Accelerating Impact: Immersive Summer Bootcamp in Implementation Science and Biostatistics

Georgian Implementation Science Fogarty Training (GIFT) Program

Ilia State University & Yale University

Yale Center for Analytical Sciences

Study Design

EXPOSURE

A characteristic (variable) potentially associated with a disease or health state – the independent/explanatory variable

- Drug/Intervention
- Demographic characteristic
- Behavioral factor
- Psychosocial
- Genetic
- etc

OUTCOME

A characteristic (variable) that is a marker of disease or a health state – the dependent/response variable

- Death
- Disease
- Morbidity
- Biological
- Behavioral
- Psychosocial
- Genetic
- etc

Study Design

- Observational – collect information about 1 or more groups but do nothing to affect them
- Experimental – researcher affects or manipulates what happens to all or some individuals


Systematic reviews
Strength of evidence pyramid
RCTs
Cohorts
Case controls
Surveys
Animal research
Expert opinion

Study Design Features:
Timing of Data Collection

- Prospective – data collected forwards in time after the onset of the study and prior to occurrence of outcomes

- Retrospective – data collected after outcomes have occurred and usually acquired from existing sources (eg. medical records) or recall

Study Design Features:
Period of Time

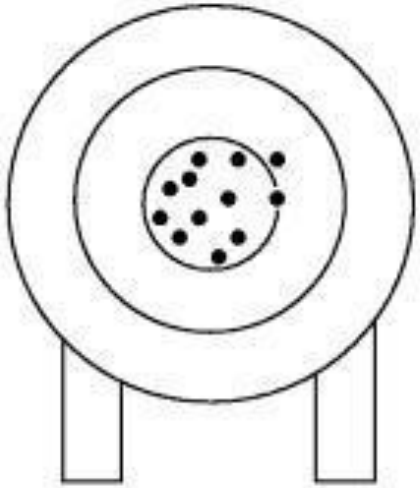
- Cross-sectional – individuals observed only once – snapshot in time

- Longitudinal – study investigates changes over time

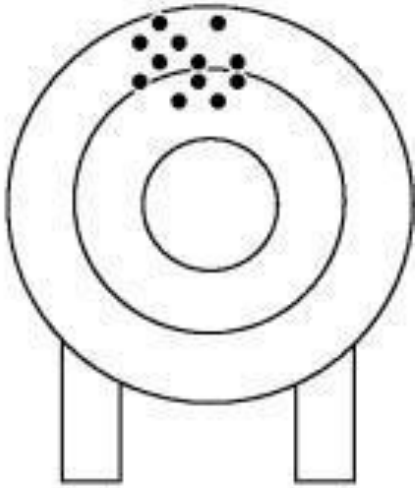
Bias

- Non-random, systematic error in the design or conduct of the study – leads to deviation from the truth
  - Selection bias – distortions that result from procedures used to select subjects and from factors that influence study participation (Rothman, 1998)
  - Information bias – a flaw in measuring exposure or outcome data that results in different quality (accuracy) of information between comparison groups (Last, 1995)

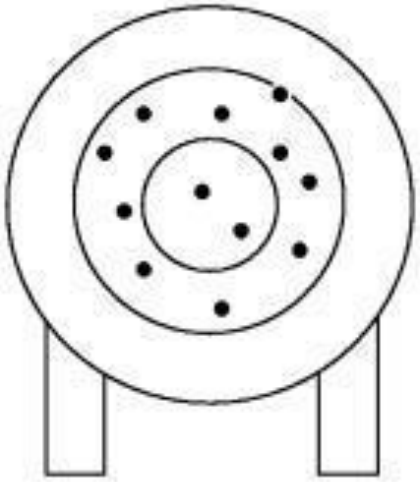
Low Bias, Low Variance



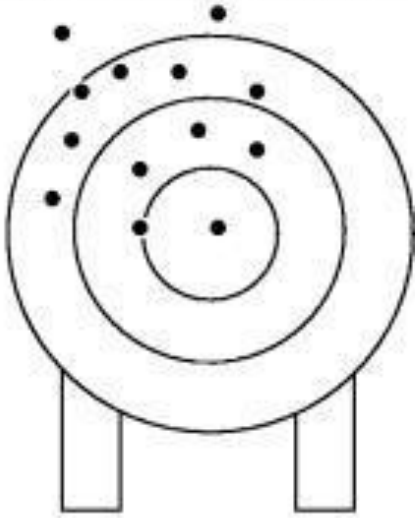
High Bias, Low Variance



Low Bias, High Variance



High Bias, High Variance



Selection Bias

- Example 1– if in a clinical trial subjects were allowed to choose whether they were in the dietary fiber or control group, subjects that were more health conscious may choose the fiber group and subsequent differences may be due partly or entirely to differences between the subjects rather than the intervention

- Example 2 – if hospital-based controls are selected in a case control study examining the association between alcohol intake and colon cancer, the effect may be underestimated as hospital patients are known to drink more alcohol than the general public

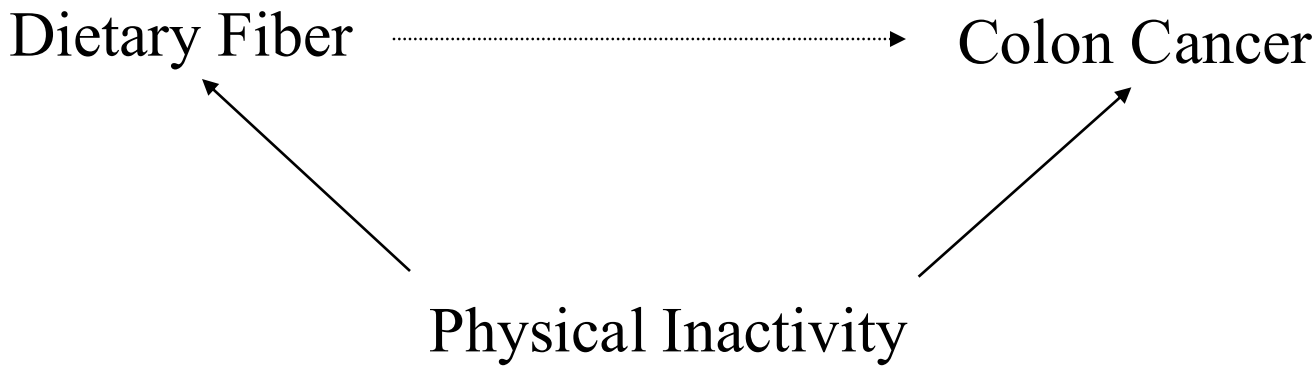
Information Bias

- Recall Bias – In a case-control study of colon cancer and dietary fiber, the cases may be more likely to accurately recall their dietary fiber consumption

- Interviewer Bias – In a case-control study using a dietary interview, if the interviewer knows a subject is a control they may search more for sources of dietary fiber consumption

Confounding

- The confusion of effects whereby part or all of the purported association between two variables (i.e. exposure and disease) is the result of a third variable.



- “Bias and confounding are not synonymous. Bias arises from flawed information or subject selection so that a wrong association is found. Confounding produces relations that are factually right, but that cannot be interpreted causally because some underlying, unaccounted-for factor is associated with both exposure and outcome.”

Vandenbrouchke et al. 2007

Observational Studies

- Case Series
- Case-control
- Cross-sectional
- Cohort

Case Series

8. CDC. State Medicaid coverage for tobacco-dependence treatments—United States, 1998 and 2000. MMWR 2002;50:979–82.
9. Solanki G, Halpin Schauffler H. Cost-sharing and the utilization of clinical preventive services. Am J Prev Med 1999;17:127–33.
10. George Washington University Medical Center School of Public Health and Health Services. Sample purchasing specifications related to tobacco-use prevention and cessation: a technical assistance document. Available at http://www.gwhealthpolicy.org/newsps/tobacco.

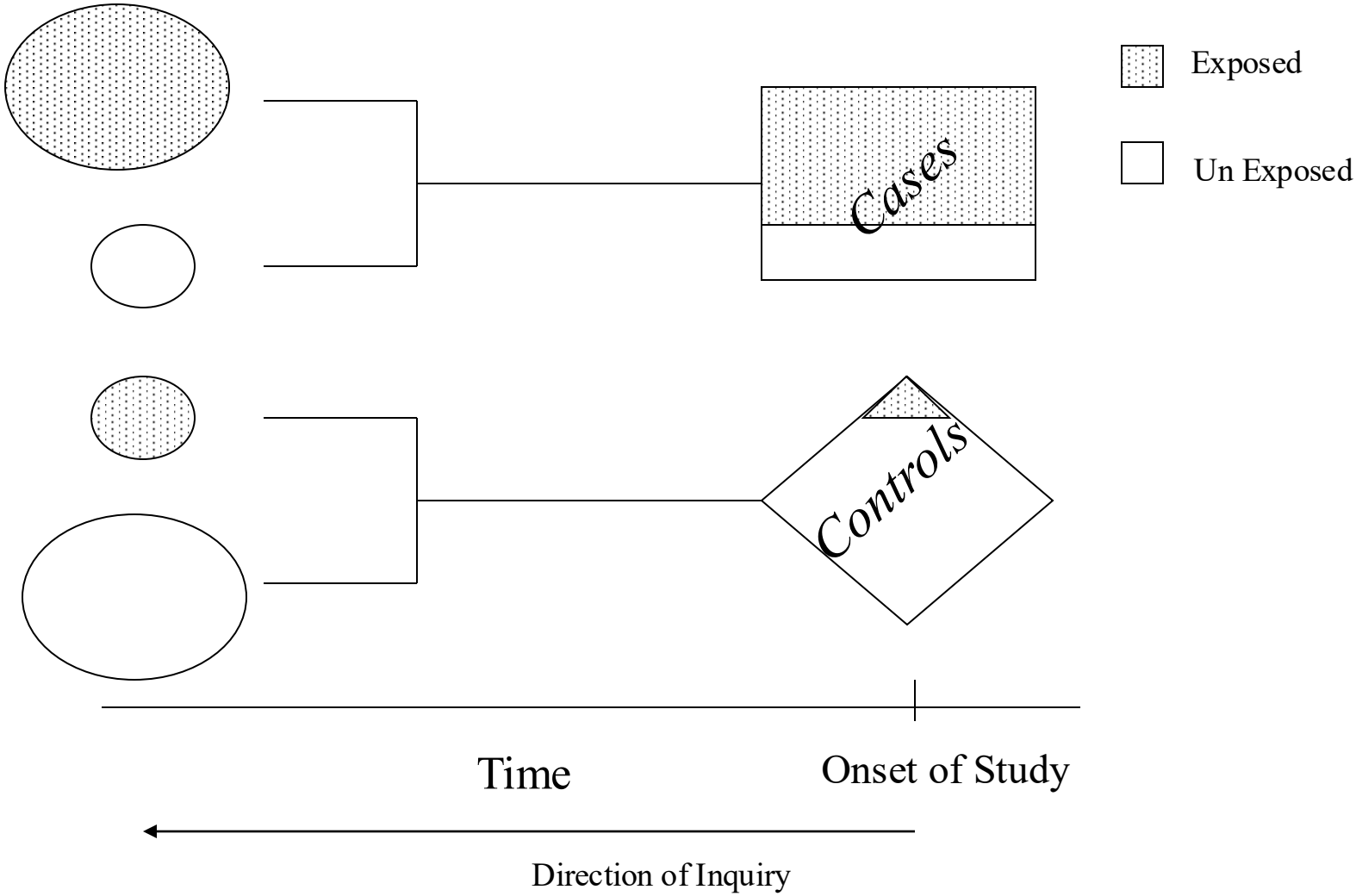
Update: Severe Acute Respiratory Syndrome — United States, May 28, 2003

CDC continues to work with state and local health departments, the World Health Organization (WHO), and other partners to investigate cases of severe acute respiratory syndrome (SARS). This report updates SARS cases reported worldwide and in the United States and reports a seventh laboratory-confirmed U.S. case.

During November 1, 2002–May 28, 2003, a total of 8,240 SARS cases were reported to WHO from 28 countries, including the United States; 745 deaths (case-fatality proportion: 9.0%) have been reported (*1*). The 363 SARS cases identified in the United States have been reported from 41 states and Puerto Rico, with 297 (82%) cases classified as suspect SARS and 66 (18%) classified as probable SARS (more severe illnesses characterized by the presence of pneumonia or acute respiratory distress syndrome) (Figure, Table) (*2*). Of the 66 probable SARS patients, 43 (65%) were hospitalized, and two (3%) required mechanical ventilation. No SARS-related deaths have been reported in the United States. Of 66 probable cases, 64 (97%) were attributed to international travel to areas with documented or suspected community transmission of SARS within the 10 days before illness onset; the remaining two (3%) probable cases occurred in a health-care worker who provided care to a SARS patient and a household contact of a SARS patient. Since the last update, new cases of SARS have been reported in Toronto, Canada, and CDC has

Case Control Studies:

What happened?



Case Control Study

American Journal of Epidemiology

Vol. 157, No. 2
Printed in U.S.A.
DOI: 10.1093/aje/kwf181

Case-Control Study of Physical Activity and Breast Cancer Risk among Premenopausal Women in Germany

Karen Steindorf[1], Martina Schmidt[1], Silke Kropp[2], and Jenny Chang-Claude[2]

[1] Unit of Environmental Epidemiology, German Cancer Research Center, Heidelberg, Germany.
[2] Division of Clinical Epidemiology, German Cancer Research Center, Heidelberg, Germany.

Received for publication June 3, 2002; accepted for publication August 5, 2002.

Important aspects of the recognized inverse relation between physical activity and breast cancer risk are still under discussion. Data on physical activity from sports, occupational activity, household tasks, walking, and cycling by reported frequency, duration, and intensity during adolescence and young adulthood were collected in 1999–2000 from 360 premenopausal breast cancer cases and 886 controls who had previously participated in a German population-based case-control study. In multivariate conditional logistic regression, no association between total physical activity and premenopausal breast cancer was found in two age periods. For women who were active during both periods, the adjusted odds ratio was 0.83 (95% confidence interval: 0.60, 1.14). When both age periods were combined, higher quartiles of total physical activity compared with the lowest quartile showed adjusted odds ratios of 0.97, 0.68, and 0.94. Only the effect of moderately high physical activity was statistically significant. Analyses by type of activity revealed significant protective effects for women who reported the highest levels of cycling activities (adjusted odds ratio = 0.66, 95% confidence interval: 0.45, 0.97). These data do not suggest an inverse monotonic association between total physical activity and breast cancer risk in premenopausal women. The study prevalence of cycling and walking for transportation demonstrated that national habits need consideration in the exposure assessment.

breast neoplasms; case-control studies; exercise; premenopause; questionnaires

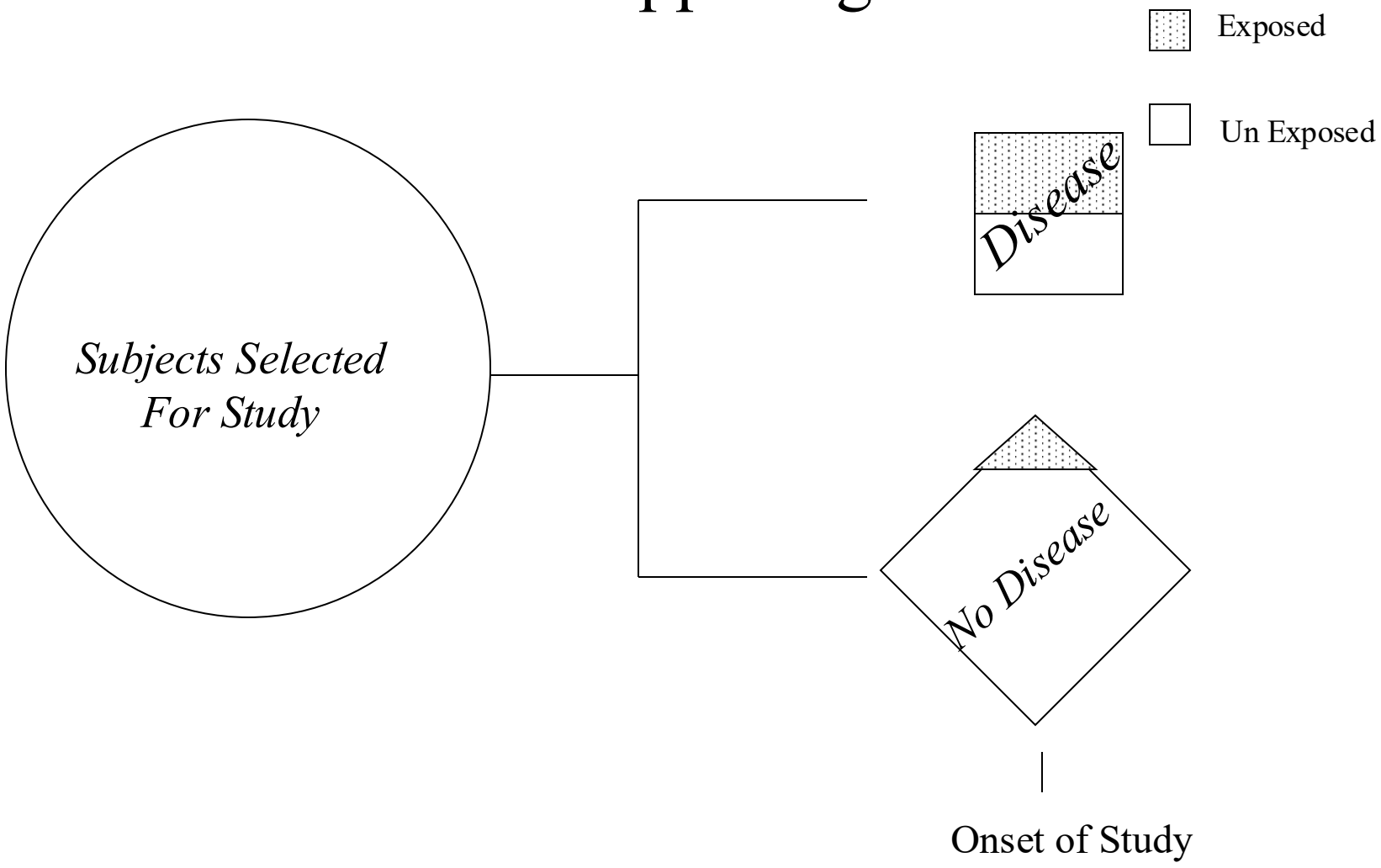
Case Control Studies

- Advantages
  - Practical – quick and cheap
  - Good for rare diseases
  - Can simultaneously study multiple exposures

- Disadvantages
  - Susceptible to many biases as collect data retrospectively
  - Outcome may change exposure
  - Can’t calculate incidence
  - Difficulty in selecting appropriate control group

Cross-sectional Studies:

What's happening?



Cross Sectional Study

Dietary Magnesium, Potassium, Sodium, and Children's Lung Function

Frank D. Gilliland,[1] Kiros T. Berhane,[1] Yu-Fen Li,[1] Deborah H. Kim,[1] and Helene G. Margolis[2]

To investigate the effects of dietary magnesium, potassium, and sodium on children's lung function, the authors examined cross-sectional dietary data and pulmonary function tests from 2,566 children aged 11–19 years who attended schools in 12 southern California communities during 1998–1999. During school visits, each child completed a health update questionnaire, a validated food frequency questionnaire, and spirometric lung function testing. Low magnesium and potassium intakes were associated with lower lung function. Girls with low magnesium intake had lower forced expiratory flow at 75% of the forced vital capacity (FEF_{75}) (–8.3%, 95% confidence interval: –14.8, –1.4) than did girls with higher intake; reductions were larger in girls with asthma (forced expiratory flow between 25% and 75% of the forced vital capacity ($FEF_{25–75}$) (–16.2%, 95% confidence interval: –22.7, –9.1) and FEF_{75} (–24.9%, 95% confidence interval: –32.8, –16.1)) than in girls without asthma ($FEF_{25–75}$ (–2.0%, 95% confidence interval: –7.4, 3.8) and FEF_{75} (–4.1%, 95% confidence interval: –11.3, 3.7)). Boys with low magnesium intake showed deficits in forced vital capacity (–2.8%, 95% confidence interval: –5.4, –0.2) compared with boys with higher intake. The effects of low magnesium intake did not vary substantially in boys with and without asthma. Among girls, low potassium intake was also associated with deficits in forced expiratory volume in 1 second (–2.7%, 95% confidence interval: –5.2, –0.1) and forced vital capacity (–2.4%, 95% confidence interval: –4.7, –0.1). In summary, low magnesium and potassium intakes were associated with lower lung volumes and flows. *Am J Epidemiol* 2002;155:125–31.

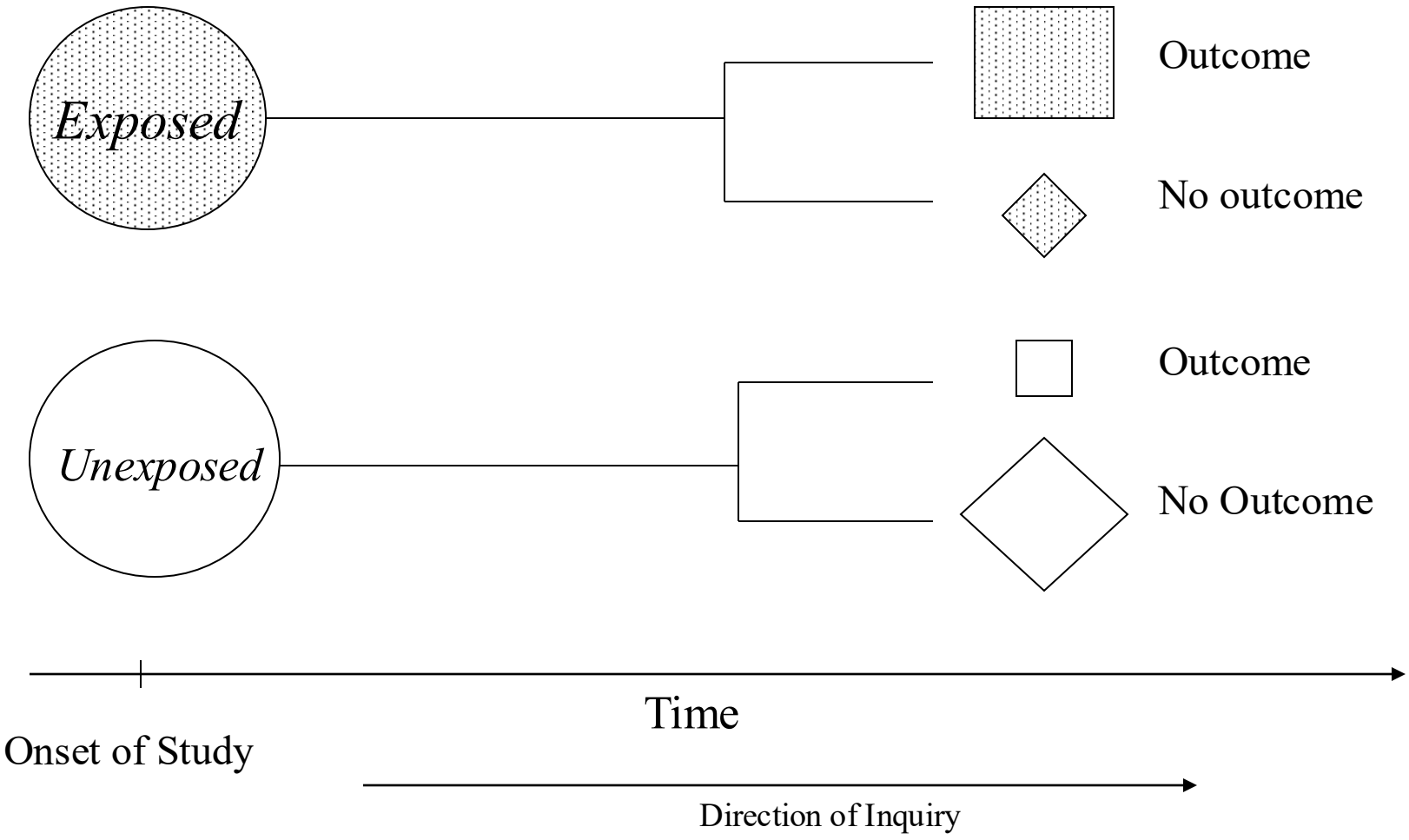
Cross Sectional Studies

- Advantages
  - Practical – quick and cheap
  - Useful for generating hypotheses
  - Can estimate exposure and outcome prevalences
  - Can study multiple exposures and outcomes

- Disadvantages
  - Temporal sequence is not established (not very good for establishing causal relations)
  - Selective survival
  - Outcomes may have changed exposures
  - Not incident outcomes

Cohort Studies:

What will happen?



Cohort Study

Television Watching and Other Sedentary Behaviors in Relation to Risk of Obesity and Type 2 Diabetes Mellitus in Women

Frank B. Hu, MD, PhD; Tricia Y. Li, MD; Graham A. Colditz, MD, DrPH; Walter C. Willett, MD, DrPH; JoAnn E. Manson, MD, DrPH

JAMA. 2003;289:1785-1791.

Context Current public health campaigns to reduce obesity and type 2 diabetes have largely focused on increasing exercise, but have paid little attention to the reduction of sedentary behaviors.

Objective To examine the relationship between various sedentary behaviors, especially prolonged television (TV) watching, and risk of obesity and type 2 diabetes in women.

Design, Setting, and Participants Prospective cohort study conducted from 1992 to 1998 among women from 11 states in the Nurses' Health Study. The obesity analysis included 50 277 women who had a body mass index (BMI) of less than 30 and were free from diagnosed cardiovascular disease, diabetes, or cancer and completed questions on physical activity and sedentary behaviors at baseline. The diabetes analysis included 68 497 women who at baseline were free from diagnosed diabetes mellitus, cardiovascular disease, or cancer.

Main Outcome Measures Onset of obesity and type 2 diabetes mellitus.

Results During 6 years of follow-up, 3757 (7.5%) of 50 277 women who had a BMI of less than 30 in 1992 became obese (BMI 30). Overall, we documented 1515 new cases of type 2 diabetes. Time spent watching TV was positively associated with risk of obesity and type 2 diabetes. In the multivariate analyses adjusting for age, smoking, exercise levels, dietary factors, and other covariates, each 2-h/d increment in TV watching was associated with a 23% (95% confidence interval [CI], 17%-30%) increase in obesity and a 14% (95% CI, 5%-23%) increase in risk of diabetes; each 2-h/d increment in sitting at work was associated with a 5% (95% CI, 0%-10%) increase in obesity and a 7% (95% CI, 0%-16%) increase in diabetes. In contrast, standing or walking around at home (2 h/d) was associated with a 9% (95% CI, 6%-12%) reduction in obesity and a 12% (95% CI, 7%-16%) reduction in diabetes. Each 1 hour per day of brisk walking was associated with a 24% (95% CI, 19%-29%) reduction in obesity and a 34% (95% CI, 27%-41%) reduction in diabetes. We estimated that in our cohort, 30% (95% CI, 24%-36%) of new cases of obesity and 43% (95% CI, 32%-52%) of new cases of diabetes could be prevented by adopting a relatively active lifestyle (<10 h/wk of TV watching and 30 min/d of brisk walking).

Conclusions Independent of exercise levels, sedentary behaviors, especially TV watching, were associated with significantly elevated risk of obesity and type 2 diabetes, whereas even light to moderate activity was associated with substantially lower risk. This study emphasizes the importance of reducing prolonged TV watching and other sedentary behaviors for preventing obesity and diabetes.

Cohort Studies

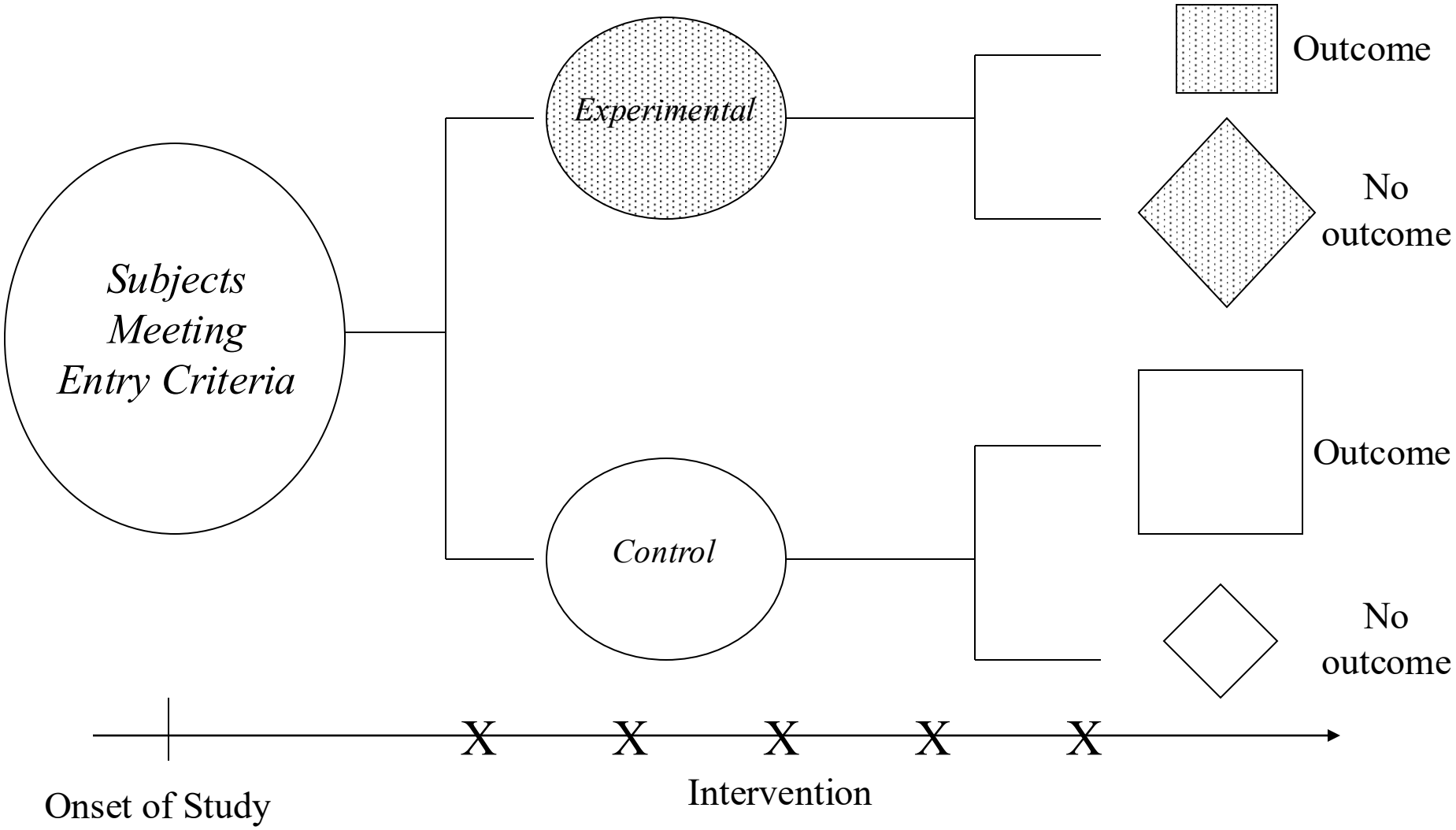
- Advantages
  - Proper temporal sequence
  - Can eliminate a lot of biases
  - Good for a rare exposure
  - Can examine multiple outcomes
  - Can estimate incidence

- Disadvantages
  - Not randomized
  - Expensive and long
  - Usually larger sample size than case-control
  - Bias and decreased power from loss to follow-up

Experimental Studies

- Parallel-group (concurrent controls)
- Self-controlled
- Cluster-randomized

Parallel Groups



Parallel Groups

A Randomized Trial of a Low-Carbohydrate Diet for Obesity

Gary D. Foster, Ph.D., Holly R. Wyatt, M.D., James O. Hill, Ph.D., Brian G. McGuckin, Ed.M., Carrie Brill, B.S., B. Selma Mohammed, M.D., Ph.D., Philippe O. Szapary, M.D., Daniel J. Rader, M.D., Joel S. Edman, D.Sc., and Samuel Klein, M.D.

ABSTRACT

BACKGROUND

Despite the popularity of the low-carbohydrate, high-protein, high-fat (Atkins) diet, no randomized, controlled trials have evaluated its efficacy.

METHODS

We conducted a one-year, multicenter, controlled trial involving 63 obese men and women who were randomly assigned to either a low-carbohydrate, high-protein, high-fat diet or a low-calorie, high-carbohydrate, low-fat (conventional) diet. Professional contact was minimal to replicate the approach used by most dieters.

RESULTS

Subjects on the low-carbohydrate diet had lost more weight than subjects on the conventional diet at 3 months (mean [±SD], −6.8±5.0 vs. −2.7±3.7 percent of body weight; P=0.001) and 6 months (−7.0±6.5 vs. −3.2±5.6 percent of body weight, P=0.02), but the difference at 12 months was not significant (−4.4±6.7 vs. −2.5±6.3 percent of body weight, P=0.26). After three months, no significant differences were found between the groups in total or low-density lipoprotein cholesterol concentrations. The increase in high-density lipoprotein cholesterol concentrations and the decrease in triglyceride concentrations were greater among subjects on the low-carbohydrate diet than among those on the conventional diet throughout most of the study. Both diets significantly decreased diastolic blood pressure and the insulin response to an oral glucose load.

CONCLUSIONS

The low-carbohydrate diet produced a greater weight loss (absolute difference, approximately 4 percent) than did the conventional diet for the first six months, but the differences were not significant at one year. The low-carbohydrate diet was associated with a greater improvement in some risk factors for coronary heart disease. Adherence was poor and attrition was high in both groups. Longer and larger studies are required to determine the long-term safety and efficacy of low-carbohydrate, high-protein, high-fat diets.

Design Techniques to Avoid Bias

- Blinding- knowledge of treatment may influence recruitment, treatment allocation, assessment of endpoints, handling of withdrawals, exclusion of data from analysis
- Randomization – tends to produce treatment groups in which distribution of prognostic factors, known and unknown, are similar
- Minimizing anticipated irregularities – protocol violations, withdrawals, missing values

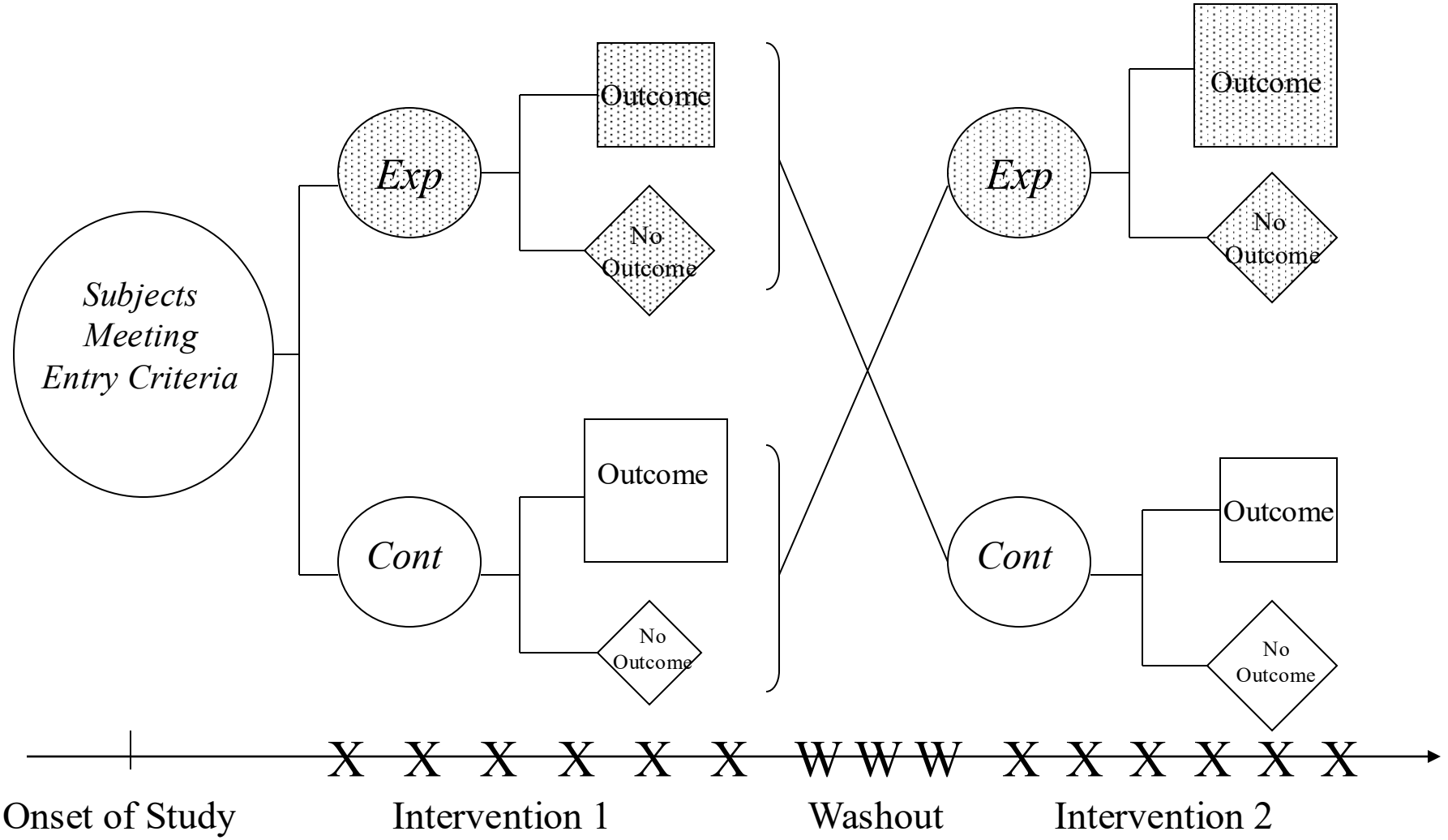
http://www.fda.gov/cder/guidance/ICH_E9-fnl.PDF

Treatment Assignment

- Randomized – best way
- Systematic – easy to manipulate
- Non randomized - open to bias in patient assignment – are groups comparable?
  - Maybe stronger patients get more aggressive treatment and higher risk patients treated more conservatively
  - Historical controls – biases from factors that change over time – Controls need:
    - Precisely defined treatment
    - Same eligibility criteria, procedures and evaluations
    - Important prognostic variables must be known and similar
    - No known external factor that could lead to different results

Cross-over Studies:

Self-controls



Cross-over Study

Morphine, Gabapentin, or Their Combination for Neuropathic Pain

Ian Gilron, M.D., Joan M. Bailey, R.N., M.Ed., Dongsheng Tu, Ph.D., Ronald R. Holden, Ph.D., Donald F. Weaver, M.D., Ph.D., and Robyn L. Houlden, M.D.

ABSTRACT

BACKGROUND

The available drugs to treat neuropathic pain have incomplete efficacy and dose-limiting adverse effects. We compared the efficacy of a combination of gabapentin and morphine with that of each as a single agent in patients with painful diabetic neuropathy or postherpetic neuralgia.

METHODS

In this randomized, double-blind, active placebo–controlled, four-period crossover trial, patients received daily active placebo (lorazepam), sustained-release morphine, gabapentin, and a combination of gabapentin and morphine — each given orally for five weeks. The primary outcome measure was mean daily pain intensity in patients receiving a maximal tolerated dose; secondary outcomes included pain (rated according to the Short-Form McGill Pain Questionnaire), adverse effects, maximal tolerated doses, mood, and quality of life.

RESULTS

Of 57 patients who underwent randomization (35 with diabetic neuropathy and 22 with postherpetic neuralgia), 41 completed the trial. Mean daily pain (on a scale from 0 to 10, with higher numbers indicating more severe pain) at a maximal tolerated dose of the study drug was as follows: 5.72 at baseline, 4.49 with placebo, 4.15 with gabapentin, 3.70 with morphine, and 3.06 with the gabapentin–morphine combination ($P<0.05$ for the combination vs. placebo, gabapentin, and morphine). Total scores on the Short-Form McGill Pain Questionnaire (on a scale from 0 to 45, with higher numbers indicating more severe pain) at a maximal tolerated dose were 14.4 with placebo, 10.7 with gabapentin, 10.7 with morphine, and 7.5 with the gabapentin–morphine combination ($P<0.05$ for the combination vs. placebo, gabapentin, and morphine). The maximal tolerated doses of morphine and gabapentin were lower ($P<0.05$) with the combination than for each drug as single agent. At the maximal tolerated dose, the gabapentin–morphine combination resulted in a higher frequency of constipation than gabapentin alone ($P<0.05$) and a higher frequency of dry mouth than morphine alone ($P<0.05$).

CONCLUSIONS

Gabapentin and morphine combined achieved better analgesia at lower doses of each drug than either as a single agent, with constipation, sedation, and dry mouth as the most frequent adverse effects.

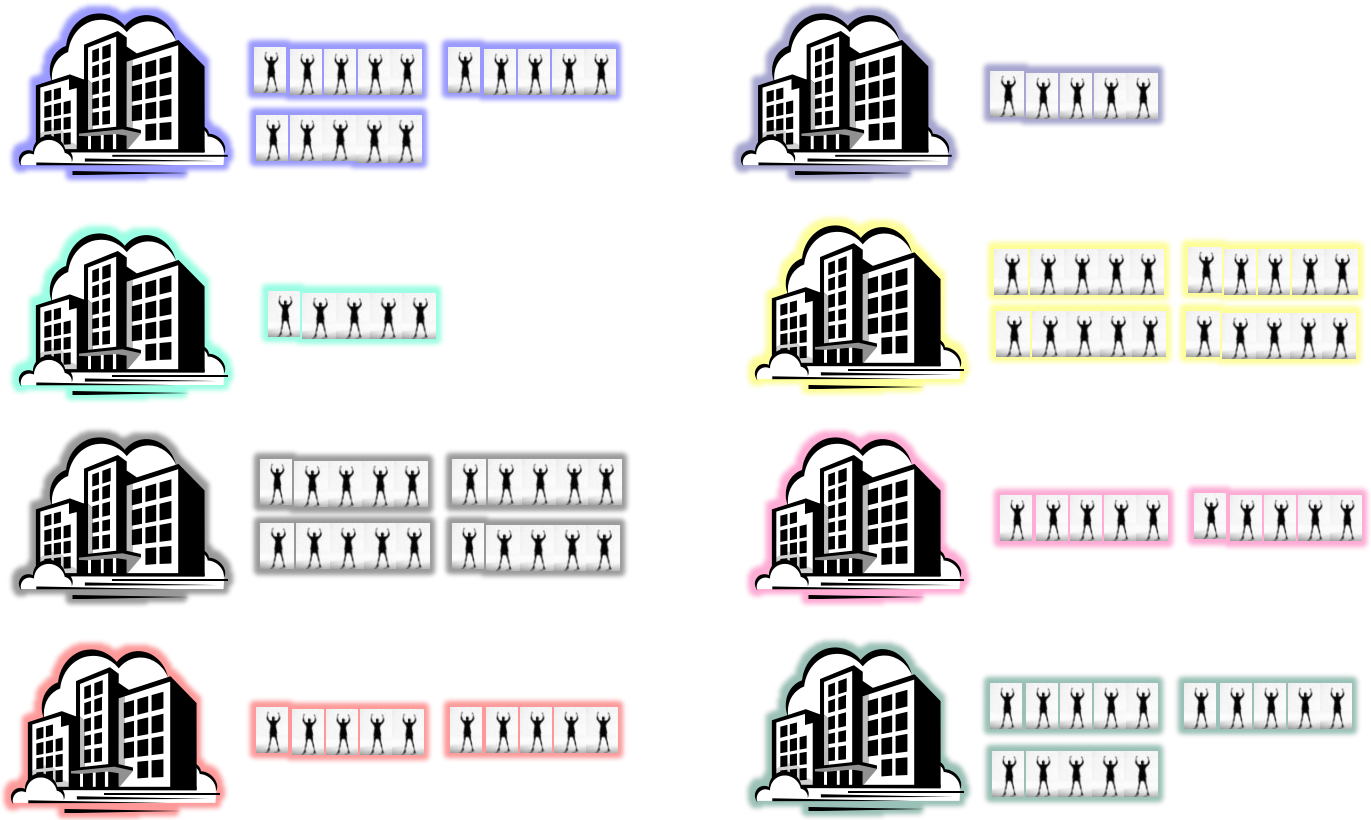
Avoiding Carryover

- Disease should be chronic and stable
- Effects of the medication should develop fully within the treatment period
- Washout periods should be sufficiently long for complete reversibility of effect

http://www.fda.gov/cder/guidance/ICH_E9-fnl.PDF

Cluster Randomized Trials

- Experimental units are clusters of individuals – randomize the clusters



Cluster Randomized Trials

Insecticide impregnated curtains to control domestic transmission of cutaneous leishmaniasis in Venezuela: cluster randomised trial

Axel Kroeger, Elci Villegas Avila, Linda Morison

School of Tropical Medicine, Liverpool L3 5QA
Axel Kroeger
professor of international community health

Research Centre "José W Torrealba," Núcleo Universitario "Rafael Rangel" Universidad de los Andes, Trujillo, Venezuela
Elci Villegas Avila
lecturer

London School of Hygiene and Tropical Medicine, London WC1E 7HT
Linda Morison
lecturer in medical statistics

Correspondence to:
A Kroeger
A.Kroeger@liverpool.ac.uk

BMJ 2002;325:810–3

Abstract

Objective To measure the impact on transmission of leishmaniasis of curtains impregnated with insecticide.
Design Cluster randomised controlled trial: household interview survey, observational study of people's behaviour, entomological study with light trap captures of sandflies inside houses.
Setting 14 urban sectors in Trujillo, Venezuela.
Participants 2913 inhabitants of 569 houses.
Intervention Sectors were paired according to their 12 month cumulative incidence of cutaneous leishmaniasis, one sector in each pair was randomly allocated to receive polyester curtains impregnated with lambdacyhalothrin (intervention group) while the other sector received curtains without insecticide or no curtains (control groups). After 12 months a follow up household survey was conducted.
Main outcome measures Reduction in abundance of sandflies indoors and 12 month incidence of clinical cases of cutaneous leishmaniasis.
Results Transmission of cutaneous leishmaniasis occurred mainly in the domestic setting, with the incidence over 12 months of 4%. The mean number of sandflies per trap per night was 16. After follow up the 12 month incidence of cutaneous leishmaniasis was 0% in the intervention group and 8% in the six pairs in the control group that received unimpregnated curtains (mean difference 8, 95% confidence interval 4.22 to 11.78; P=0.001). There were significantly fewer sandflies in the intervention group (2 *v* 15, mean difference 13 sandflies per trap; 9 to 17; $P<0.001$).
Conclusion Curtains impregnated with insecticide provide a high degree of protection against indoor transmission of cutaneous leishmaniasis.

Introduction

American cutaneous leishmaniasis is transmitted by sandflies in rain forest areas mainly among mammals and occasionally humans. However, deforestation has brought vectors and some animal hosts closer to humans, and domestic animals have emerged as alternative reservoirs resulting in an increase in cases of cutaneous leishmaniasis in urban areas.[1] This trend has been described for several countries.[2–8]

The domestication of transmission has also increased the opportunities for control. Indoor house spraying in Peru,[9] space spraying in Venezuela,[10] and curtains impregnated with pyrethroid insecticide in Burkina Faso,[11] Sudan,[12 13] and Colombia[14] have considerably reduced sandfly populations.

We carried out a study in an urban area of Venezuela with intense transmission of cutaneous leishmaniasis. We assessed any reduction in abundance of sandflies indoors and of clinical cases in areas with houses protected by curtains impregnated with pyrethroid insecticide compared with areas with houses using non-impregnated curtains or with no curtains at all. Curtains are preferred to bed nets in urban areas. We received approval for the study from the ethics committee at the Universidad de los Andes, Venezuela.

Methods

Study area

Trujillo, Venezuela, is 800 metres above sea level, has 33 399 inhabitants, and is divided into 22 sectors. It has an annual average temperature of 23.3°C and two annual rainfall periods (July and November) of 750 mm each. The city has many green areas where opossum and other woodland reservoir hosts of *Leishmania* live.

Reasons to Use Cluster Randomized Trials

- Intervention applied at cluster level
- Administrative convenience
- Cooperation of investigators
- Improve participant adherence
- Avoidance of contamination of the intervention

Disadvantages of Cluster Randomized Trials

- Loss of statistical efficiency
- Selection bias
- Decrease in cooperation of investigators

Systematic Reviews

- A synthesis of research studies that has been prepared using a systematic approach to minimising biases and random errors which is documented in the methods section. (Chalmers and Altman, 1995)
- Meta-analysis – statistical pooling of results to provide a single combined estimate of effect

Cross-sectional studies are efficient for determining causal relation between exposure and outcome.

A. True
B. False

All properly randomized trials are not susceptible to selection bias.

A. True
B. False

An investigator wishes to understand the relation between smoking and a rare variant of stomach cancer. Absent any other information, which study design would you initially propose to most efficiently examine this association?

A. Cross-sectional
B. Parallel Group Randomized Trial
C. Crossover Trial
D. Case Control
E. Prospective Cohort

Early hospital-based studies found a strong association between oral contraceptives and thromboembolism. However, given the suspicion over the risk of thromboembolism due to oral contraceptives, physicians were more likely to admit women who had symptoms if they were taking oral contraceptives. The differential probability of hospitalization depending on oral contraceptive use resulted in an overestimated odds ratio. This is an example of:

A. Selection Bias
B. Information Bias
C. Confounding

In a trial to prevent serious fall injuries in older adults, individuals were randomized to either standard of care or the delivery of care through a falls care manager (FCM). The primary outcome was serious fall-related injuries requiring medical attention. The investigators were worried that the relationship that patients had with their FCM would lead to an over-reporting of the primary outcome in the intervention group. This is an example of:

A. Selection bias
B. Information bias
C. Confounding

Smoking is a known risk factor for pancreatic cancer and coffee-drinkers are more likely to smoke. When smokers and non-smokers are combined, the risk of pancreatic cancer is 1.5 times higher in coffee drinkers than in non-drinkers. However, when analysis is restricted to either smokers only or non-smokers only, there is no elevated risk of pancreatic cancer associated with coffee drinking. This is an example of:

A. Selection bias
B. Information bias
C. Confounding

Her attempt to stay blinded was ruined by the sign-building side-effect of the treatment.

Intention to Treat

- This is a principle that arises from protocol violations in an experiment.
  - Subjects receive wrong treatment
  - Subjects do not comply with assigned treatment

- Intention to Treat Principle – follow all subjects according to the protocol regardless of actual treatment received, analyze as you randomize

- This is a conservative approach that evaluates effectiveness and provides proper statistical inference