2019-2

Cohort Studies and Relative Risks

Richard Ssekitoleko
Yale University

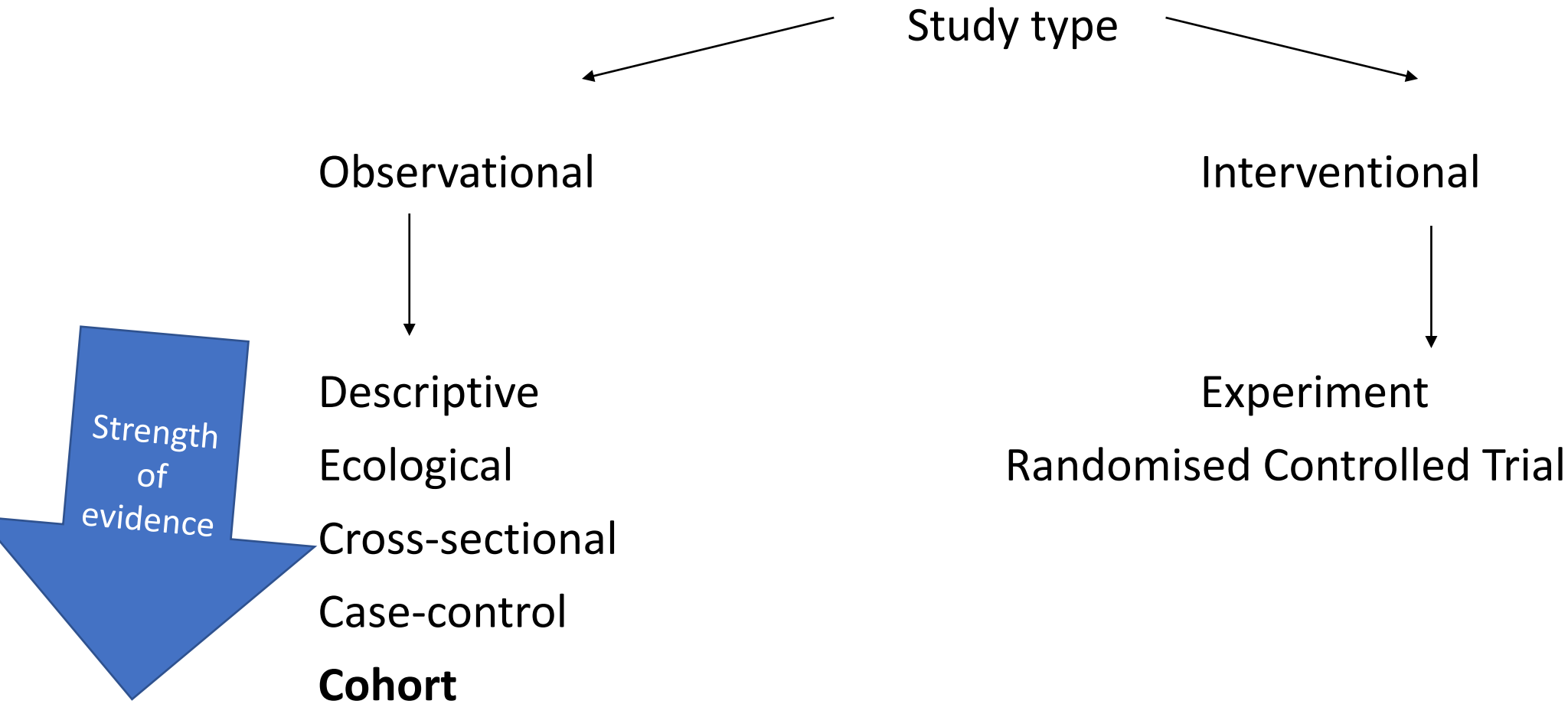
Cohort studies and Relative risks

Richard Ssekitoleko

Objectives

- Define a cohort study and the steps for the study
- Understand the populations in a cohort study
- Understand timing in a cohort study and the difference between retrospective, prospective and ambi-directional cohort studies
- Understand the selection of the cohort population and the collection of exposure and outcome data
- Understand the sources of bias in a cohort study
- Understand the calculation and interpretation of the relative risk
- Understand use of the new-castle Ottawa quality assessment score for cohort studies

Hierarchy of Evidence



Study designs and the basic principles

- Randomised controlled trial

- ***Cohort study*** *(Also called Longitudinal, follow up or Incidence study)*

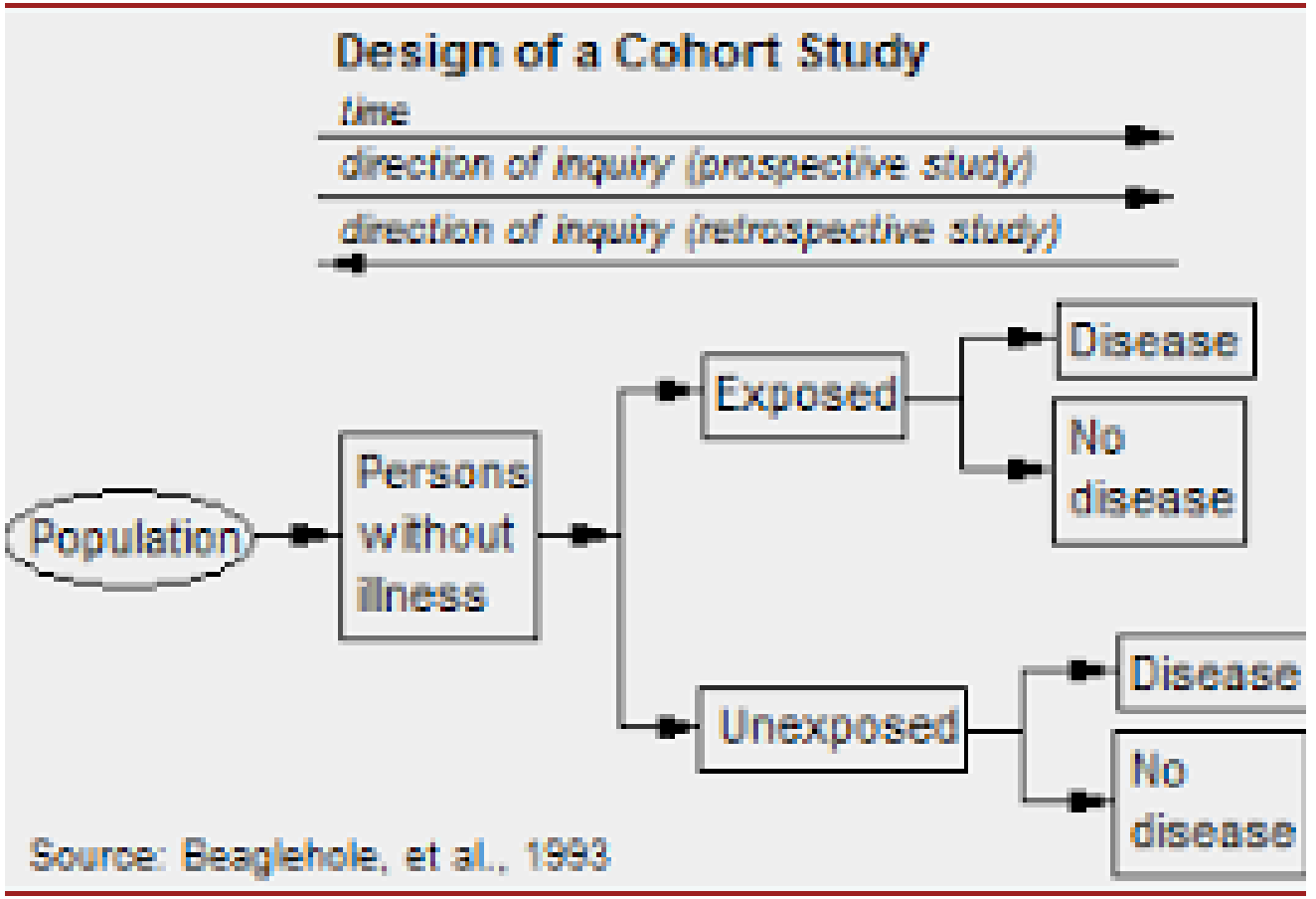
- Case-control study

- Cross-sectional study

- Ecological study

Cohort studies

- Observational study type
- Steps in a cohort study
  - Define the target population
  - Get a sample of the target population
  - Identify the exposure status of the sampled members
  - Follow up the members over time to identify new(Incident cases) of the disease (outcome)
  - Compare the risk of the outcome in those who are exposed at baseline to those who are not exposed(Risk Ratio)



Populations studied in Cohort studies

- Open/Dynamic cohort
  - Individuals may enter or leave at anytime
  - losses may occur
  - Defined by changeable characteristic
  - Measure incidence rate
- Fixed Cohort
  - Irrevocable event
  - Does not gain members/Losses may occur
  - Measure incidence rate
- Closed cohort
  - Irrevocable event
  - Does not gain members; no losses occur
  - Measure cumulative incidence

Timing of cohort studies

- Events in a cohort study defined by 3 terms
  - Prospective/ concurrent: Meaning to look forward in time
  - Retrospective: Meaning to look back in time
  - Ambi directional: Meaning to look both ways

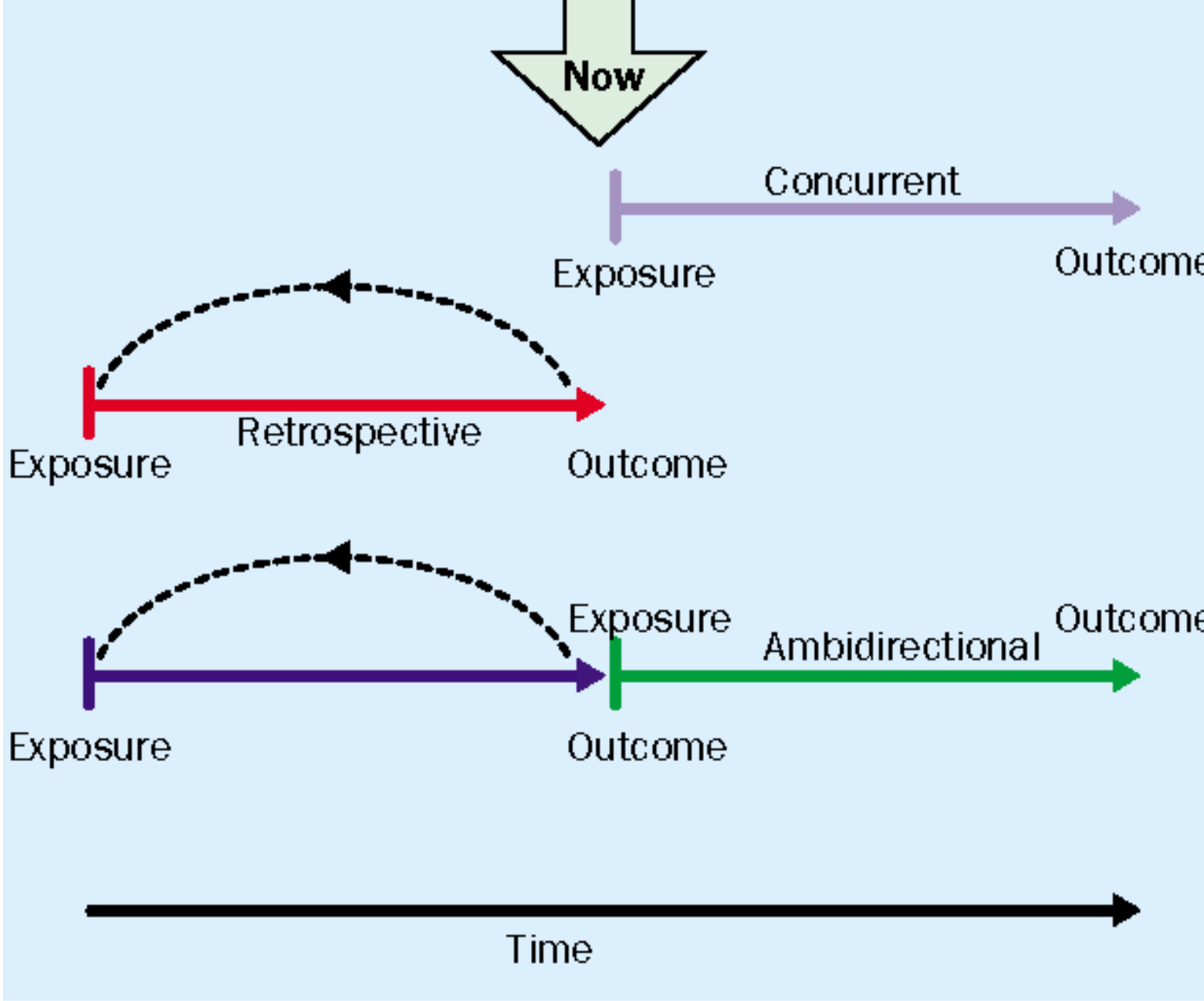


Figure 2: **Schematic diagram of concurrent, retrospective, and ambidirectional cohort studies** *Grimes et al. Lancet 2002;359:341-45*

BMJ Global Health

Derivation and validation of a universal vital assessment (UVA) score: a tool for predicting mortality in adult hospitalised patients in sub-Saharan Africa

Christopher C Moore,[1] Riley Hazard,[2] Kacie J Saulters,[3] John Ainsworth,[4] Susan A Adakun,[5] Abdallah Amir,[6] Ben Andrews,[7] Mary Auma,[6] Tim Baker,[8] Patrick Banura,[9] John A Crump,[10] Martin P Grobusch,[11] Michaëla A M Huson,[11] Shevin T Jacob,[12] Olamide D Jarrett,[13] John Kellett,[14] Shabir Lakhi,[15] Albert Majwala,[6] Martin Opio,[16] Matthew P Rubach,[17] Jamie Rylance,[18] W Michael Scheld,[1] John Schieffelin,[19] Richard Ssekitoleko,[5] India Wheeler,[18] Laura E Barnes[20]

BMJ Glob Health
2017;**2**:e000344. doi:10.1136/bmjgh-2017-000344

Retrospective cohort example

- Derivation and validation of a universal vital assessment score

- Methods: Pooled data from hospital based cohort studies from 2009 to 2015
- Analysis involved 5573 patients

- Exposure: Baseline UVA score

- Outcome: Inpatient mortality 996(17.3%)

- Temporal association between exposure and outcome clear

- By time study occurred both exposures and outcomes had occurred
- Lots of missing data with imputation → Information bias

- 2829 (50.8%) were female.
- Median (IQR) age was 36 (27–49) years
- The UVA score included points for temperature, heart and respiratory rates, systolic blood pressure, oxygen saturation, GCS and HIV serostatus.
- The UVA score had an area under the receiver operating characteristic curve (AUC) of 0.77 (95% CI 0.75 to 0.79)
- UVA score Outperformed other scoring systems (MEWS and qSOFA)
- UVA score could help with triage decisions in the study settings

Table 3 Population characteristics of patients pooled from hospital-based cohort studies conducted in six African countries from 2009 to 2015 according to outcomes

Variable	Overall (n=5573)	Survived (n=4607)	Died in-hospital (n=966)	p
Female, n (%)	2829 (50.8)	2323 (50.4)	506 (52.4)	0.27
Age (years), median (IQR)	36 (27–49)	36 (26–50)	36 (29–46)	0.20
Temperature (°C), median (IQR)	37.4 (36.3–38.5)	37.4 (36.3–38.5)	37.6 (36.2–38.4)	0.36
Heart rate (bpm), median (IQR)	100 (85–120)	100 (84–116)	110 (92–128)	<0.001
Respiratory rate (brpm), median (IQR)	26 (22–32)	24 (20–32)	30 (24–40)	<0.001
SBP (mm Hg), median (IQR)	100 (90–120)	100 (90–120)	90 (80–110)	<0.001
DBP (mm Hg), median (IQR)	62 (55–80)	65 (60–80)	60 (50–70)	<0.001
Oxygen saturation (%), median (IQR)	96 (94–98)	96 (94–98)	96 (93–98)	<0.001
GCS score, median (IQR)	15 (15–15)	15 (15–15)	15 (13–15)	<0.001
HIV-infected, n (%)	2122 (38.1)	1537 (33.4)	585 (60.1)	<0.001
CD4 (cells/µL), median (IQR)	72 (23–156)	79 (28–175)	46 (14–112)	<0.001
WBC (103/µL), median (IQR)	6.0 (3.7–9.7)	6.1 (3.9–9.7)	5.6 (3.1–9.7)	0.003
Haemoglobin (g/dL), median (IQR)	10.1 (7.8–12.2)	10.4 (8.1–12.5)	8.7 (6.9–11.2)	<0.001
Platelets (103/µL), median (IQR)	180 (105–270)	188 (113–275)	159 (81–256)	<0.001
Lactate (mmol/L), median (IQR)	3.5 (2.4–4.9)	3.3 (2.4–4.5)	4.2 (2.7–6.1)	<0.001

bpm, beats per minute; brpm, breaths per minute; CD, cluster of differentiation; DBP, diastolic blood pressure; GCS, Glasgow Coma Scale; SBP, systolic blood pressure; WBC, white blood cell concentration.

Hypoglycemia at admission is associated with in-hospital mortality in Ugandan patients with severe sepsis

Richard Ssekitoleko, MBChB, MMed; Shevin T. Jacob, MD, MPH; Patrick Banura, MBChB, MPH; Relana Pinkerton, PhD; David B. Meya, MBChB, Mmed; Steven J. Reynolds, MD, MPH; Nathan Kenya-Mugisha, MBChB; Harriet Mayanja-Kizza, MBChB, MS; Rose Muhindo, MBChB, MMed; Sanjay Bhagani, MBBS; W. Michael Scheld, MD; Christopher C. Moore, MD, FACP

Crit Care Med 2011 Vol. 39, No. 10

Prospective cohort example

- Prospective observational study on patients with Sepsis in 3 Ugandan hospitals.
- Analysis involved 418 admitted patients
- Exposure: Admission blood glucose concentration
- Outcome: In hospital mortality 113(27%)
- Measure of association: Hazard ratio
- Results: Significantly higher rates of mortality in patients with hypoglycemia: HR 95% CI 1.9(1.1-3.3)

Table 3. Univariate predictors of survival meeting ≤0.30 criteria and final multivariate model results using Cox regression

Outcome	Survived	Died	Hazard Ratio (95% Confidence Interval) and *p*			
			Univariate	*p*	Multivariate	*p*
Admission glucose concentration, n (%)						
Euglycemia (4.4–6.1 mmol/L)	113 (80.7)	27 (19.3)				
Hypoglycemia (<4.4 mmol/L)	44 (64.7)	24 (35.3)	2.0 (1.2–3.6)	.013	1.9 (1.1–3.3)	.03
Hyperglycemia (>6.1 mmol/L)	148 (70.5)	62 (29.5)	1.5 (0.96–2.4)	.08	1.6 (0.97–2.5)	.07
AMS, n (%)						
No AMS	274 (77.6)	79 (22.4)				
AMS	31 (47.7)	34 (52.3)	2.5 (1.6–3.7)	<.001	2.2 (1.5–3.4)	<.001
White blood cell count, n (%)						
≥4,000 to ≤12,000 cells/μL	162 (79.8)	41 (20.2)				
<4,000 or >12,000 cells/μL	136 (67.0)	67 (33.0)	1.7 (1.1–2.4)	.01	1.7 (1.1–2.5)	.013
Heart rate, n (%)						
≤90 beats/min	18 (85.7)	3 (14.3)				
>90 beats/min	286 (72.4)	109 (27.6)	1.9 (0.60–6.0)	.27	—	Not significant
Bacteremia or fungemia						
Negative	248 (74.7)	84 (25.3)				
Positive	57 (66.3)	29 (33.7)	1.3 (0.88–2.0)	.18	—	Not significant
Platelets, n (%)						
≥100,000 cells/μL	245 (79.5)	63 (20.5)				
<100,000 cells/μL	48 (52.2)	44 (47.8)	2.4 (1.6–3.5)	<.001	1.8 (1.2–2.7)	.007
Hospital site,[a] n (%)						
Mulago or Masaka	249 (77.3)	73 (22.7)				
Mbarara	56 (58.3)	40 (41.7)	2.4 (1.6–3.5)	<.001	1.8 (1.2–2.9)	.004

- Temporal association between admission blood glucose and mortality clear

- Researchers identified baseline exposures and then followed up patients (Prospective observational study)

Selecting the cohort population

- Based on study hypothesis
  - Guided by the exposure to be studied e.g smokers vs non smokers

- May be population based cohort based on common exposures e.g smoking, alcohol consumption, exercise and common chronic illnesses

- May be exposure based e.g occupational groups such as road builders

Collection of Exposure and Outcome data

- Study outcomes have not occurred at the beginning of the cohort follow up period.

- Exposures of interest may vary during the study period.
  - May be present at the beginning
  - May occur during the study
  - May stop during the study period

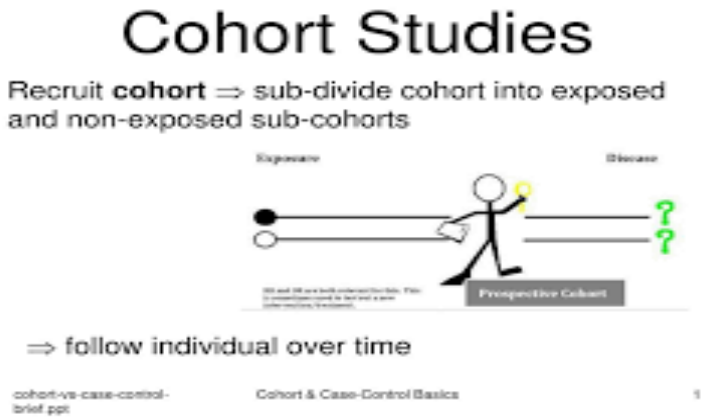
- Temporal association between exposure and outcome is clear

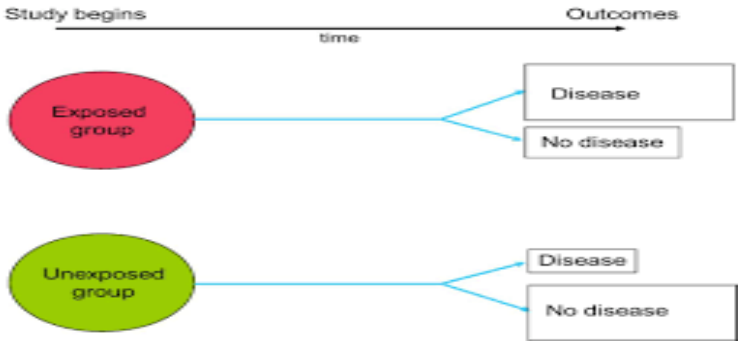
Study population

- At the beginning of follow up all cohort members should be alive not have the outcome of interest

- All members should be at risk of getting the outcome of interest
  - E.g In a study of women involving an outcome of uterine ca, one cannot include women who had a hysterectomy at baseline

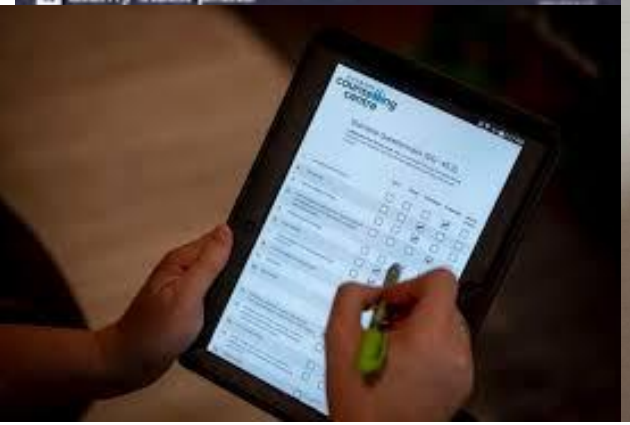
Comparison populations in cohort studies

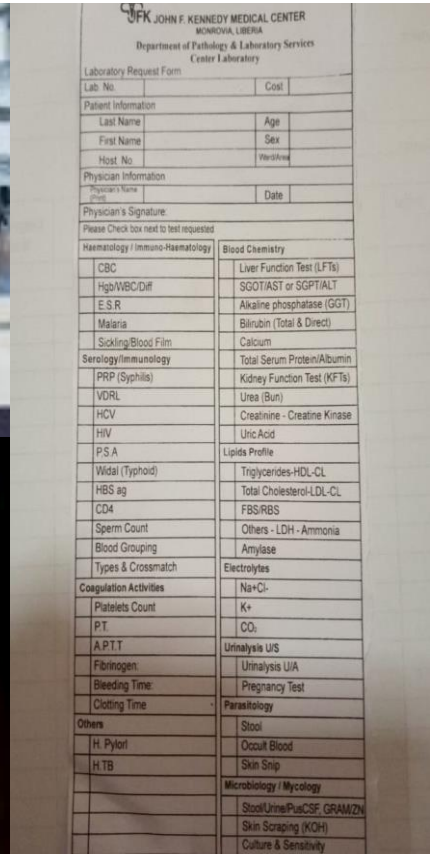




- Single cohort
  - Members of a single population are classified by levels of exposure
  - Comparison group is unexposed or less exposed group
  - Need to account for confounding factors
  - Common exposures: alcohol, smoking, exercise
- Double cohort
  - Involves an exposed population and an unexposed population
  - Comparison group may be the general population

Sources of exposure information





JFK JOHN F. KENNEDY MEDICAL CENTER
MONROVIA, LIBERIA
Department of Pathology & Laboratory Services
Center Laboratory
Laboratory Request Form

Lab. No.		Cost	
Patient Information			
Last Name		Age	
First Name		Sex	
Host No.		Ward/Area	
Physician Information			
Physician's Name (Print)		Date	
Physician's Signature:			
Please Check box next to test requested			

Haematology / Immuno-Haematology		Blood Chemistry	
	CBC		Liver Function Test (LFTs)
	Hgb/WBC/Diff		SGOT/AST or SGPT/ALT
	E.S.R		Alkaline phosphatase (GGT)
	Malaria		Bilirubin (Total & Direct)
	Sickling/Blood Film		Calcium
Serology/Immunology			Total Serum Protein/Albumin
	PRP (Syphilis)		Kidney Function Test (KFTs)
	VDRL		Urea (Bun)
	HCV		Creatinine - Creatine Kinase
	HIV		Uric Acid
	P.S.A	Lipids Profile	
	Widal (Typhoid)		Triglycerides-HDL-CL
	HBS ag		Total Cholesterol-LDL-CL
	CD4		FBS/RBS
	Sperm Count		Others - LDH - Ammonia
	Blood Grouping		Amylase
	Types & Crossmatch	Electrolytes	
Coagulation Activities			Na+Cl-
	Platelets Count		K+
	P.T.		CO_2
	A.P.T.T	Urinalysis U/S	
	Fibrinogen:		Urinalysis U/A
	Bleeding Time:		Pregnancy Test
	Clotting Time	Parasitology	
Others			Stool
	H. Pylori		Occult Blood
	H.TB		Skin Snip
		Microbiology / Mycology	
			Stool/Urine/PusCSF, GRAM/ZN
			Skin Scraping (KOH)
			Culture & Sensitivity

- Data may be collected routinely during follow up period
- Sources of data
  - Participant interviews
  - Monitoring data from home or workplace
  - Laboratory monitoring
  - Medical records

Outcome data in a cohort study

- Reports of symptoms and signs
- Medical assessment results
- Medical records
- Disease registry results
- Medical examination results
- Death certificates

JFK JOHN F. KENNEDY MEDICAL CENTER
P. O. BOX 1973
MONROVIA, LIBERIA

DEATH CERTIFICATE

Deceased Last Name	First Name	Middle Name	Hospital No.	Age	Sex	Date of Death / Hour:

Address	PLACE OF DEATH JOHN F. KENNEDY MEDICAL CENTER

Place of Birth:	Citizenship:	MARITAL STATUS				
		M	S	W	D	SEP

CAUSE OF DEATH

The certified cause of death (has, has not) been confirmed by post mortem.
Date last seen alive by me:_____

Death was caused by: (condition, if any which give rise to immediate cause, stating the underlying cause last)	Give appropriate interval between onset and death.
(A)	

Due to, or as a consequence of:
(B)

Other significant conditions contributing to death, but not related to the cause of death given above.

I,_____ M.D., do hereby certify that I have medically and surgically attended that above named deceased and that the particulars and cause of death written above are abstracts from the deceased Medical record and are true to the best of my knowledge and belief.

_____ M.D.

MEDICAL BOARD LICENSE NO.:_____

Date certificate was written:_____

NOTE: This document is not valid unless it bears the MEDICAL RECORD STAMP

Follow up in a cohort study

- All participants need to be tracked throughout the study
  - To get their true outcome
  - To get their person time contribution to the study
- Loss to follow up is a form of bias and reduces validity of results
  - Decreased sample size reducing ability of the study to detect an association if present
  - Those lost to follow up may differ in important ways from those who stay
- May occur due to death, change of residence, migration or participant decision to stop taking part in the study

Minimizing loss to follow up in a cohort study

- Explain need to follow up with participants at start
- Get contact details for participant, friends, relatives or physician
- Maintain regular follow up (Mail, phone or personal contact)
- Follow up on non responses and disappearances promptly
- Offer incentives for follow up e.g transport refund

Analysis in a cohort study

Contingency (or 2 x 2) Table

	Cases	Controls	Total
Exposed	a	b	a+b
Unexposed	c	d	c+d
Total	a+c	b+d	a+b+c+d

- Need to calculate incidence in the exposed and unexposed groups
- May calculate cumulative incidence or incidence density rate depending on the available information
- Comparing incidence in the exposed and unexposed groups will enable estimation of the relative risk

Risk Ratio (relative risk)

$$\frac{\text{RISK of outcome occurrence in exposed}}{\text{RISK of outcome occurrence in unexposed}} = \text{Risk Ratio (RR)}$$

RR > 1 suggests exposure predisposes to outcome

RR < 1 suggests exposure protects against outcome

RR = 1 is null and indicates no association between exposure and outcome

Interpreting the relative risk

- Gives the strength of association between the exposure and outcome
- May not be causal
- Could be explained by random error, confounding or bias
- May represent the cumulative incidence ratio or the incidence density ratio depending on how it is calculated

Risk Ratio Calculations

If, after follow up, the following is seen:

Exposure	Disease: Yes	Disease: No	Total
Yes	d_1	h_1	$n_1 = d_1 + h_1$
No	d_0	h_0	$n_0 = d_0 + h_0$
Total	$d = d_1 + d_0$	$h = h_1 + h_0$	$n = d + h$

Then, Simple Cumulative incidence(risk ratio(RR) = $$\frac{\text{risk in exposed}}{\text{risk in unexposed}} = \frac{d_1 / n_1}{d_0 / n_0}$$

Simple Cumulative incidence example

- The table below summarizes a population of 1000 subjects with respect to a particular disease D broken down by sex

	Disease		
	D	$\bar{D}$	Total
Men	140	60	200
Women	180	620	800
Total	320	680	1000

- What is the relative risk of getting the disease associated with being a man as opposed to being a woman?

Relative risk= Risk of disease in men/Risk in women=(140/200)/(180/800)=0.7/0.225=3.1

Odds ratios and risk ratios

- How do you interpret the relative risk?

  The risk of getting the disease in males is 3.1 times the risk of getting the disease in females

- What is the odds ratio for the disease among men as opposed to women?

  Odds of the disease in men: Odds=Risk of disease in men/risk of no disease in men= (140/200)/(60/200)=0.7/0.3=2.3

  Odds of the disease among women: Probability of disease in women/Probability of no disease (180/800)/(620/800)=0.29

  The odds ratio for disease associated with being a man as opposed to a woman. Odds ratio= Odds in men/Odds in women= 2.3/0.29=7.93

- In which type of study is the odds ratio the preferred measure of association?
- Compare the risk ratio to the odds ratio. What do you conclude?

Incidence density

- Person-time at risk
  - Length of time for each individual that they are in the population at risk
  - Sum of person time for each individual during their stay in study is the total person-time
- When a person is no longer at risk, they no longer contribute to person time e.g when they get the outcome
- Incidence density
  - Rate of occurrence of new cases of disease during person time of observation in a population at risk of getting the disease
  - Numerator =Number of new cases of disease
  - Denominator=Total person time of observation in population at risk
- A rate and the units are Inverse time (1/time)

Incidence density ratio

- Incidence density ratio= Incidence density in exposed group/Incidence density in unexposed group

	Total person time of observation	Number of persons with outcome	Incidence density
Exposed group	A	C	C/A
Unexposed group	B	D	B/D

- Incidence density ratio= (C/A)/(B/D)

Incidence rate ratio example

A study examined mortality among homeless shelter residents in New York City from 1987 to 1994. There were 15 deaths observed among women aged 25-34, with 728 person-years of observation. Among men aged 25-34, 31 deaths were observed, with 1988 person-years of observation. (Am J Public Health. 1999 Apr;89(4):529-34).

	Death	Person-time
Women	15	728
Men	31	1988

- The measure of relative risk appropriate for this data is the Incidence density ratio.
  - Incidence density =number of new cases/total person time at risk

- The relative risk of mortality among women aged 18-24 compared to men aged 18-24 is the incidence density ratio and is given by:

  Incidence density women/Incidence density men

  =(15/728)/(31/1988)=1.32

- Interpretation
  - The rate of mortality among women was 1.32 times the rate of mortality among men in New York City between 1987 and 1994.
- What is the difference between the Incidence density ratio and the cumulative incidence ratio?

- How do we get the person time?

Limitations of cohort studies

- Measurement error (A form of information bias)
  - Commonly errors in exposure measurement
  - Errors in outcome assessment(People may die from competing risks, actual onset of the disease may be missed)
- Confounding- Occurs when a factor is causally associated with both the outcome and exposure under study
- Selection bias (To the different groups and loss to follow up)
- Loss to follow-up (A form of selection bias)
  - If it is related to the exposure or outcome of interest
  - May be differential or non-differential

Loss to follow-up

- A problem with cohort studies is loss to follow-up
- Loss to follow-up may be non-differential i.e. not related to exposure and outcome
- Or differential i.e. is related to exposure and/or outcome. e.g. subjects with poor education who contract HIV die very quickly and do not present to health centres or hospitals. Affects the measure of effect

Cohort Studies

Advantages

- Clear temporal relationship: between exposure and outcome (Compare cross sectional studies)
- Good for rare exposures
- Can evaluate multiple effects of an exposure
- Can minimise biases in exposure measurement
- Directly measures disease incidence or risk

Disadvantages

- Usually expensive and Time consuming (Prospective)
- Poor information on exposures and other key variables (Retrospective)
- Inefficient for disease with long induction and latent periods(Prospective)
- Bias/ confounding
- Changes over time can affect exposure and disease classification

Critical review for cohort studies

J Korean Med Assoc. 2011 Apr;54(4):419-429.

https://doi.org/10.5124/jkma.2011.54.4.419

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort
   a) truly representative of the average _____ (describe) in the community ★
   b) somewhat representative of the average _____ in the community ★
   c) selected group of users eg nurses, volunteers
   d) no description of the derivation of the cohort
2) Selection of the non exposed cohort
   a) drawn from the same community as the exposed cohort ★
   b) drawn from a different source
   c) no description of the derivation of the non exposed cohort
3) Ascertainment of exposure
   a) secure record (eg surgical records) ★
   b) structured interview ★
   c) written self report
   d) no description
4) Demonstration that outcome of interest was not present at start of study
   a) yes ★
   b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis
   a) study controls for _____ (select the most important factor) ★
   b) study controls for any additional factor ★ (This criteria could be modified to indicate specific control for a second important factor.)

Outcome

1) Assessment of outcome
   a) independent blind assessment ★
   b) record linkage ★
   c) self report
   d) no description
2) Was follow-up long enough for outcomes to occur
   a) yes (select an adequate follow up period for outcome of interest) ★
   b) no
3) Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for ★
   b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost) ★
   c) follow up rate < ____% (select an adequate %) and no description of those lost
   d) no statement


WELCOME TO
JFK MEMORIAL
HOSPITAL

References

- Essentials of Epidemiology in Public health: Ann Aschengrau and George R Seage
- Introduction to the field of Statistics: David S Moore, George P McCabe and Bruce A Craig.

Acknowledgements

- Jack Colford: School of Public Health. University of California, Berkeley
- Michael Bates: School of Public Health. Univeristy of California, Berkeley.
- Sarah Lewis: Department of Social Medicine. University of Bristol UK.