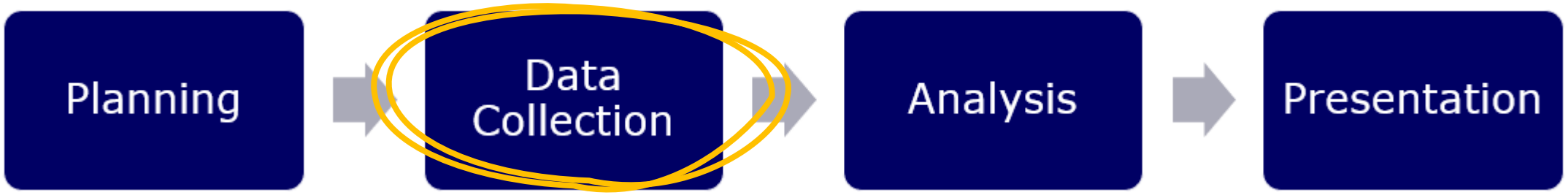
How PHS & Biostatistics can help with data collection

EMR data queries – PHS has staff trained in programming to extract data from the medical record (EPIC) and billing (CDS).

- Hourly charge for programming, so the price depends on the complexity of the request
- The benefit is that large amounts of data can be acquired following consistent rules

NEW – EPIC COSMOS – Biostatistics has two COSMOS superusers

- Query and analyze data across EPIC sites – great for rare diseases or outcomes
- Seats for a slicer-dicer version may be available
- Deeper analysis must be done in the COSMOS platform by a superuser



How PHS & Biostatistics can help with data collection

Surveys & Manual data collection – The RedCap team can help to open a new RedCap survey form. The charge is different for a self-build or hourly support to create the project.

- Surveys can be sent as a survey to participants by email
- OR it can be used to collect data that the researchers identifies in the medical record.
- The benefit is that it uses quality checks for data entry and keeps track of the history of data changes.

Survey development – Biostatisticians are experienced in survey question development to make sure that the needed information is robustly captured to meet your research goals

Platforms for Data Collection and Entry

Excel

Easiest to use, but also easiest to mess up

No tracking of change history

Few field format checks

Available to all

REDCap®

REDCap

Good for large or small projects

Direct data entry or survey

Can use data entry forms

Some reporting features

Tutorials to support setup

$300/project for tech support and data storage

ACCESS

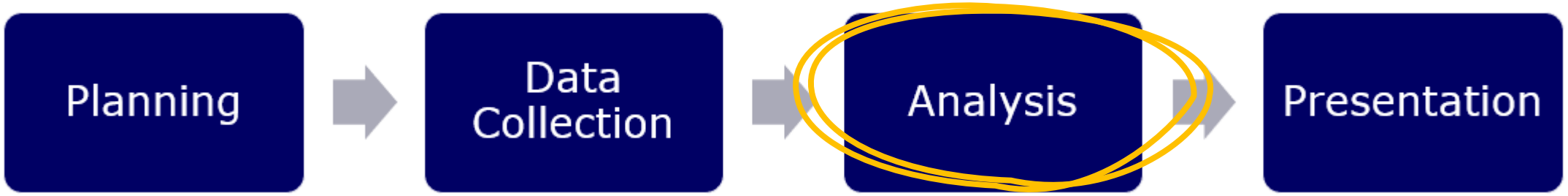
Useful for longer term projects where data will be reused

More powerful for doing queries across multiple tables

Can use data entry forms

Likely need support to set-up and query

HENRY FORD HEALTH®



How PHS & Biostatistics can help with data analysis

The Division of Biostatistics has MS and PhD trained biostatisticians to support data analysis

- Our goal is to use statistical techniques to summarize and analyze your data in a robust and meaningful way to meet your study objectives
- We will communicate these results with written reports, featuring tables and graphs
- We are available to meet via Teams/Webex or in-person at OFP or HFH

- Initial meeting – understand the project goals, explore the collected data, and set/review analysis plan
- Interim meetings and reports – to communicate data preparation steps and analysis results

Your collaborating biostatistician will help with the presentation of work

- Creating presentation-level tables and graphics from the selected work presented in reports
- Writing the statistical methods section
- Drafting an initial interpretation of data analysis results
- Reviewing the final product for accuracy of reporting and interpretation
- Helping to conduct revisions as requested by reviewers

Authorship Expectations

The Division of Biostatistics follows the recommendations of the ICJME for authorship

- https://www.icmje.org/recommendations/browse/roles-and-responsibilities/
- Authorship should be given to those who have “Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work”
- We do not accept acknowledgements
- Authors should be given opportunity to contribute to the writing and comment on the final draft before submission

- These expectations apply for both published abstracts and full manuscripts

Study Designs and other considerations

- Three most common study designs
  - Case control study
  - Cohort study
  - Randomized clinical trial

Case-Control Study

Study participants are enrolled into the study based on their disease status

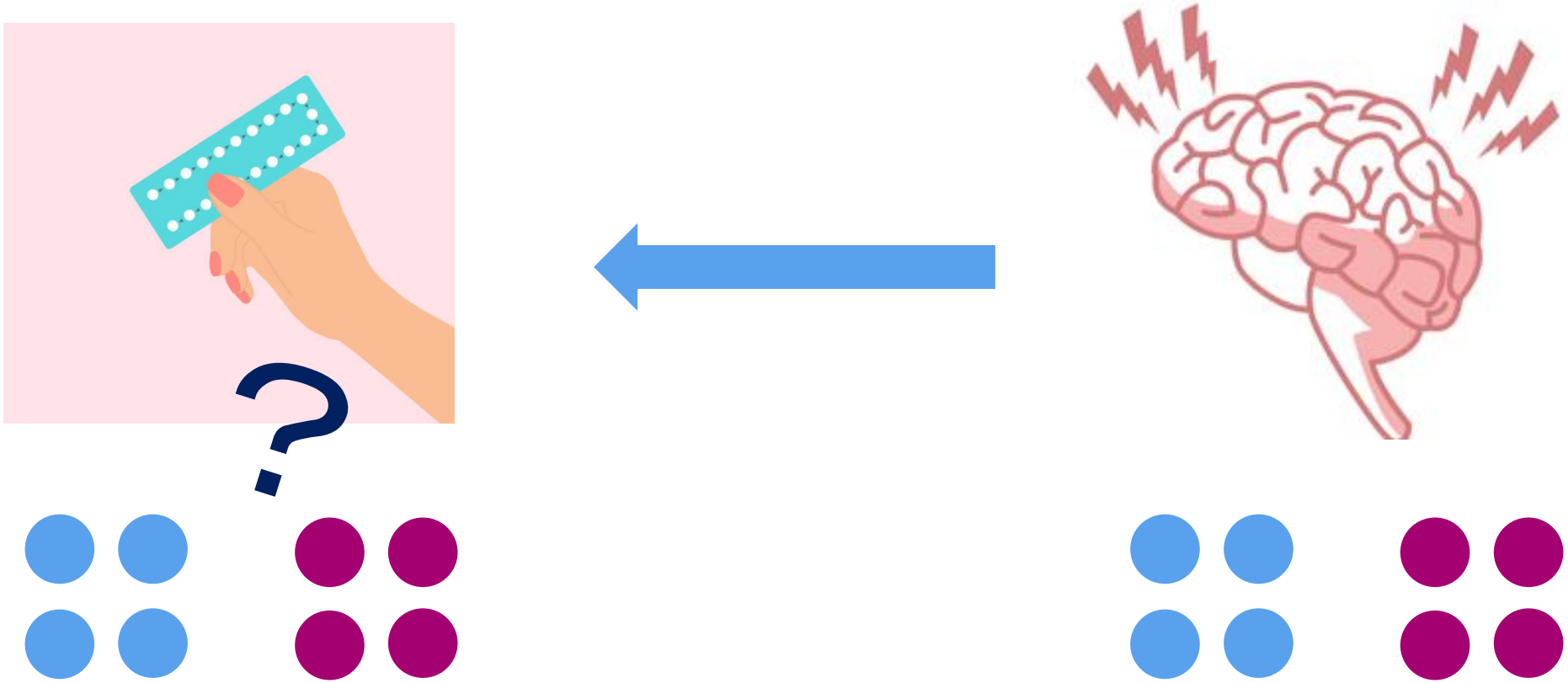
Cases – those with the disease

Controls – those without the disease

Look back in time to determine the characteristics of the study participants before the onset of disease

Observational study design

Case-Control Study



Cohort Study

Study participants are enrolled before their disease status is known

Study group – those with a characteristic or risk factor of interest

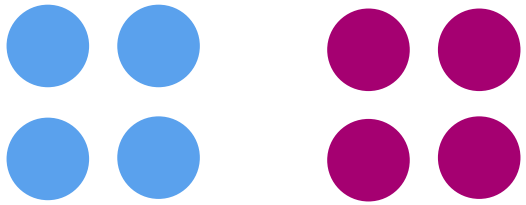
Control group – those without the risk factor of interest

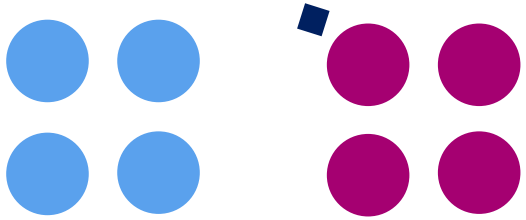
Subjects are then followed forward in time to see who does develops the disease

Observational study

Maybe prospective or retrospective

Cohort Study





Randomized Clinical Trial

Study participants are randomly assigned to a study or control group

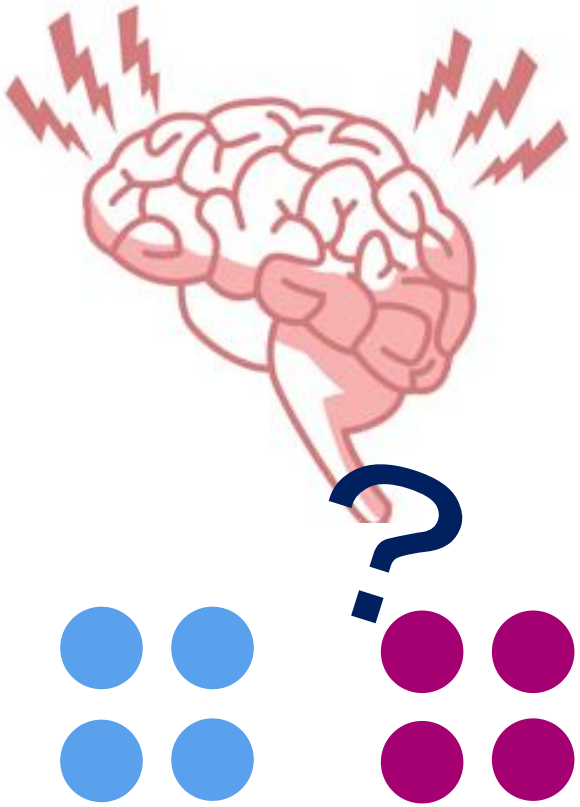
Subjects are then followed forward to see who develops the disease

Usually subjects, as well as investigators, are not aware of group assignment

- Single and/or Double Blind or Masking

Intervention study

Randomized Clinical Trial



What study design?

- **Case-Control Study**
  - Rare conditions or disease: using fewer individuals
  - Examine multiple variables associated with disease
- **Cohort Study**
  - Greater assurance that a particular characteristic preceded a particular outcome
  - Examine relationship of a risk factor and multiple outcomes
- **Randomized Clinical Trials**
  - Study interventions and its efficacy

Research Questions and Hypothesis

- A clear, focused **research question** guides the study.
  - **P** (Population): Who is being studied?
  - **I** (Intervention/Exposure): What is the treatment or risk factor?
  - **C** (Comparison): What is the control or comparison group?
  - **O** (Outcome): What is the desired effect or result?
  - "In elderly patients with hypertension (P), does a low-sodium diet (I) compared to no dietary changes (C) reduce the incidence of heart attack (O)?"
- A **hypothesis** is a testable statement based on your research question. It predicts the relationship between variables.
  - Null Hypothesis (H_0): No difference or effect.
  - Alternative Hypothesis (H_1): There is a difference or effect.

Outcomes or endpoints

- Have determined hypothesis to test, now what measure will you use to test it?
- **Outcomes** (or endpoints) are the specific events or measurements that the study is designed to assess. They help determine whether an intervention, exposure, or treatment had the desired effect.
  - Primary Outcome: The main result the study is designed to evaluate.
  - Secondary Outcomes: Additional effects explored in the study.
- **Examples** :
  - Clinical Trials: Symptom relief, recovery time, side effects.
  - Cohort Studies: Incidence of disease or complications over time.
  - Case-Control Studies: Presence or absence of conditions (e.g., infections).

Sample Size and Power

Testing level (alpha) usually set at 0.05

Power usually a minimum of 80%

- Power is the probability of demonstrating statistical significance if the study hypothesis is true.

Meaningful difference to detect

- May need some information about variability in the outcome measure

Sample size

One or two sided testing

- Usually do two sided

Example of “bad” data

	A	B	C	D	E	F	G
1						Baseline	Follow-up
2	Patient No	DOB	Sex	Race	Status	Creatinine	Creatinine
3	1	19-Aug-55	m	Black	Alive	None	None
4	2	4/23/1953	F	African American	Dead - 6/12/2008	<.5	>0.5
5	3	31/10/1942	male	White	Alive	55 mmol/L	55 mmol/L
6	Patient No	DOB	Sex	Race	Status	Creatinine	Creatinine
7	4	5.6.70	Male	Caucasian	A	N/A	N/A
8	5	12-Nov-32	f	A	Dead1/12/2009	2	1.9
9	6	9-Aug-52	female	??	D 9/25/2007	200 mmol/L	200 mmol/L
10	GroupA						
11	GroupB		The average systolic blood pressure for Group A was 125				
12			The average systolic blood pressure for Group B was 115				
13	AVG Creatinine						
14	Group A	0.6					
15	Group B	1.9					
16	TTEST	p = 0.028					

“Better” data

ID	Group	DOB	Sex	Race	Status	Death Date	Creatinine1	Creatinine2	SBP
1	1	8/19/1955	M	1	0				123
2	1	4/23/1953	F	1	1	6/12/2008	0.4	0.6	125
3	1	10/31/1942	M	2	0		0.8	0.8	127
4	0	5/6/1970	M	2	0				116
5	0	11/12/1932	F	3	1	1/12/2009	2	1.9	115
6	0	8/9/1952	F		1	9/25/2007	1.8	1.8	114

Group 1 = Treatment	DOB = Date of Birth
Group 0 = Control	SBP = Systolic Blood Pressure
Status 1 = Dead	
Status 0 = Alive	
Creatinine units = mg/dL	
Race 1 = Black	
Race 2 = White	
Race 3 = Middle Eastern	

When to contact PHS

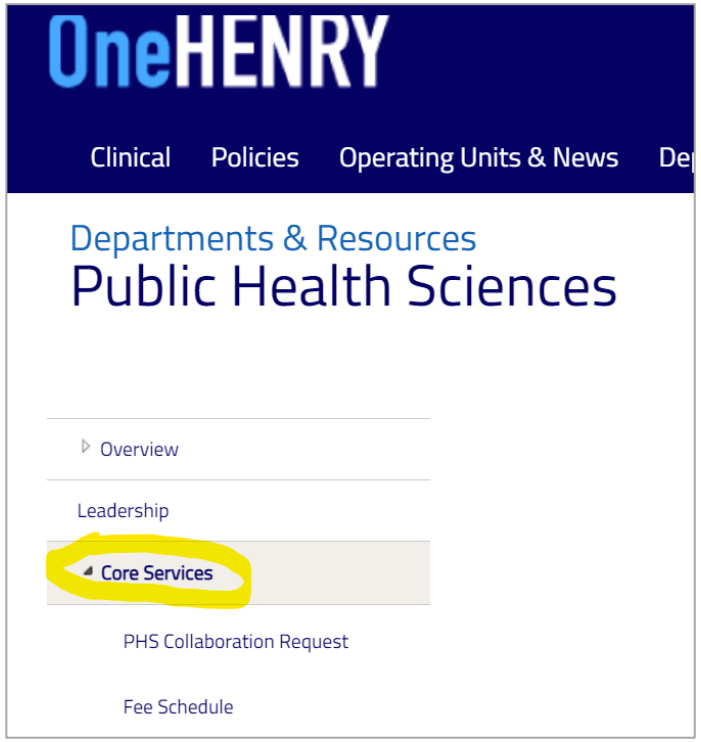
- The sooner the better!
- Can help with selecting most appropriate study design and patient population to avoid potential biases
- Give advice on designing data collection and data entry templates
- Doing these things before will save time and money when we start analysis

What PHS charges

- Do not charge for proposal and grant development
  - HFMG / Research office provides subsidy for this
- Graduate Medical Education (GME) does cover 10 hours of biostatistics support for residents and fellows, which can be renewed
  - Must complete the GME proposal and get signatures before submitting intake
- Hourly rates for other PHS services are computed to cover our costs
  - Can do on an hourly rate or percent effort from grant or cost center

Website References

https://onehenry.hfhs.org/departments/publichealthsciences/Pages/Core-Services.aspx



https://www.henryford.com/hcp/research/public-population-research/public-health-sciences

New projects?

https://onehenry.hfhs.org/departments/publichealthsciences/Pages/Overview.aspx

Contact

- Xuan Ma, PhD – Biostatistician II

  xma1@hfhs.org

- Yueren Zhou, MS – manager of biostatistics staff

  YZHOU1@hfhs.org

- Laila Poisson, PhD - Vice Chair of Biostatistics

  LPOISSO1@hfhs.org