Research Course

October 23, 2024

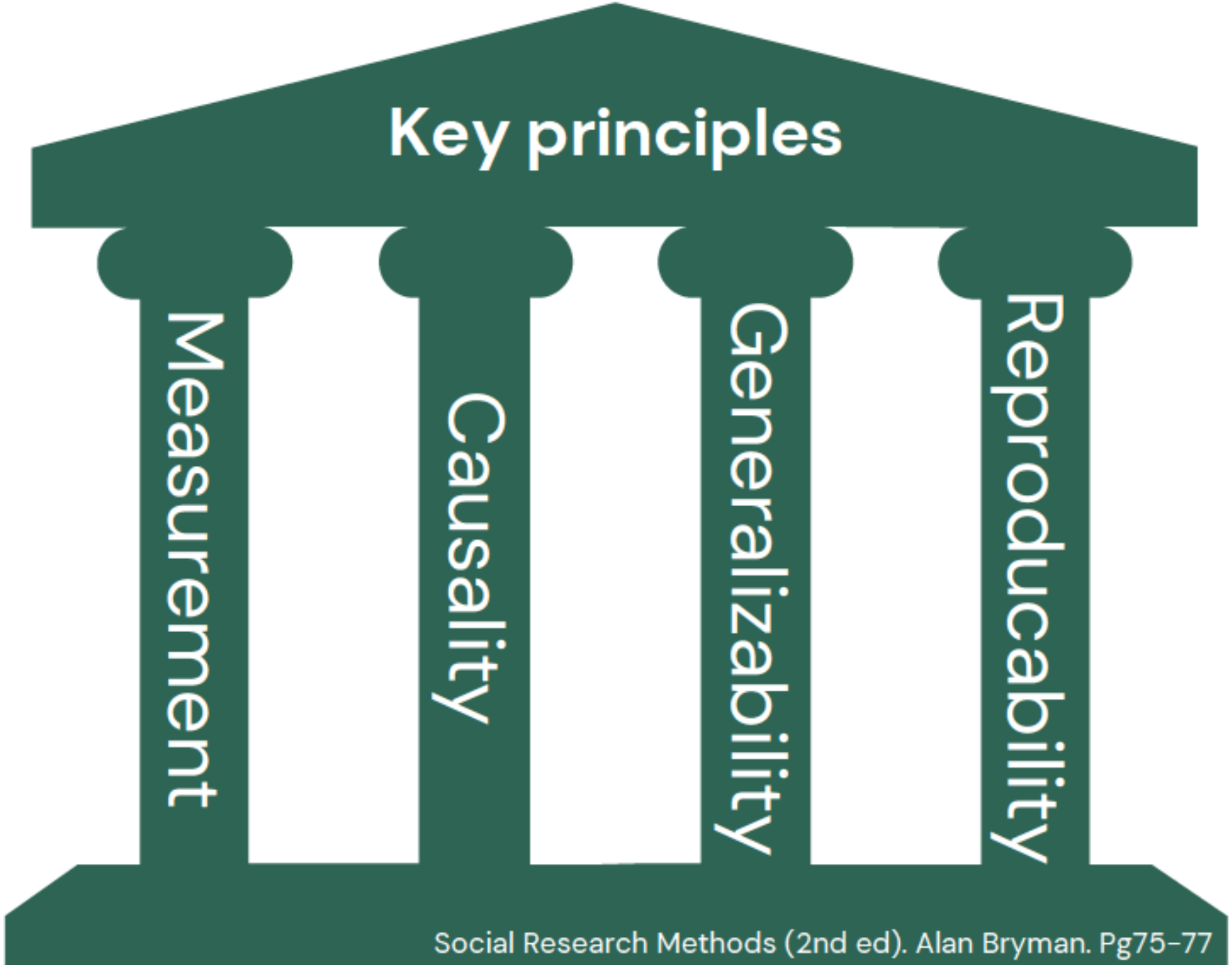
Study Design

Obervational & more

Khara Sauro
Associate Professor
Surgery | Community Health Sciences | Oncology
Cumming School of Medicine | University of Calgary

Learning Objectives

1. Be able to differentiate different types of observational study designs
2. Know when to use different study design based on research question
3. Critically appraise studies based on study design



Social Research Methods (2nd ed). Alan Bryman. Pg75-77

Before we talk about study design, let's review types of Bias...

1. Selection
2. Measurement/ Information
3. Confounding

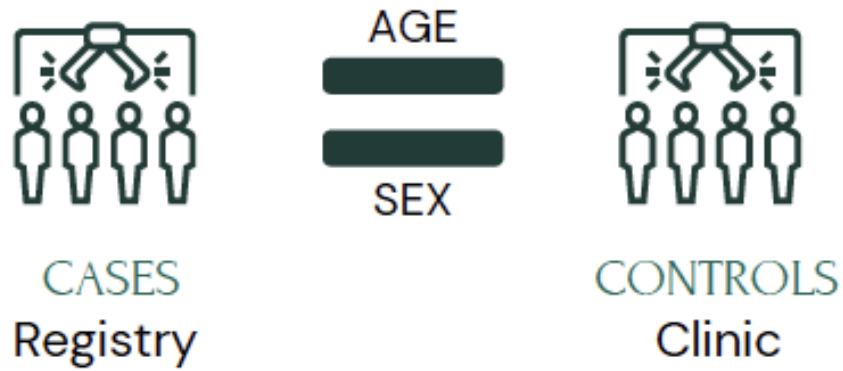
Types of Bias - Selection

1

Selection

A systematic difference in how participants are classified or chosen

Example



Types of Bias - Measurement

VALIDITY

RELIABILITY

Types of Internal Validity

1

FACE

Does the indicator make intuitive sense?

How to measure:
Survey or consensus among experts. No statistical test.

2

CONTENT

Degree to which instrument measures depth & breadth of construct or concept.

How to measure:
Survey or consensus among experts. No statistical test.

3

CRITERION

Degree to which measure relates to a criterion. Predictive.

How to measure:
Statistical agreement (e.g. kappa, correlation)

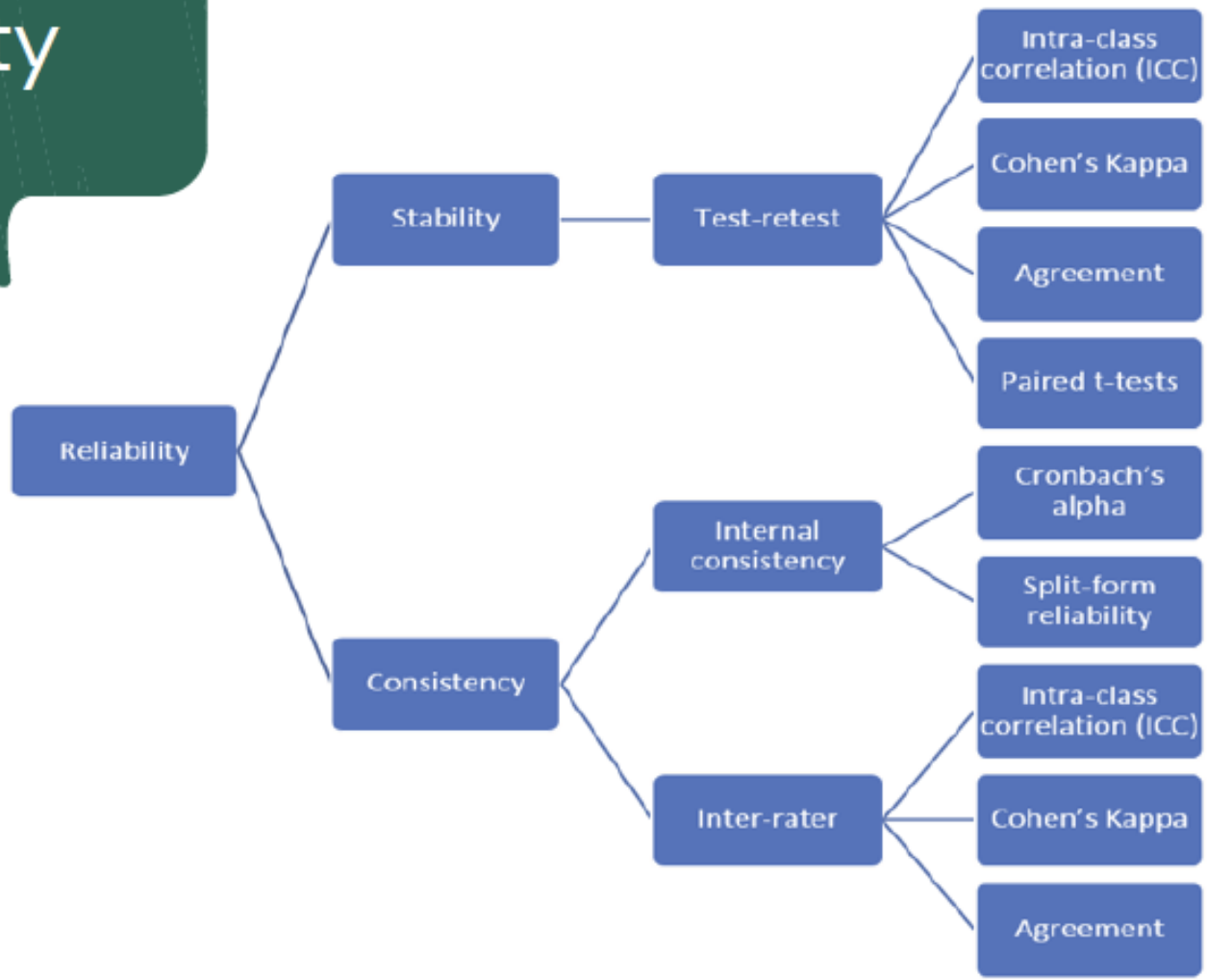
4

CONSTRUCT

Degree to which measure relates to other variables within a system/theory

How to measure:
Statistical measures of association

Reliability



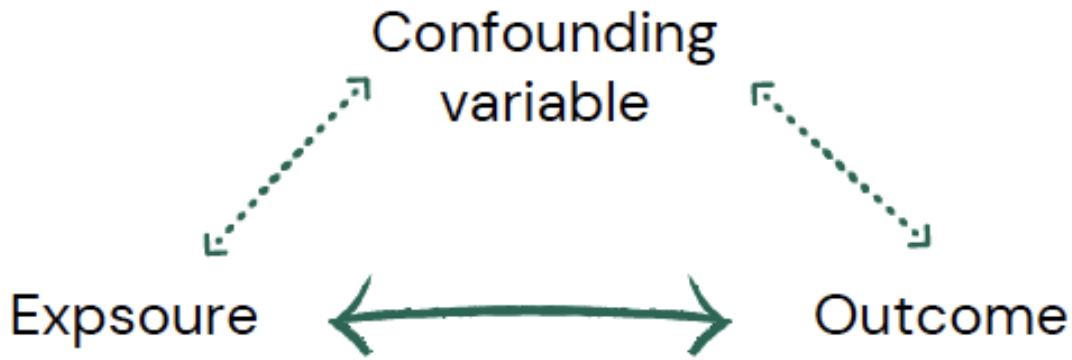
Slide used with permission from: Dr. Gavin McCormack. MDCH681 2020

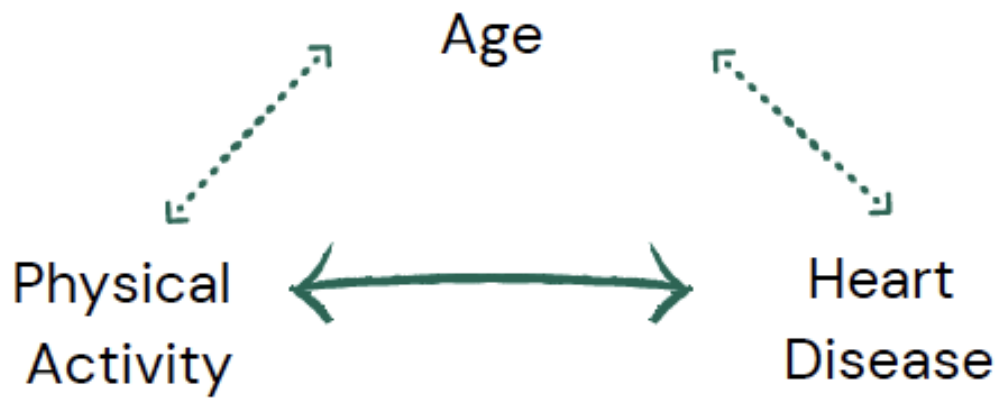
Types of Bias - Confounding

3

Confounding

Exposure and outcome are both associated with a third variable (confounder).





Mitigating Risk of Bias

1

Study Design

Randomization
Restricting
Matching

2

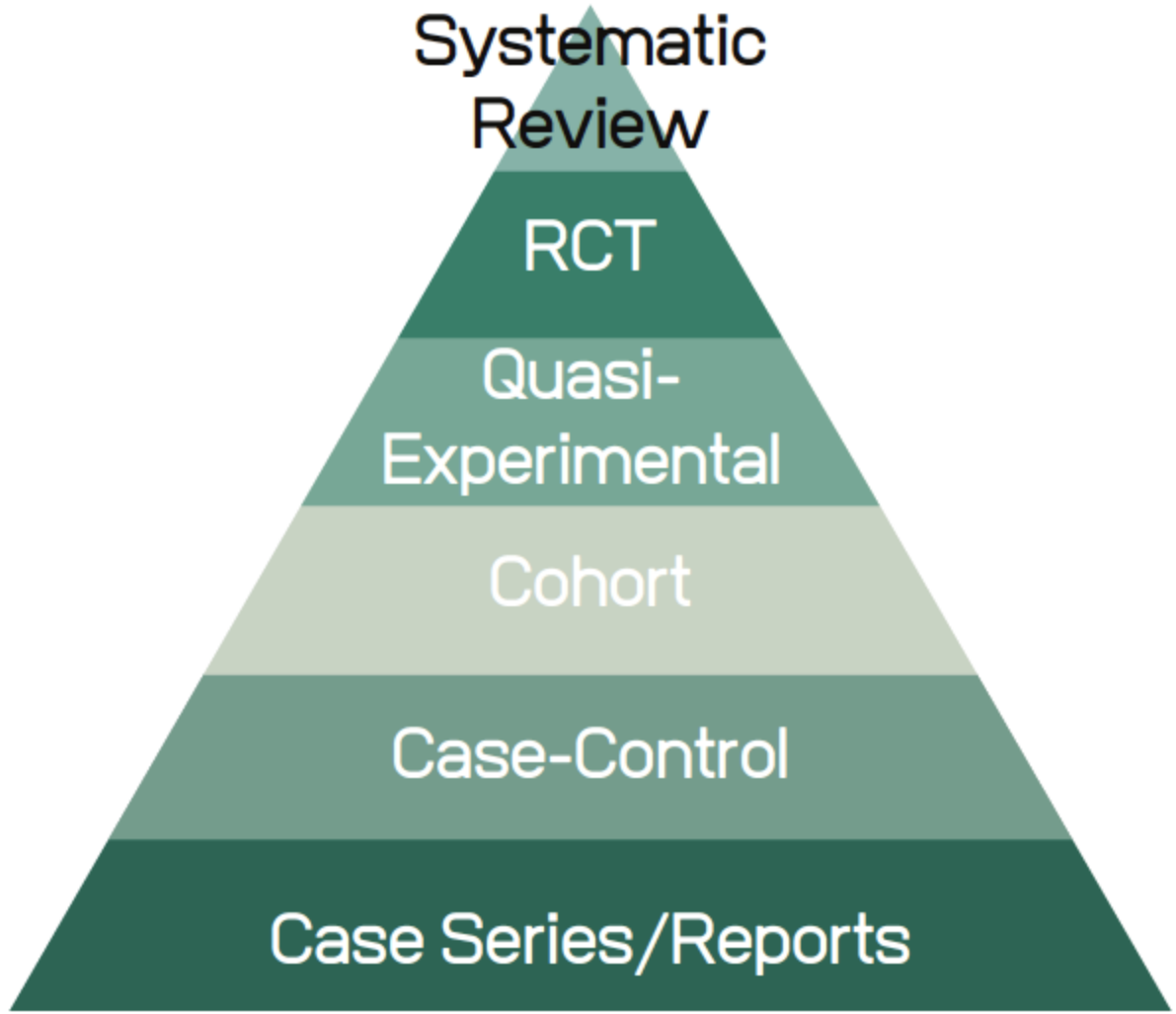
Measurement

Blinding
Standardizing
Valid
Reliable

3

Statistical

Stratification
Modeling
Matching


Systematic Review
RCT
Quasi-Experimental
Cohort
Case-Control
Case Series/Reports

Cohort and Case Control Studies

Additional Considerations, Reporting Guidelines, and Critical Appraisal

Some Additional Considerations

Advantages and Disadvantages

Examples

Classifying Exposures and Outcomes

Biased selection of individuals for your study

Cohort Studies: Advantages

- Clear temporal sequence, we know the exposure happened before the outcome as everyone start off outcome free
  - An important feature for determining causation
- Can study multiple effects of the same exposure
- Rare exposures can easily be studied
- Can truly measure risk as all individuals begin without the outcome of interest

Cohort Studies: Disadvantages

- Can be very time consuming and expensive
  - Can end up following people for decades
- Loss to follow up: whenever you are following individuals through time there is a chance you lose contact with some of the participants
  - It is important to minimize this as much as possible to avoid biased results (attrition bias)
- Potential for outcome misclassification if there are major advances in disease detection during the follow up period
- Difficult to study rare outcomes

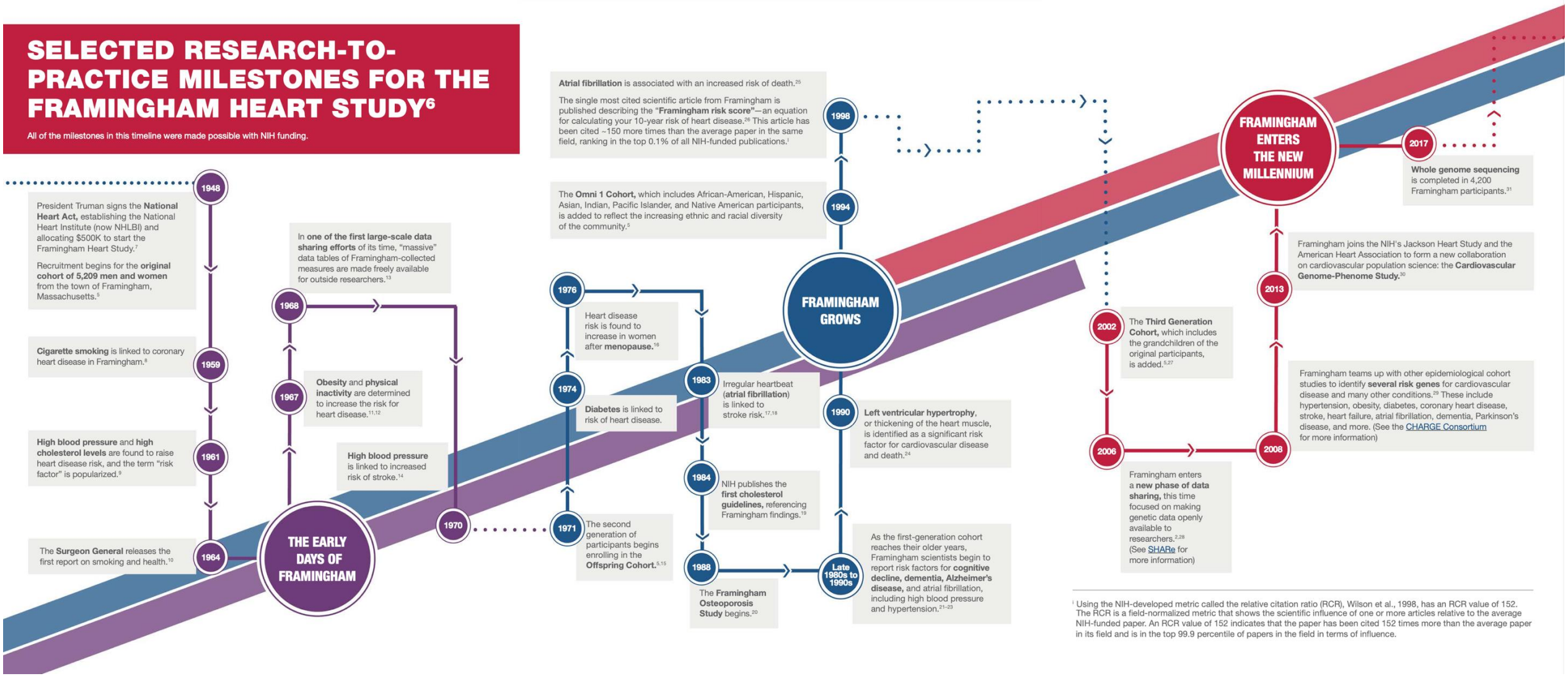
Example: Framingham Heart Study (Multipurpose Longitudinal Cohort)

WHAT IS THE FRAMINGHAM HEART STUDY?[3]

The study, which aimed to unravel the underlying causes of heart disease, started in 1948 with 5,209 participants in the town of Framingham, Massachusetts. Framingham is a longitudinal cohort study, a type of epidemiological study that follows a group of individuals over time to determine the natural history of certain diseases, explore the behavior of those diseases, and identify the factors that might explain their development. Part of the reason Framingham, Massachusetts was picked as the study site was because it was just big enough to provide a sufficient number of individuals for the study, while also small enough to be suited to the community approach of recruiting and effectively following participants over time.[4,5] Participants underwent physical examinations, gave blood samples for laboratory tests, and provided lifestyle and medical history information at regular intervals. Now a joint project of the NHLBI and Boston University, Framingham has expanded over the years, both in geographical and population scope. Today it includes many grandchildren and spouses in three generations of participants, as well as two cohorts of minority participants (the Framingham Omni Cohorts).

SELECTED RESEARCH-TO-PRACTICE MILESTONES FOR THE FRAMINGHAM HEART STUDY[6]

All of the milestones in this timeline were made possible with NIH funding.



[i] Using the NIH-developed metric called the relative citation ratio (RCR), Wilson et al., 1998, has an RCR value of 152. The RCR is a field-normalized metric that shows the scientific influence of one or more articles relative to the average NIH-funded paper. An RCR value of 152 indicates that the paper has been cited 152 times more than the average paper in its field and is in the top 99.9 percentile of papers in the field in terms of influence.

Circulation
Volume 98, Issue 10, 8 September 1998; Pages 946-952
https://doi.org/10.1161/01.CIR.98.10.946

CLINICAL INVESTIGATION AND REPORTS

Impact of Atrial Fibrillation on the Risk of Death

The Framingham Heart Study

Emelia J. Benjamin, Philip A. Wolf, Ralph B. D'Agostino, Halit Silbershatz, William B. Kannel, and Daniel Levy

ABSTRACT: *Background*—Atrial fibrillation (AF) causes substantial morbidity. It is uncertain whether AF is associated with excess mortality independent of associated cardiac conditions and risk factors.*Methods and Results*—We examined the mortality of subjects 55 to 94 years of age who developed AF during 40 years of follow-up of the original Framingham Heart Study cohort. Of the original 5209 subjects, 296 men and 325 women (mean ages, 74 and 76 years, respectively) developed AF and met eligibility criteria. By pooled logistic regression, after adjustment for age, hypertension, smoking, diabetes, left ventricular hypertrophy, myocardial infarction, congestive heart failure, valvular heart disease, and stroke or transient ischemic attack, AF was associated with an OR for death of 1.5 (95% CI, 1.2 to 1.8) in men and 1.9 (95% CI, 1.5 to 2.2) in women. The risk of mortality conferred by AF did not significantly vary by age. However, there was a significant AF-sex interaction: AF diminished the female advantage in survival. In secondary multivariate analyses, in subjects free of valvular heart disease and preexisting cardiovascular disease, AF remained significantly associated with excess mortality, with about a doubling of mortality in both sexes.*Conclusions*—In subjects from the original cohort of the Framingham Heart Study, AF was associated with a 1.5- to 1.9-fold mortality risk after adjustment for the preexisting cardiovascular conditions with which AF was related. The decreased survival seen with AF was present in men and women and across a wide range of ages.

Key Words: fibrillation, atrial ▪ mortality ▪ prognosis ▪ stroke ▪ cerebrovascular disorders ▪ risk factors ▪ aging

Cohort Study Example

Case Control Studies: Advantages

- Generally, less resource intensive than a cohort study
- Can study rare outcomes in an efficient manner
- Allow for multiple exposures to be studied at the same time

Case Control Studies: Disadvantages

Hard to study rare exposures

Can only study one outcome at a time

Can be difficult to determine if the exposure truly happened before the outcome

Did the outcome appear today, or did it go undiagnosed until today?

Subject to selection and recall bias

Cohort vs. Case-Control Studies

	Forward Directionality (exposure to outcome)	Backwards Directionality (outcome to exposure)
Retrospective	Retrospective Cohort Study	Case-Control Study
Prospective	Prospective Cohort Study	-

Example

GLOBAL AND POPULATION HEALTH

INTERHEART

Global Risk Factors for Acute Myocardial Infarction

SHARE:

COMPLETED

Homepage > Research > Global and Population Health > INTERHEART

The INTERHEART study found that nine easily measurable and modifiable risk factors could explain more than 90 per cent of the risk of a heart attack globally and in all regions and major ethnic groups of the world.

This landmark study emphasized that avoidance of tobacco, daily consumption of fruits and vegetables and regular exercise could potentially avoid two-thirds of heart disease.

The INTERHEART results also indicated that the two most important risk factors for myocardial infarction (MI) globally are:

- Tobacco: Smoking even one cigarette per day increases the risk of MI by five per cent.
- Abnormal lipids (fats in the blood).

As well, INTERHEART found that the markers of abdominal obesity and hip size (waist-to-hip-ratio) are far more predictive than body mass index (BMI) in predicting MI. Furthermore, stress and psychosocial factors were found to be important risk factors for MI.

INTERHEART - DOWNLOAD PDF

STUDY TYPE
Observational

STUDY DESIGN
Case-control

NO. OF COUNTRIES
52

NO. OF SITES
262

NO. OF PARTICIPANTS
29972

STUDY PERIOD
1999-2003

SPONSOR
PHRI

Case-Control Study Example

Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study

*Salim Yusuf, Steven Hawken, Stephanie Ôunpuu, Tony Dans, Alvaro Avezum, Fernando Lanas, Matthew McQueen, Andrzej Budaj, Prem Pais, John Varigos, Liu Lisheng, on behalf of the INTERHEART Study Investigators**

Summary

Background Although more than 80% of the global burden of cardiovascular disease occurs in low-income and middle-income countries, knowledge of the importance of risk factors is largely derived from developed countries. Therefore, the effect of such factors on risk of coronary heart disease in most regions of the world is unknown.

Methods We established a standardised case-control study of acute myocardial infarction in 52 countries, representing every inhabited continent. 15 152 cases and 14 820 controls were enrolled. The relation of smoking, history of hypertension or diabetes, waist/hip ratio, dietary patterns, physical activity, consumption of alcohol, blood apolipoproteins (Apo), and psychosocial factors to myocardial infarction are reported here. Odds ratios and their 99% CIs for the association of risk factors to myocardial infarction and their population attributable risks (PAR) were calculated.

Findings Smoking (odds ratio 2·87 for current *vs* never, PAR 35·7% for current and former *vs* never), raised ApoB/ApoA1 ratio (3·25 for top *vs* lowest quintile, PAR 49·2% for top four quintiles *vs* lowest quintile), history of hypertension (1·91, PAR 17·9%), diabetes (2·37, PAR 9·9%), abdominal obesity (1·12 for top *vs* lowest tertile and 1·62 for middle *vs* lowest tertile, PAR 20·1% for top two tertiles *vs* lowest tertile), psychosocial factors (2·67, PAR 32·5%), daily consumption of fruits and vegetables (0·70, PAR 13·7% for lack of daily consumption), regular alcohol consumption (0·91, PAR 6·7%), and regular physical activity (0·86, PAR 12·2%), were all significantly related to acute myocardial infarction ($p<0·0001$ for all risk factors and $p=0·03$ for alcohol). These associations were noted in men and women, old and young, and in all regions of the world. Collectively, these nine risk factors accounted for 90% of the PAR in men and 94% in women.

Interpretation Abnormal lipids, smoking, hypertension, diabetes, abdominal obesity, psychosocial factors, consumption of fruits, vegetables, and alcohol, and regular physical activity account for most of the risk of myocardial infarction worldwide in both sexes and at all ages in all regions. This finding suggests that approaches to prevention can be based on similar principles worldwide and have the potential to prevent most premature cases of myocardial infarction.

Lancet 2004; 364: 937–52

Published online
September 3, 2004
http://image.thelancet.com/extras/04art8001web.pdf

See Comment page 912

*Listed at end of report.

Population Health Research Institute, Hamilton General Hospital, 237 Barton Street East, Hamilton, Ontario, Canada L8L 2X2
(Prof S Yusuf DPhil, S Ôunpuu PhD, S Hawken MSc, T Dans MD, A Avezum MD, F Lanas MD, M McQueen FRCP, A Budaj MD, P Pais MD, J Varigos BSc, L Lisheng MD)

Correspondence to:
Prof Salim Yusuf
yusufs@mcmaster.ca

Defining and Classifying Exposures

- Clear definitions for exposure are very important
- Is the exposure chronic?
  - In assessing a relationship between Type I Diabetes and stroke diabetes status could be considered a chronic exposure
- Or transient?
  - In assessing the relationship between smoking and stroke smoking status might be transient. What if a smoker quits smoking mid cohort?
- Are there different levels of exposure?
  - Sometimes smoking exposure is defined using “pack-years”
- Is there a latency period between the exposure and when risk due to the exposure may begin?
  - Are you considered at increased risk of stroke immediately after your first cigarette or after you accumulate a certain exposure level?

Defining and Classifying Outcome Events

- A clear definition of what does and does not constitute and outcome event is important
- The time at which the outcome occurs defines the person-time contributed to the study so gathering this information as precisely as possible is important
  - For some events like death, a stroke, a heart attack this may be clear
  - For other events like the development of cancer this may be ambiguous
  - Do you classify an event having occurred at time of diagnosis, time of first symptoms, something else?

Misclassification Bias

- Systematic error in measurement causing individuals exposure or outcome status to be misclassified (or mis-measured in the case of a continuous variable)
- Imperfect diagnostic tests (low sensitivity/specificity)
- Imperfect measurement instruments

QUASI-EXPERIMENTAL

Design Components

1. Intervention / Control
2. Observe outcome

QUASI-EXPERIMENTAL DESIGNS

1. Case study
2. One-group Pretest-Posttest
3. Posttest Control Group
4. Pretest-posttest Control Group
5. Interrupted time series

QUASI-EXPERIMENTAL

Design Components

1. Intervention / Control
2. Observe outcome

Diagnostic Tests

- A test with low sensitivity will misclassify some diseased individuals as healthy
  - In a study assessing disease prevalence this will result in an underestimate of the prevalence
- A test with low specificity will misclassify some healthy individuals as diseased
  - In a study assessing disease prevalence this will result in an overestimate of the prevalence

Differential vs. Non-differential Misclassification Bias

- Imprecise tools can lead to misclassification of outcomes and/or exposures leading to the prevalence of the exposure or outcome to be misestimated
- If the misclassification of the outcome does not depend on the exposure status (or vice versa) the misclassification bias is non-differential
- If the misclassification of the outcome depends on exposure status (or vice versa) the misclassification bias is differential

Differential Misclassification Bias Example

	Cases	Controls
Exposed	a	b
Not Exposed	c	d

- Case control studies often rely on recall of past exposures
- Individuals with a disease are more likely to recall a past exposure than healthy individuals: *recall bias*
- Questions about past exposures are more sensitive in cases than controls
  - Differential – depends on case/control status
- Typically leads to an overestimation of the odds ratio

Interpreting Differential Misclassification Bias

- Differential misclassification can lead to either an over or underestimate of the association
- It is up to the reader to consider carefully
  - If differential misclassification has occurred
  - If this might have over or underestimated the association
  - What magnitude of misestimation might be present

Non-Differential Misclassification Bias Example

	Cases	Controls
Exposed	a	b
Not Exposed	c	d

- To avoid using recall a case control study may use administrative health data to determine exposure status
- The classification of exposure status in healthcare data may be imperfect
- But if there is no reason to believe the magnitude of inaccuracy depends on outcome status the bias is *non-differential*
- The direction of non-differential misclassification bias is always in the direction of the null

Interpreting Non-Differential Misclassification Bias

- The direction of non-differential misclassification bias is always in the direction of the null
- It is up to the reader to consider carefully
  - If non-differential misclassification has occurred
  - What magnitude of misestimation might be present
- In a study finding no association where non-differential misclassification may have occurred this might be the reason no association was found
- In a study where an association was found, and non-differential misclassification occurred then the association may be underestimated

Some ways to avoid (or lessen) misclassification bias

- Use valid measures with as high sensitivity/specificity as possible
- Blinding: individuals assessing exposures should be blinded to outcome status (and vice versa)
  - In a case control study if an interviewer assessing exposure status knew the person had the disease they might be tempted to probe more deeply for evidence of exposure: Interviewer Bias or Diagnostic Suspicion Bias

Selection Bias

- Different from sampling error (which is random)
- Selection bias is systematic
  - It results from a flaw in the study design (flawed sampling procedures)
  - Or other factors related to study participation (like withdrawing from the study)
- Your study sample is systematically different from the population you intended to study and this is somehow related to your exposure or outcome
- There are many different sub-types of selection bias

Selection Bias Example

- The Canadian Community Health Survey (CCHS) is an annual survey facilitated by Statistics Canada to gain health information on Canadians
- Households are randomly selected to participate and face-to-face interviews were conducted with a randomly selected member of the household (prior to widespread internet use, now the CCHS is mostly performed online)

- In 2002 there was a mental health focused version of the CCHS.
- Among all households selected in 2002, 77% participated
- Using the results from the survey it was estimated that the prevalence of schizophrenia was 1.1%

Selection Bias Example

- If the target population is all adults in Canada – what are some ways that this study may have been impacted by selection bias?

Selection Bias Example

- Based on the study methodology it is not possible for individuals experiencing homelessness or individuals living in institutions (care homes, prison, etc.) to be included in the study.
  - What if there is a differing prevalence of schizophrenia in these individuals than in those studied? Sampling bias or ascertainment bias (some members of the population are less likely to be included than others – and this relates to the outcome of interest)

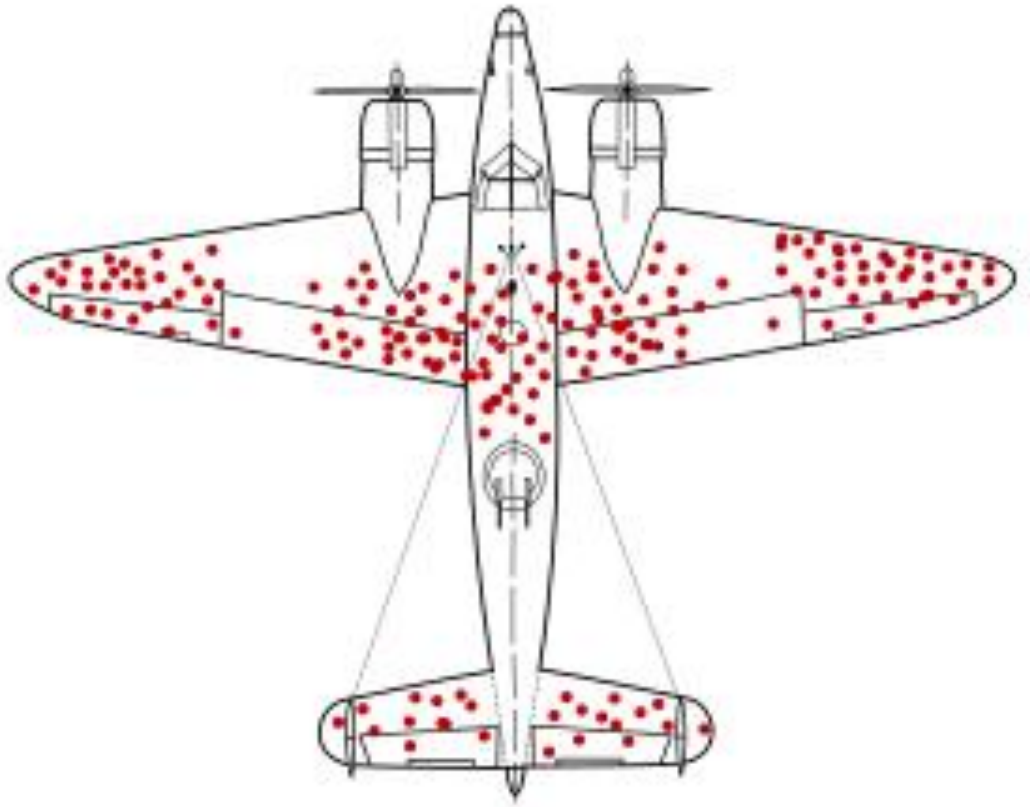
- What about the 23% of households who declined to participate in the study?
  - Could this be related to having schizophrenia and being either unable or unwilling to answer a health survey? Non-response bias (some members of the population are less likely to respond than others – and this relates to the outcome of interest)

Other forms of selection bias

- Self-selection or volunteer bias: individuals who volunteer for a study may be systematically different than the population of interest

- Attrition bias: individuals who drop out of study may be systematically different from individuals who stay in a study

- Survivorship bias: non-survivors may be systematically different from survivors



Assessing Selection Bias

- There is not a statistical procedure for determining if selection bias has occurred or not
- You need to think critically when designing or reading a study to ensure that the intended population is being appropriately represented in the study
- Often the effects of selection bias are beyond that which can be fixed or adjusted for through statistics
- You should always consider selection bias when evaluating a studies merit

Reporting Guidelines

Search for reporting guidelines

Browse for reporting guidelines by selecting one or more of these drop-downs:

Study type | Clinical area | and | Section of report

Please select...

- Please select...
- Animal pre-clinical research
- Artificial intelligence/Machine learning studies
- Clinical practice guidelines
- Diagnostic and prognostic studies
- Economic evaluations
- Experimental studies
- Mixed methods studies
- Observational studies
- Patent landscapes
- Qualitative research
- Quality improvement studies
- Reliability and agreement studies
- Study protocols
- Systematic reviews/Meta-analyses/Reviews/HTA/Overviews

Please select...

Start again | Help

Displaying 639

Most recently a

1 PRID … s in endodontics: A consensus-based deve…

2 Enhancing the quality of evidence, comparability, and reproducibility in ventriculoatrial shunt research for normal pressure hydrocephalus: A systematic review and VAS-NPH reporting guideline

Reporting guidelines for main study types

Randomised trials	CONSORT	Extensions
Observational studies	STROBE	Extensions
Systematic reviews	PRISMA	Extensions
Study protocols	SPIRIT	PRISMA-P
Diagnostic/prognostic studies	STARD	TRIPOD
Case reports	CARE	Extensions
Clinical practice guidelines	AGREE	RIGHT
Qualitative research	SRQR	COREQ
Animal pre-clinical studies	ARRIVE	
Quality improvement studies	SQUIRE	Extensions
Economic evaluations	CHEERS	Extensions

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

Continued on next page

Critically Appraising Observational Studies

Critically Appraising Research

- It is important that all research is critically evaluated such that the reader can decide if they are willing to accept the research claims to be true

- It is reasonable to view research results with a certain level of skepticism initially and carefully examine all aspects of the research study for potential flaws

- Perhaps, if no substantial flaws can be identified it is then reasonable to drop our skepticism and accept the study results

Critically appraising research

- There are many proposed frameworks for critically appraising a research study

- We will work through the framework proposed by Dr. Scott Patten from the textbook Epidemiology for Canadian Students

Step 1: Identifying the research question and hypotheses

- The author's question or hypothesis (or both) should be clearly stated
- If the purpose of critical appraisal is to determine how well an author answered their question the question must not be vague

- A research question may look like this: "Is there a difference in 90-day outcome among ischemic stroke patients treated with endovascular therapy vs. standard of care?"

- A research hypothesis may look like this: "We hypothesized that ischemic stroke patients treated with endovascular therapy would have better 90-day outcomes than those treated with standard of care."

Step 2: identify the exposure and outcome variables

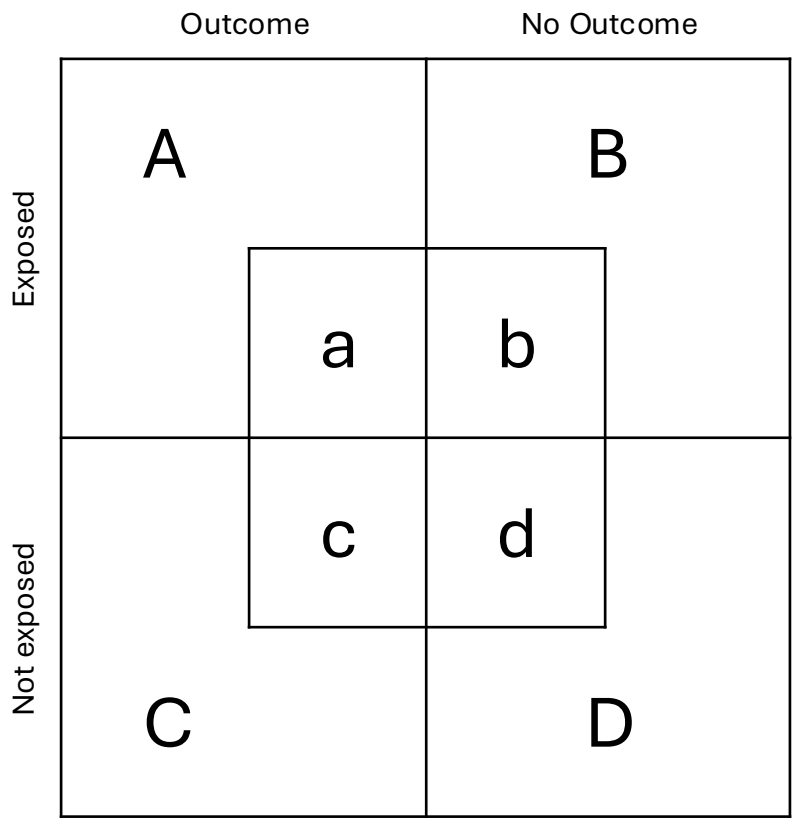
- This should be clear from the research question and methodology
- If either of these is vague it can be difficult to appraise the rigor of the study
- Sometimes there are multiple exposures and/or outcomes. The authors should clearly define one of these to be of primary interest and the others to be of secondary interest

Step 3: identify the study design

- Is the study observational or experimental?
- What is the unit of analysis (aggregate or individual)?
- If experimental, how was the intervention assigned (random or non random)?
- If observational, what type of observational study (cohort, case control, etc.)?

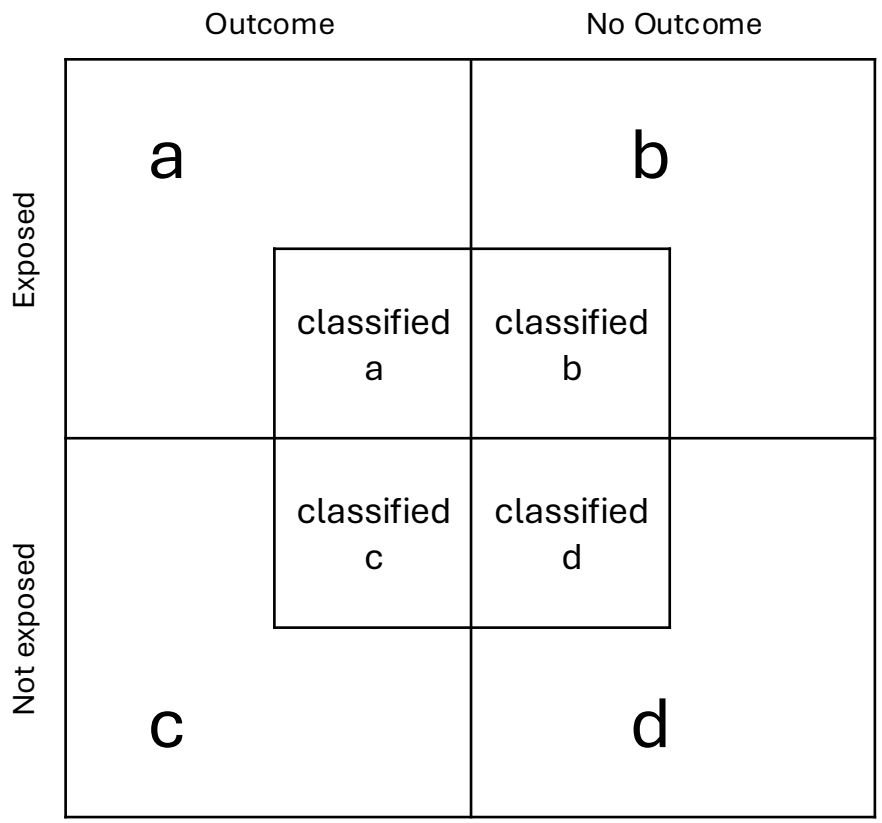
Step 4: Assessment of selection bias

- Assess how individuals were selected into the study and whether the sample is an accurate representation of the population of interest
- Common types of selection bias: ascertainment bias, non-response bias, volunteer bias, attrition bias, survivorship bias
- If you believe selection bias may have occurred, try to describe and quantify it as best you can



Step 5: assessment of misclassification bias

- Assess whether you think the exposure and/or outcome was vulnerable to misclassification bias
- If so, is the bias differential or non-differential in nature?
- If you think there is bias what is the magnitude of the bias?
- Recall examples of misclassification bias: low sensitivity/specificity of diagnostic tools, vague definitions or classification criteria, recall bias, interviewer bias, diagnostic suspicion bias



Step 6: Assessment of Confounding and Effect Modification

- Did the study consider that other variables may impact the exposure/outcome association?
  - Did the study miss any confounders?
  - Was effect modification (heterogeneity of effect) considered?

- Did the study employ any methods to control for this?

Step 7: Assessing the role of chance

- Examine the confidence intervals reported in the study
  - Are they very large and imprecise? Or are they narrow indicating good precision?

- Are you concerned the study has made a Type I Error?
  - How many comparisons were made? Were any methods used to conserve the Type I Error rate?

- Are you concerned the study has made a Type II Error?
  - Did the study justify the chosen sample size with a power calculation?

Step 8: Assessing causality

- If the study is asserting a causal relationship, what type of evidence was presented to support this?

- Recall the Bradford-Hill causal criteria: consistency, biologic plausibility, dose-response, temporality, strength, reversibility

Step 9: Assessing generalizability

- Do you think the study findings might apply beyond the intended target population?
  - Ex. Do you think a study performed in the US could be generalized to Canada?
  - Do you think a study performed in younger people could be generalized to older people?

- Sometimes this is reasonable to consider as there may be limited resources to perform a follow up study in a new population

- This is more a manner of subjective opinion rather than fact
- Just because a study cannot be generalized beyond its target population does not make it poor quality

Approaching critical appraisal

- Critical appraisal should be approached systematically, following the steps laid out

- If at any time a fatal flaw in the study is found, you may decide to halt the appraisal process and decide there is no value in further assessment
  - Ex. if a study is found to be subject to large misclassification bias, assessing things like precision is unnecessary. It does not matter the width of a confidence interval if it surrounds an invalid point estimate.

Critical Appraisal Example

BMC Public Health

RESEARCH ARTICLE Open Access

Alcohol, psychoactive substances and non-fatal road traffic accidents - a case-control study

Stig Tore Bogstrand[1,2*], Hallvard Gjerde[3], Per Trygve Normann[3], Ingeborg Rossow[4,5] and Øivind Ekeberg[6]

Abstract

Background: The prevalence of alcohol and other psychoactive substances is high in biological specimens from injured drivers, while the prevalence of these psychoactive substances in samples from drivers in normal traffic is low. The aim of this study was to compare the prevalence of alcohol and psychoactive substances in drivers admitted to hospital for treatment of injuries after road traffic accidents with that in drivers in normal traffic, and calculate risk estimates for the substances, and combinations of substances found in both groups.

Methods: Injured drivers were recruited in the hospital emergency department and drivers in normal conditions were taken from the hospital catchment area in roadside tests of moving traffic. Substances found in blood samples from injured drivers and oral fluid samples from drivers in moving traffic were compared using equivalent cut off concentrations, and risk estimates were calculated using logistic regression analyses.

Results: In 21.9% of the injured drivers, substances were found: most commonly alcohol (11.5%) and stimulants eg. cocaine or amphetamines (9.4%). This compares to 3.2% of drivers in normal traffic where the most commonly found substances were z-hypnotics (0.9%) and benzodiazepines (0.8%). The greatest increase in risk of being injured was for alcohol combined with any other substance (OR: 231.9, 95% CI: 33.3- 1615.4, $p < 0.001$), for more than three psychoactive substances (OR: 38.9, 95% CI: 8.2- 185.0, $p < 0.001$) and for alcohol alone (OR: 36.1, 95% CI: 13.2- 98.6, $p < 0.001$). Single use of non-alcohol substances was not associated with increased accident risk.

Conclusion: The prevalence of psychoactive substances was higher among injured drivers than drivers in normal moving traffic. The risk of accident is greatly increased among drivers who tested positive for alcohol, in particular, those who had also ingested one or more psychoactive substances. Various preventive measures should be considered to curb the prevalence of driving under the influence of psychoactive substances as these drivers constitute a significant risk for other road users as well as themselves.

Keywords: Alcohol, Case–control, Emergency treatment, Injury, Psychoactive substances, Road traffic accident

Step 1: Identifying the research question and hypotheses

Step 1: Identifying the research question and hypotheses

“The aim of our study was to compare the prevalence of alcohol and psychoactive drugs in drivers admitted to hospital for treatment of non-fatal injuries after road traffic accidents with prevalence in drivers in normal traffic, and to calculate odds ratios for single and multiple substance use.”

Step 2: identify the exposure and outcome variables

Step 3: identify the study design

Step 4: Assessment of selection bias

Step 4: Assessment of selection bias

Description of cases

Inclusion of injured drivers

Drivers older than 18 year of age admitted to the emergency department at Oslo University Hospital, Ullevål, were included. Criteria for inclusion was informed consent from the patient or his or hers next of kin if the patient was not able to give an informed consent. If the patient was unavailable to give consent in the emergency department because of acute medical treatment or for other reasons, blood samples were obtained and the patients were asked to give informed consent later during the stay. If the patient refused to participate, was discharged before giving consent or for other reasons could not give an informed consent, the blood sample was destroyed. No data was registered on non participants. The data was collected over a 12 month period from December 2007 to December 2008. Further details on study inclusion have been published in an earlier paper [1].

Description of controls

Selection of drivers in roadside survey

The selection of drivers of cars and vans (reference group) was carried out in collaboration with the Mobile Police Service (MPS) in south-eastern Norway as a part of the DRUID project. Drivers were selected from April 2008 to March 2009 using a stratified multi-stage cluster sampling procedure. The first stage consisted of selecting regions in south-eastern Norway. The second stage covered selection of road sites, and dates and times for the sampling, while the third stage consisted of randomly stopping of cars and vans for routine control of breath alcohol or driving licence by the MPS. Participation was voluntary and anonymous. Our study team asked as many drivers as possible for participation in the project during two two-hour periods for each study day. Oral and written information was given to each driver; information leaflets were available in 12 languages. If informed consent was given, a sample of oral fluid was taken and a short questionnaire was filled in. The following data were recorded: gender, age, nationality of the driver, and type of vehicle. Drivers did not receive any reward for taking part in the survey.

Step 5: assessment of misclassification bias

Step 5: assessment of misclassification bias

Blood sample analysis

The blood samples were analyzed according to the forensic toxicology analytical programme using two different methods for screening and confirmation analyses of the drugs. The samples were first screened for amphetamines, cannabis, cocaine metabolites and opiates by an immunological method [16]. Screening for other drugs was carried out using high-performance liquid chromatography with mass spectroscopy detection (LC-MS) [17]. Then, all drug findings were confirmed and quantified using gas or liquid chromatography with mass spectroscopy detection (GC-MS or LC-MS) [17–20]. An enzymatic dehydrogenase method was used for ethanol analysis [21]. Analysis was carried out for a total of 20 different substances, constituting all impairing drugs on the Norwegian market which have been shown to be linked to increased accident risk [22]. Analytical cut-off values corresponding to the cut-off used in the DRUID project (http://www.druid-project.eu) were used, as described in Table 1[23].

Analysis of oral fluid

The amount of collected oral fluid was determined by weighing, and the dilution with buffer was taken into account when calculating analytical results expressed in g per kg or ng per ml of undiluted (native) oral fluid. Samples of less than 0.2 ml oral fluid were regarded as failed samples and therefore excluded.

Alcohol was analysed by an automated enzymatic method using alcohol dehydrogenase [21]. Drug concentrations in oral fluid-buffer mixtures were determined by liquid chromatography – tandem mass spectrometry [24]. Cut-off concentrations are presented in Table 1. Samples with drug concentrations above the linearity limits were diluted and re-analysed.

Step 6: Assessment of Confounding

Table 3 Crude and adjusted odds ratio of accident risk (n = 5401)

From: Alcohol, psychoactive substances and non-fatal road traffic accidents - a case-control study

Alcohol	***Crude OR (95% CI)***	***Adjusted OR (95% CI)*** [1]
No alcohol (referent)		
Alcohol alone	29.0 (11.5- 73.0)**	36.1 (13.2- 98.6)**
Alcohol combined	124.4 (22.5- 688.5)**	231.9 (33.3- 1615.4)**
Positive samples with no alcohol	***Crude OR (95% CI)***	***Adjusted OR (95% CI)*** [2]
No non-alcohol psychoactive substances (referent)		
One psychoactive substance	1.6 (0.5- 5.0)(ns)	1.4 (0.4- 4.4)(ns)
Two psychoactive substances	17.1 (5.6- 52.4)**	13.3 (4.2- 41.3)**
Three or more psychoactive substances	51.4 (11.3- 233.5)**	38.9 (8.2- 185.0)**

Chi-square test: (ns) = Not statistically significant, , ** = P < 0.001.

[1]Adjusted for age group and day and time.

[2]Adjusted for age group.

Step 7: assessing the role of chance

Table 3 Crude and adjusted odds ratio of accident risk (n = 5401)

From: Alcohol, psychoactive substances and non-fatal road traffic accidents - a case-control study

Alcohol	***Crude OR (95% CI)***	***Adjusted OR (95% CI)*** [1]
No alcohol (referent)		
Alcohol alone	29.0 (11.5- 73.0)**	36.1 (13.2- 98.6)**
Alcohol combined	124.4 (22.5- 688.5)**	231.9 (33.3- 1615.4)**
Positive samples with no alcohol	***Crude OR (95% CI)***	***Adjusted OR (95% CI)*** [2]
No non-alcohol psychoactive substances (referent)		
One psychoactive substance	1.6 (0.5- 5.0)(ns)	1.4 (0.4- 4.4)(ns)
Two psychoactive substances	17.1 (5.6- 52.4)**	13.3 (4.2- 41.3)**
Three or more psychoactive substances	51.4 (11.3- 233.5)**	38.9 (8.2- 185.0)**

Chi-square test: (ns) = Not statistically significant, , ** = $P < 0.001$.

[1]Adjusted for age group and day and time.

[2]Adjusted for age group.

Step 8: assessing causality

Conclusion

Prevalence of psychoactive substances was higher among injured drivers than drivers in normal moving traffic. Alcohol and stimulant drugs were particularly prevalent among drug positive injured drivers. The adjusted OR was high for alcohol combined with drugs; for alcohol alone, and for combinations of two or several non-alcohol substances. Various preventive measures should be considered to curb the prevalence of driving under the influence of psychoactive substances as these drivers constitute a significant risk for other road users as well as themselves.

Step 9: assessing generalizability