Information Bias in Epidemiological Studies

Madhukar Pai, MD, PhD
Assistant Professor
Department of Epidemiology & Biostatistics
McGill University, Montreal, Canada
Email: madhukar.pai@mcgill.ca

Lets say you decide to do a case-control study on dietary fat and breast cancer for your thesis…

Breast cancer

Dietary fat over the past decade		Yes	No
	High	a	b
	Low	c	d

How will you estimate dietary fat intake over the past decade?
What tools could you use? How accurate and precise are these tools?
Is the study worth doing???

Misclassification of exposure

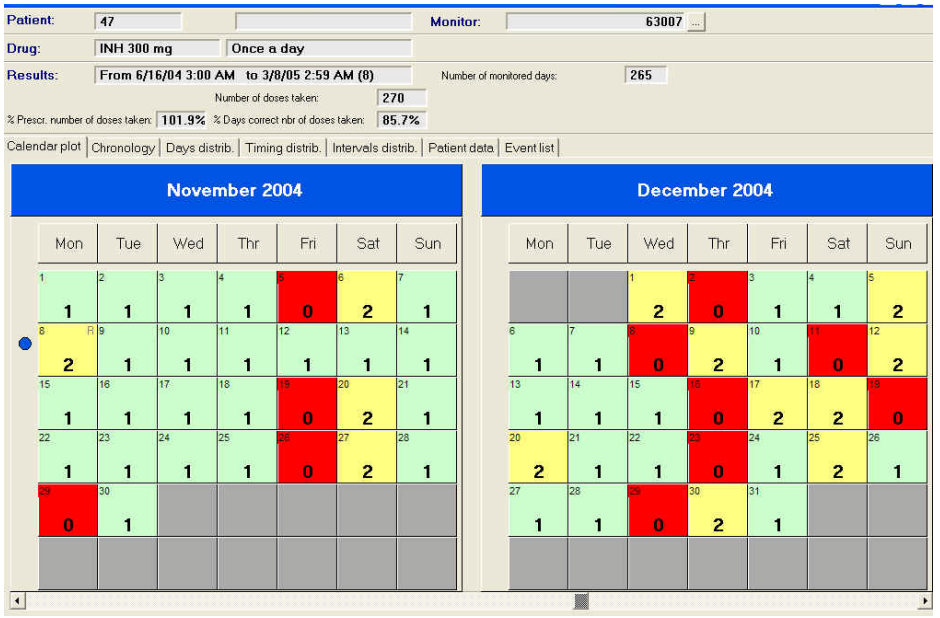
- How accurately can these commonly studied exposures be measured?
  - Age
  - Race
  - Dietary intake
  - Physical activity
  - Pain
  - Stress
  - Socioeconomic status
  - Smoking
  - Alcohol
  - Sexual behavior
  - Adherence to medications
  - Caffeine intake
  - Blood pressure
  - Intelligence

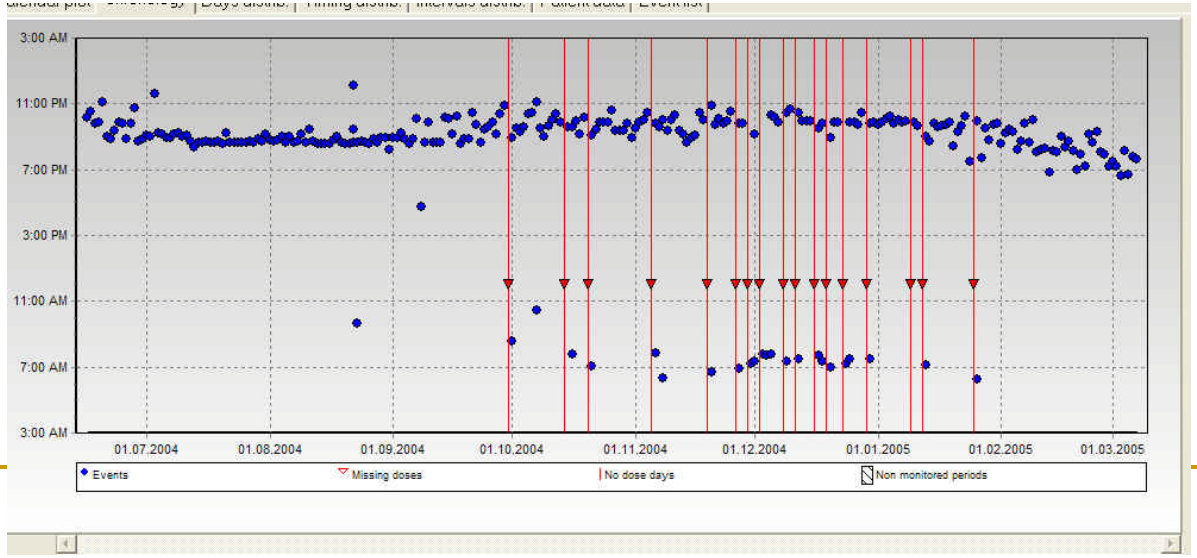
How to measure adherence?

- **Is there a gold standard?**
  - No gold standard method
- **What are the available methods?**
  - Provider's assessment of adherence
  - Self reported adherence by patient
  - Standardized, patient-administered questionnaires
  - Pill counts (e.g. remaining dosage units)
  - Pharmacy database (prescription refills, etc)
  - MEMS (medication event monitoring system)
  - Biochemical measurements (e.g. biomarkers in urine)
  - Direct observation of medication ingestion (e.g. DOT)
- **Which approach is most prone to misclassification?**
  - Provider's assessment of adherence
- **Which approach is least prone to misclassification?**
  - DOT, MEMS
- **What may be the optimal strategy, considering cost and feasibility?**
  - Overall, no single measurement strategy is optimal
    - multi-method approach that combines self-reporting with some objective measure is the current state-of-the-art in measurement of adherence

Source: WHO, 2003

MEMS: Medication Event Monitoring System





Direct observation of therapy (DOT)

Coffee: a source of great confusion and anxiety!

Home ›

HEALTH NEWS

E-mail this page Print this page

A few cups of coffee may lower colon cancer risk

Posted: 01 August 2007 1708 hrs

Photos 1 of 1

TOKYO : Drinking a few cups of coffee a day may lower the risk of advanced colon cancer, at least for women, Japanese researchers said Wednesday.

The study, supported by Japan's health ministry, showed women who drink more than three cups of coffee a day were 56 percent less likely to develop advanced colon cancer than those who drink no coffee at all.

"Drinking coffee sustains the secretion of bile acid and keeps down cholesterol levels, the mechanisms thought to prevent colon cancer," the report said.

But unfortunately the effect was not seen in men, the medical research team said.

Many men smoke and drink alcohol more than women, and those habits probably offset the effect of coffee, the study said.

The research team tracked down about 96,000 people in Japan aged from 40 to 69 between the early 1990s and 2002, of whom 726 men and 437 women later suffered colon cancer.

BBC NEWS OPEN The News in 2 minutes

lews Front Page
Africa
Americas
Asia-Pacific
Europe
Middle East
South Asia
UK
Business
Health
Medical notes
Science/Nature
Technology
Entertainment
Also in the news
Video and Audio

Last Updated: Thursday, 17 November 2005, 10:45 GMT

E-mail this to a friend Printable version

Decaf coffee linked to heart risk

Drinking decaffeinated coffee could increase the risk of heart disease, a study has suggested.

Experts say pregnant women can drink a small amount of coffee

It could lead to a rise in harmful cholesterol levels, the US National Institutes of Health study found.

The finding comes as a Danish team reiterated that drinking eight or more cups of coffee a day while pregnant may double the risk of losing the baby.

They advised pregnant women to drink no more than three cups of coffee a day, in line with existing UK advice.

How to measure caffeine intake?

Measurement of Coffee and Caffeine Intake: Implications for Epidemiologic Research[1]

GEORGE B. SCHREIBER, D.SC., CARLA E. MAFFEO, PH.D.,[2]
MORTON ROBINS, M.S.P.H., MARY N. MASTERS, M.S.P.H., AND
ANNELL P. BOND

Westat, Inc., 1650 Research Blvd., Rockville, Maryland 20850

Reported associations between coffee or caffeine intake and benign breast disease, cancers, and cardiovascular diseases have generally been weak and inconsistent. The apparent discrepancies in these studies might be attributable to imprecision in the measurement of coffee and caffeine intake. A study of a random sample of 2,714 U.S. adults disclosed considerable misclassification of total caffeine intake and, to a lesser extent, coffee intake when the estimates were limited to only the number of cups of coffee consumed. Adjustment for the following factors is recommended: amount of caffeinated and decaffeinated coffee consumed both on weekdays and on weekends; the size of the container used; the method used to brew caffeinated coffee; and the amount of caffeine imbibed from tea and soft drinks. Intake of coffee varied markedly between seasons of the year and over time. Random misclassification of coffee and caffeine intake would have the effect of obscuring dose–response relationships to disease incidence.

TABLE 1
CAFFEINE CONTENT OF VARIOUS SOURCES OF CAFFEINE INTAKE

Coffee by method of brewing[a] (mg/5 oz)	
Instant	60 mg
Drip	115 mg
Perked	85 mg
Other	87 mg
Instant, drip, and perked	87 mg
Drip and perked	100 mg
Decaffeinated	3 mg
Tea by type (mg/5 oz)	
Caffeinated	40 mg
Caffeinated and decaffeinated	30 mg
Caffeinated soft drinks (mg/oz)	3 mg
Chocolate: Estimated caffeine (mg) for frequency of use	
Daily use	20 mg
Almost daily use	10 mg
Sometimes, almost never	0 mg
Medications[b]	
Anacin, Excedrin, Vanquish	65 mg
NoDoz, Vivarin, Cafedrine	200 mg
Darvon compound, Fiorinol	75 mg
Midol, Easy-Mens, Cope	65 mg
Prolamine, Appedrine	140 mg
Pre-Mens Forte, Aqua-Ban	200 mg
Cafergot, Wigraine, Migral	200 mg

Caffeine content is multiplied by the recommended frequency of use for each medication to estimate caffeine intake

Prev Med, 1988

How to measure caffeine intake?

Psychological Reports, 2001, 89, 521-526. © Psychological Reports 2001

CAFFEINE CONSUMPTION QUESTIONNAIRE: A STANDARDIZED MEASURE FOR CAFFEINE CONSUMPTION IN UNDERGRADUATE STUDENTS[1]

KRISTI L. SHOHET AND R. ERIC LANDRUM

Boise State University

Summary.—Undergraduate students ($N = 691$) were given the 1992 Caffeine Consumption Questionnaire of Landrum and provided information on age, sex, and year in school. A subset ($n = 168$) of those completing the questionnaire were also given the Morningness–Eveningness Questionnaire of Horne and Ostberg. Analysis indicated that the average intake of caffeine was roughly 1,600 mg, i.e., a range from 13 mg to 21,840 mg per week. Older students consumed more caffeine than younger ones, and students with an Evening personality preference consumed more caffeine in the evening and nighttime hours than those with a Morning personality preference. These results are discussed in the context of other caffeine studies. Caffeine consumption is an important issue, and a consistent measurement system should be used by various researchers testing different populations.

Subjective (questionnaires based on recall)

RAPID COMMUNICATIONS IN MASS SPECTROMETRY
Rapid Commun. Mass Spectrom. 2007; 21: 2693–2703
Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/rcm.3137

RCM

Liquid chromatography/electrospray ionization tandem mass spectrometry assay for determination of nicotine and metabolites, caffeine and arecoline in breast milk

Manuela Pellegrini[1], Emilia Marchei[1], Silvia Rossi[1], Federica Vagnarelli[2], Abhilasha Durgbanshi[3], Óscar García-Algar[4], Oriol Vall[4] and Simona Pichini[1]*

[1]Department of Therapeutic Research and Medicines Evaluation, Istituto Superiore di Sanitá, Rome, Italy
[2]Arcispedale Santa Maria Nuova, Reggio Emilia, Italy
[3]Department of Criminology and Forensic Sciences, Dr. H.S. Gour University, Sagar, India
[4]Paediatric Service, URIE, Hospital del Mar, and Universitat Autònoma, Barcelona, Spain

Food Additives and Contaminants, 2001, Vol. 18, No. 12, 1075–1087

Urinary biomarkers for assessing dietary exposure to caffeine

H. M. Crews, L. Olivier and L. A. Wilson*
Central Science Laboratory, Sand Hutton, York YO41 1LZ, UK

1983, Klebsnoff *et al.* 1998, Yang *et al.* 1998). Accurate consumption data are required that can be

Objective (biomarkers)

Coffee consumption and risk of coronary heart disease: A meta-analysis

Francesco Sofi [a,d,*], Andrea A. Conti [a,b], Anna Maria Gori [a,d], Maria Luisa Eliana Luisi [b], Alessandro Casini [c,d], Rosanna Abbate [a,d], Gian Franco Gensini [a,b]

Abstract *Background and aims:* During the past three decades the relationship between habitual coffee drinking and coronary heart disease (CHD) has been assessed in numerous studies, with conflicting results. The aim of this study was to systematically examine the data published on the association between habitual coffee consumption and risk of CHD.

Methods and results: Thirteen case–control and 10 cohort studies were included. Case–control studies incorporated 9487 cases of CHD and 27,747 controls, and cohort studies included a total of 403,631 participants that were followed for between 3 and 44 years. The summary of odds ratios (OR) for the case–control studies showed statistically significant associations between coffee consumption and CHD for the highest intake group (>4 cups/day), OR 1.83 (95% CI 1.49–2.24; $P < 0.0001$), and for the second highest category (3–4 cups/day), OR 1.33 (95% CI 1.04–1.71; $P < 0.0001$), while no significant association emerged for low daily coffee intake (≤2 cups/day), OR 1.03 (95% CI 0.87–1.21; $P = 0.45$). The analysis of long-term follow-up cohort studies did not show any association between the consumption of coffee and CHD, with a relative risk (RR) of 1.16 (95% CI 0.95–1.41; $P = 0.14$) for the highest category, and 1.05 (95% CI 0.90–1.22; $P = 0.57$) and

Nutr Metab Cardiovasc Dis 2007

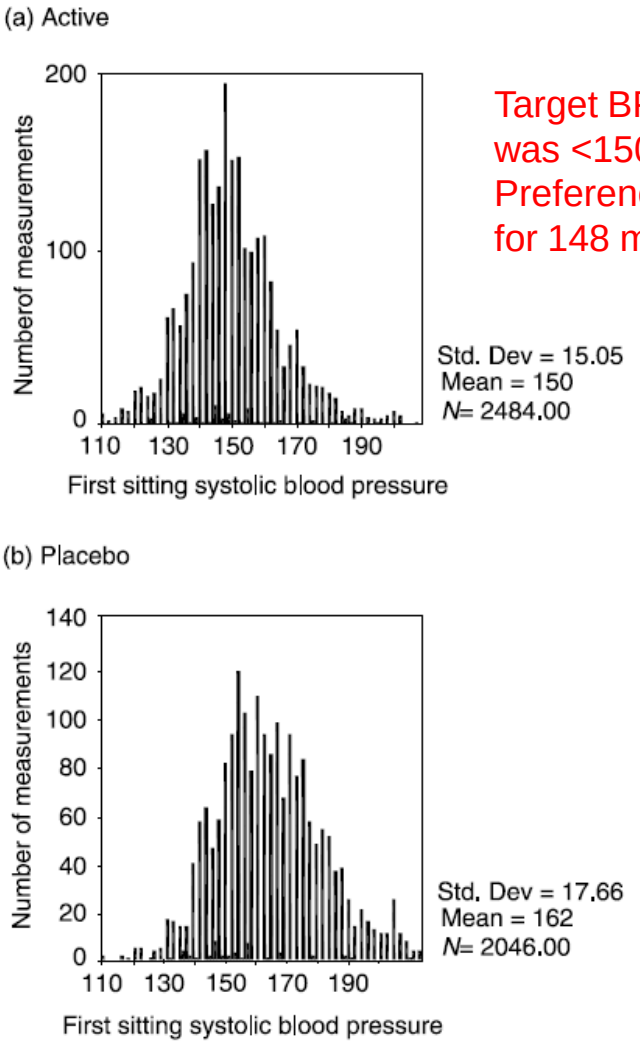
Blood pressure: digit preference bias

DATA ANALYSIS AND STATISTICAL METHODS

Terminal digit preference and single-number preference in the Syst-Eur trial: influence of quality control

David Wingfield[a], Jonathan Cooke[b], Lut Thijs[c], Jan A. Staessen[c], Astrid E. Fletcher[d], Robert Fagard[c] and Christopher J. Bulpitt[b], on behalf of the Syst-Eur Investigators

Blood Press Monit. 2002 Jun;7(3):169-77



First sitting systolic blood pressure in the fifth year following first randomization for (a) active treatment group and (b) placebo group. Note that the most common result for the active treatment mode is 148 mmHg and that the distribution is near normal.

Target BP was <150 mmHg; Preference noted for 148 mmHg

- Digit preference is a subconscious bias towards choosing numbers that end in certain digits.
- Can influence many medical readings (such as blood pressure, age, birth weight) and can reduce the power of statistical tests
- Most frequently recorded examples show preferences to figures that end in 0, 5, or even numbers

Misclassification of exposure in questionnaire studies

Table. Sources of Questionnaire Bias

Source	Bias
1. Question Design	
Problems with wording	ambiguous question complex question double-barrelled question (two questions in one) short question technical jargon uncommon word vague word
Missing or inadequate data for intended purpose	belief vs behavior (hypothetical question, personalized question) starting time data degradation insensitive measure
Faulty scale	forced choice (insufficient category) missing interval overlapping interval scale format
Leading questions	framing leading question mind-set
Intrusiveness	reporting (self-report response) sensitive question
Inconsistency	case definition change of scale change of wording diagnostic vogue
2. Questionnaire Design	
Formatting problem	horizontal response format juxtaposed scale (questionnaire format) left alignment and right alignment
Questionnaire too long	no-saying (nay-saying) and yes-saying (yea-saying) open question (open-ended question) response fatigue
Flawed questionnaire structure	skipping question
3. Administration of Questionnaire	
Interviewer not objective	interviewer nonblinding
Respondent's subconscious reaction	end aversion (central tendency) positive satisfaction (positive skew)
Respondent's conscious reaction	faking bad (hello-goodbye effect) faking good (social desirability, obsequiousness) unacceptable disease unacceptable exposure unacceptability underlying cause (rumination)
Respondent's learning	learning hypothesis guessing
Respondent's inaccurate recall	primacy and recency proxy respondent (surrogate data) recall telescope
Cultural differences	cultural

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
VOLUME 2: NO. 1 JANUARY 2005

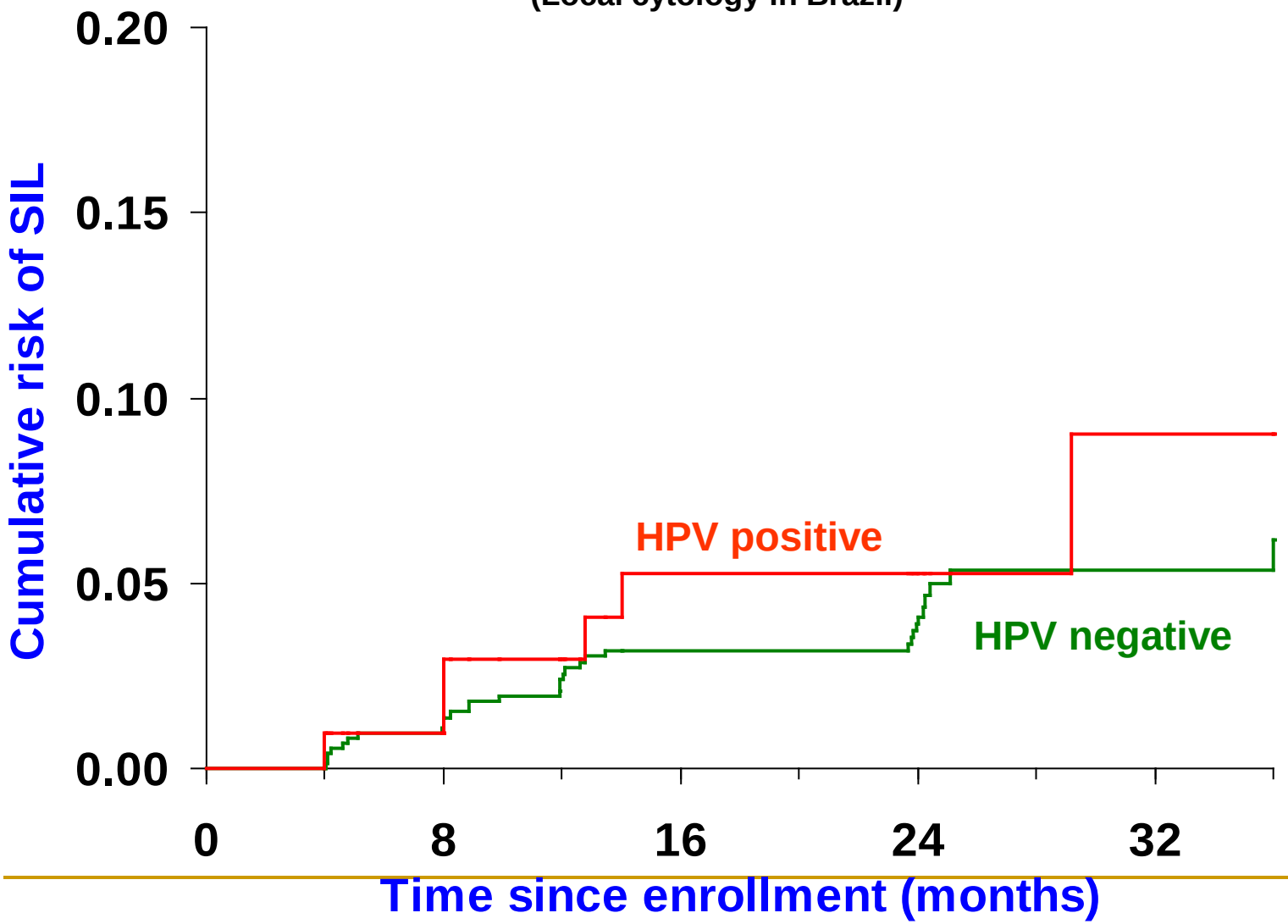
SPECIAL TOPICS

A Catalog of Biases in Questionnaires

Bernard C.K. Choi, PhD, Anita W.P. Pak, PhD

Misclassification of exposure in laboratory studies

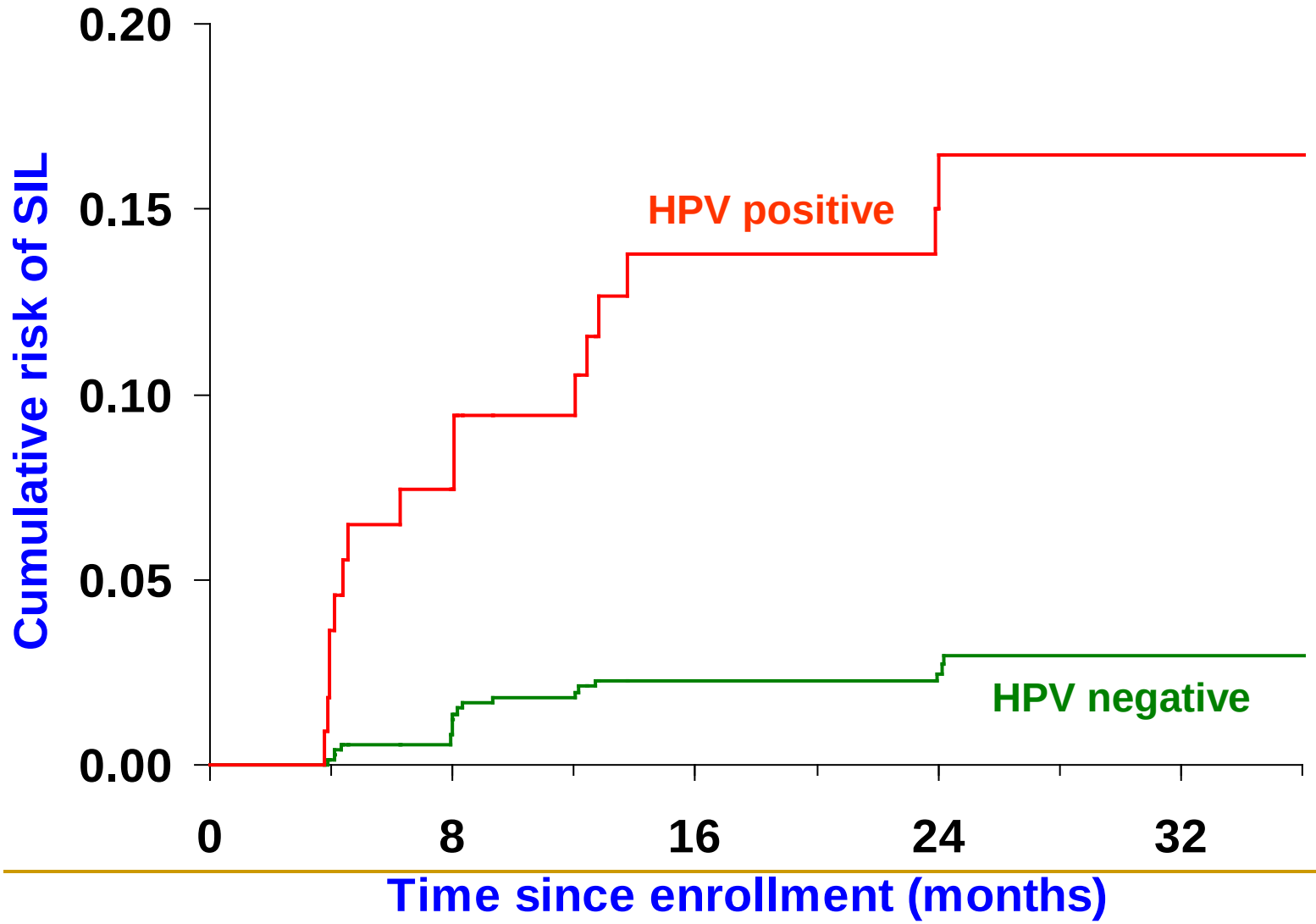
Example: Cumulative incidence of squamous intraepithelial lesions (SIL) among women with a normal Pap smear at entry
(Local cytology in Brazil)



Source: Eduardo Franco, McGill Univ.

Franco et al., PAJPH 1999; Ludwig-McGill Cohort (Follow-up data as of August 1997)

Example; Cumulative incidence of SIL among women with a normal Pap smear at entry
(Review cytology in Montreal)



Source: Eduardo Franco, McGill Univ.

Franco et al., PAJPH 1999; Ludwig-McGill Cohort (Follow-up data as of August 1997)

With better tests for HPV, the association between HPV and cervical cancer became stronger

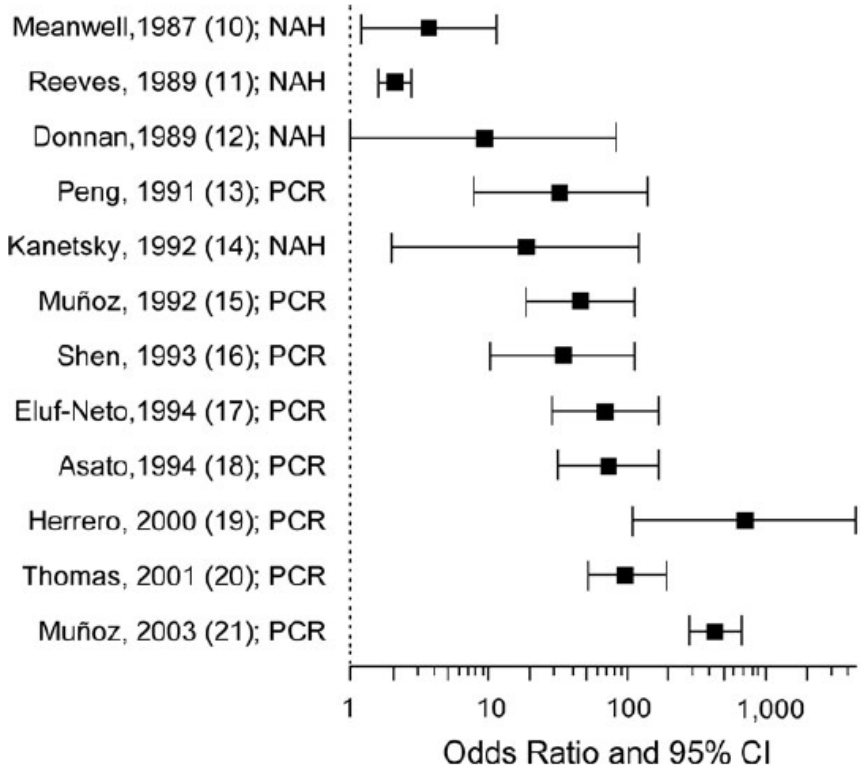


Figure 2. Odds ratios and 95% confidence intervals for the association between human papillomavirus (HPV) infection (via HPV DNA detection) and invasive cervical cancer risk in successive molecular epidemiologic studies (mostly case-control) (from top to bottom, references 10–21). CI, confidence interval; NAH, nonamplified hybridization; PCR, polymerase chain reaction.

"Studies are ordered by year of publication, which underscores the transition from nonamplified hybridization techniques to detect HPV DNA, prevailing in the 1980s, to the new era of amplified target detection via polymerase chain reaction (PCR) protocols. The graph shows that the magnitude of the association increased substantially, from 2- to 5-fold risk increases in the early studies to triple digits in the most recent investigations. "

American Journal of Epidemiology 2010 171(2):164-168;

Misclassification of outcome

- How accurately can the following be measured?
  - Depression
  - Tuberculosis in children
  - Appendicitis
  - Dementia
  - Diabetes
  - Attention deficit disorder
  - Cause of death
  - Obesity
  - Chronic fatigue syndrome
  - Angina

Measurement error: a fact of life

- Measurement error in the ascertainment of:
  - Exposure
  - Outcome/disease
  - Covariates (e.g. confounders)

- Measurement error leads to misclassification bias:
  - Non-differential misclassification bias
  - Differential misclassification bias

What is information bias?

- "A flaw in measuring exposure, covariate, or outcome variables that results in different quality (accuracy) of information between comparison groups"
- "Bias in an estimate arising from measurement errors"
  - Porta M. A dictionary of epidemiology. Oxford, 2008.
- "A distortion in the measure of effect caused by a lack of accurate measurements of exposure or disease status." [ERIC Notebook, 2001, UNC]
- Defining feature:
  - Information bias occurs at the stage of data collection
  - Misclassification of exposure and/or outcome status is the main source of error, and this, in turn, has the potential to bias the effect estimate

Example of an amazingly good measurement tool for identifying terrorists!

Do you seek to engage in espionage, sabotage, export control violations, or any other illegal activity while in the United States?	NO
Do you seek to engage in terrorist activities while in the United States or have you ever engaged in terrorist activities?	NO
Have you ever or do you intend to provide financial assistance or other support to terrorists or terrorist organizations?	NO
Are you a member or representative of a terrorist organization?	NO
Have you ever ordered, incited, committed, assisted, or otherwise participated in genocide?	NO
Have you ever committed, ordered, incited, assisted, or otherwise participated in torture?	NO
Have you committed, ordered, incited, assisted, or otherwise participated in extrajudicial killings, political killings, or other acts of violence?	NO
Have you, while serving as a government official, been responsible for or directly carried out, at any time, particularly severe violations of religious freedom?	NO

Courtesy:

US visa application

How good is the measurement tool?

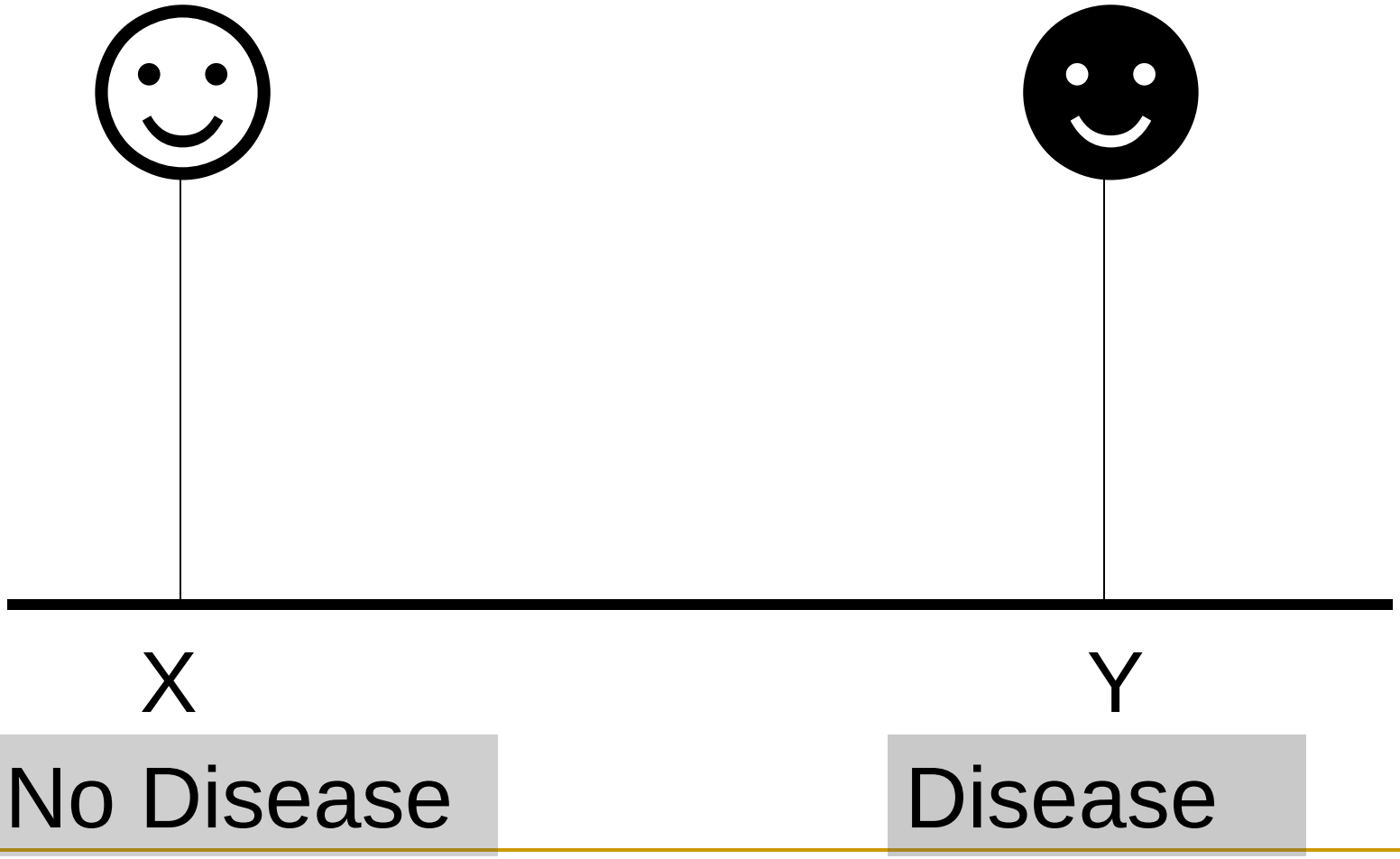
"Misclassification occurs when sensitivity and/or specificity of the procedure to detect exposure and/or effect is not perfect…"

Delgado-Rodriguez et al. J Epidemiol Comm Health 2004

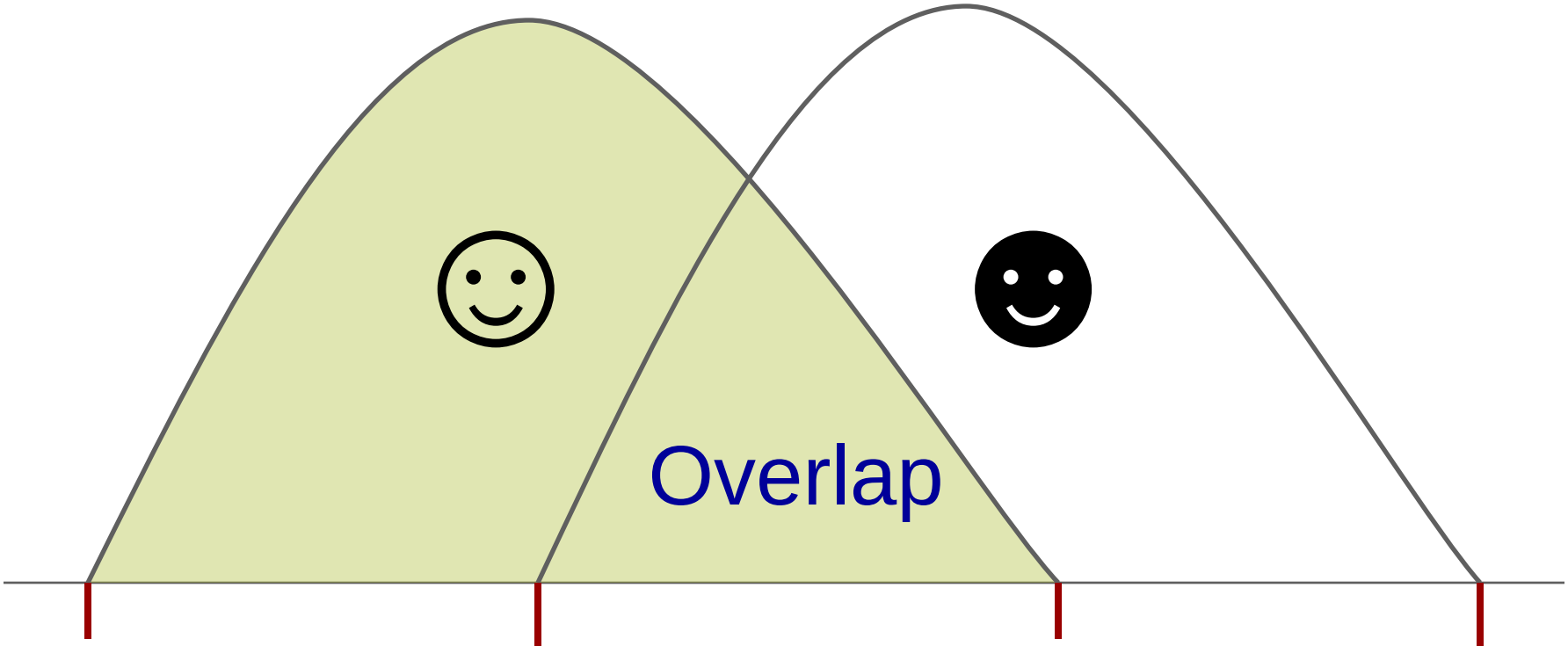
Disease +

Disease -

The ideal measurement tool (i.e. a diagnostic test) = no misclassification



Variations in test results

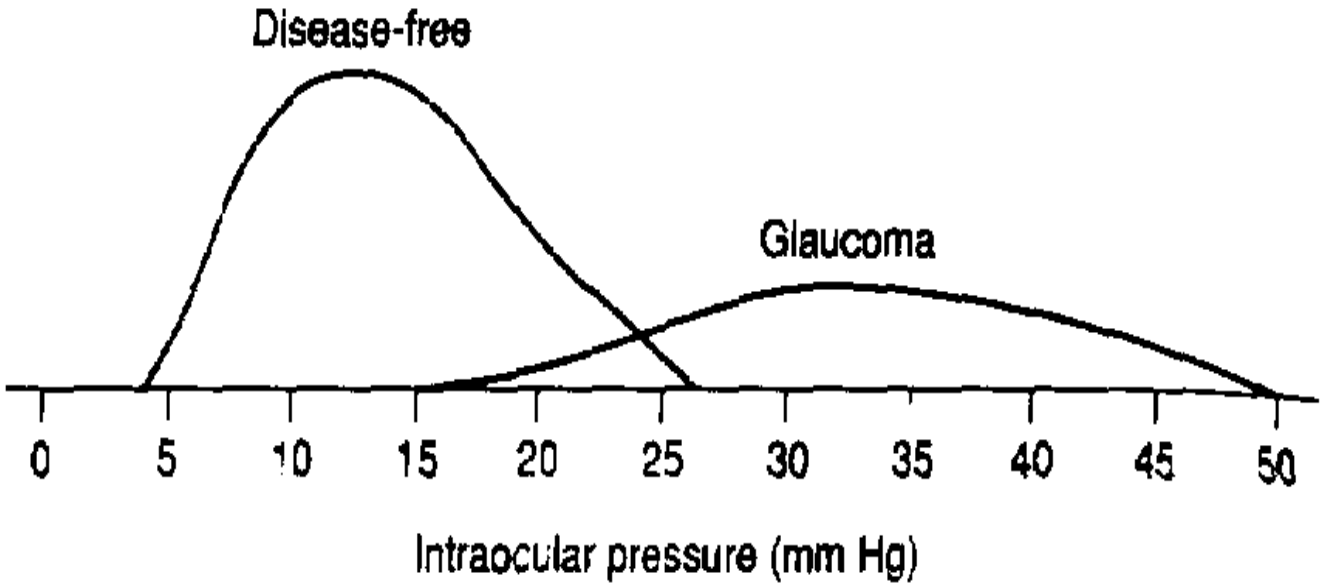


Range of Variation in Disease free

Range of Variation in Diseased

Example: intra-ocular pressure

Fig 2



Overlap of distributions of intraocular pressure among those with glaucoma and those without glaucoma

Riegelman & Hirsch 1996

Performance characteristics of a diagnostic test

- Diagnostic 2 X 2 table: need results of the "gold standard" and the index test

	Disease +	Disease -
Test +	True Positive	False Positive
Test -	False Negative	True Negative

SENSITIVITY
[true positive rate]

	Disease present	Disease absent
Test positive	True positives (TP)	False positives (FP)
Test negative	False negative (FN)	True negatives (TN)

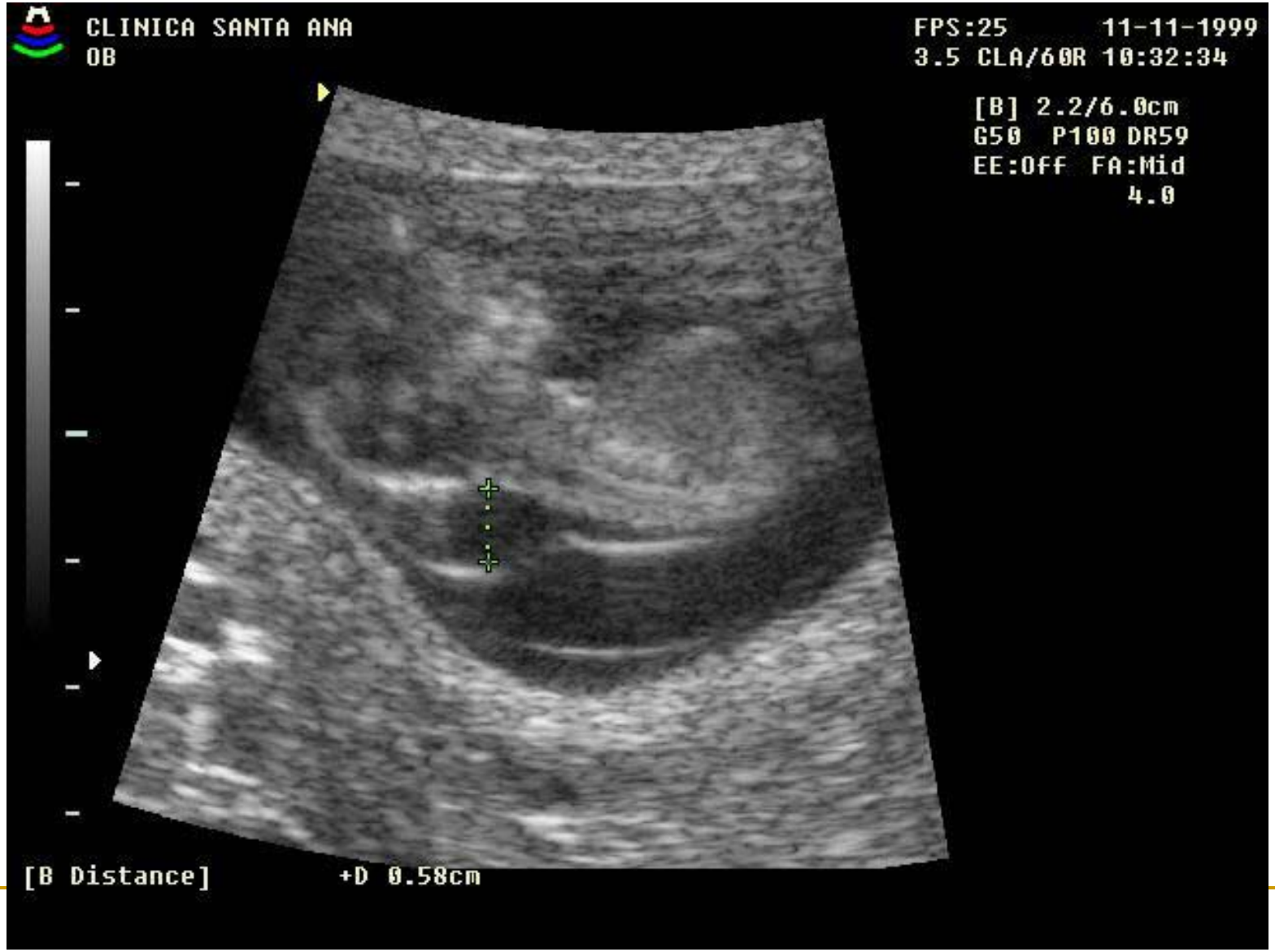
The proportion of patients with disease who test positive = P(T+|D+) = TP / (TP+FN)

SPECIFICITY
[true negative rate]

	Disease present	Disease absent
Test positive	True positives	False positives
Test negative	False negative	True negatives

The proportion of patients without disease who test negative: P(T-|D-) = TN / (TN + FP).

Example: Ultrasonography for Down Syndrome



Example: Ultrasonography for Down Syndrome

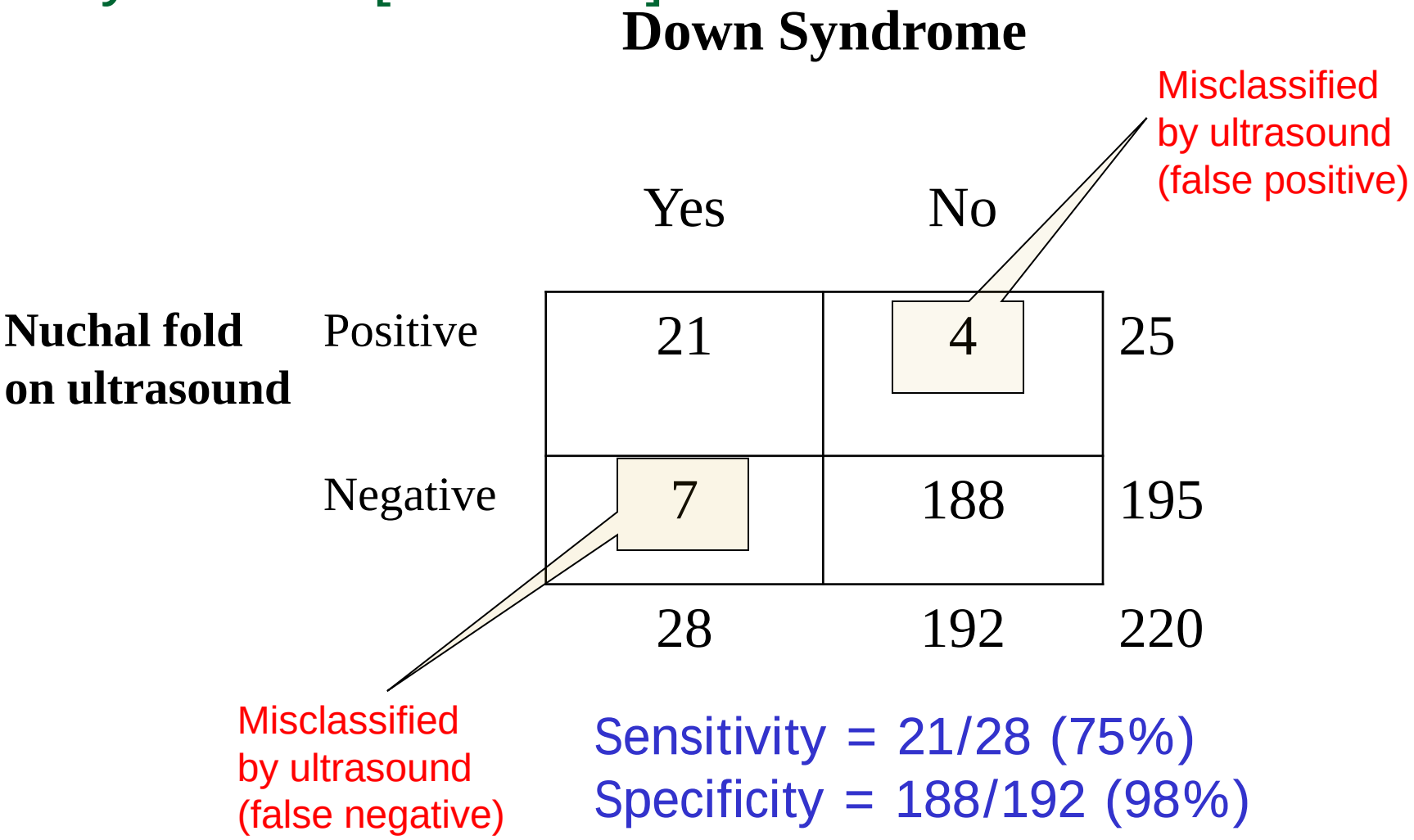
		Down Syndrome		
		Yes	No	
Nuchal fold on ultrasound	Positive	28	0	28
	Negative	0	192	192
		28	192	220

Is there misclassification in these hypothetical data?

Sensitivity = 100%
Specificity = 100%

N Engl J Med 1987;317:1371

Example: Ultrasonography for Down Syndrome [real data]



N Engl J Med 1987;317:1371

Very rarely, you get tests that are nearly perfect (i.e. 100% sensitive and 100% specific)

OPEN ACCESS Freely available online

PLoS one

Evaluation of Diagnostic Accuracy, Feasibility and Client Preference for Rapid Oral Fluid-Based Diagnosis of HIV Infection in Rural India

Nitika Pant Pai[1]*, Rajnish Joshi[2], Sandeep Dogra[3], Bharati Taksande[2], S. P. Kalantri[2], Madhukar Pai[4], Pratibha Narang[2], Jacqueline P. Tulsky[5], Arthur L. Reingold[6]

1 Immunodeficiency Service, Montreal Chest Institute, McGill University Health Center, Montreal, Canada, 2 Mahatma Gandhi Institute of Medical Sciences, Sevagram, Maharashtra, India, 3 Acharya Shri Chander College of Medical Sciences, Jammu, India, 4 Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada, 5 Department of Internal Medicine, University of California at San Francisco, San Francisco, California, United States of America, 6 Division of Epidemiology, University of California at Berkeley, Berkeley, California, United States of America

Background. Oral fluid-based rapid tests are promising for improving HIV diagnosis and screening. However, recent reports from the United States of false-positive results with the oral OraQuick® ADVANCE HIV1/2 test have raised concerns about their performance in routine practice. We report a field evaluation of the diagnostic accuracy, client preference, and feasibility for the oral fluid-based OraQuick® Rapid HIV1/2 test in a rural hospital in India. ***Methodology/Principal Findings***.. A cross-sectional, hospital-based study was conducted in 450 consenting participants with suspected HIV infection in rural India. The objectives were to evaluate performance, client preference and feasibility of the OraQuick® Rapid HIV-1/2 tests. Two Oraquick® Rapid HIV1/2 tests (oral fluid and finger stick) were administered in parallel with confirmatory ELISA/Western Blot (reference standard). Pre- and post-test counseling and face to face interviews were conducted to determine client preference. Of the 450 participants, 146 were deemed to be HIV sero-positive using the reference standard (seropositivity rate of 32% (95% confidence interval [CI] 28%, 37%)). The OraQuick test on oral fluid specimens had better performance with a sensitivity of 100% (95% CI 98, 100) and a specificity of 100% (95% CI 99, 100), as compared to the OraQuick test on finger stick specimens with a sensitivity of 100% (95% CI 98, 100), and a specificity of 99.7% (95% CI 98.4, 99.9). The OraQuick oral fluid-based test was preferred by 87% of the participants for first time testing and 60% of the participants for repeat testing. ***Conclusion/Significance***. In a rural Indian hospital setting, the OraQuick® Rapid- HIV1/2 test was found to be highly accurate. The oral fluid-based test performed marginally better than the finger stick test. The oral OraQuick test was highly preferred by participants. In the context of global efforts to scale-up HIV testing, our data suggest that oral fluid-based rapid HIV testing may work well in rural, resource-limited settings.

So, its important to note that in all epi studies:

- Exposure will be measured with some sensitivity and some specificity
- Disease will be measured with some sensitivity and some specificity
- Confounders (covariates) will be measured with some sensitivity and some specificity

- If each is measured with error, then imagine how they can all add up!

Information bias in randomized controlled trials

- Sources:
  - Lack of blinding can cause detection bias (knowledge of intervention can influence assessment or reporting of outcomes)
    - Subjects (“participant expectation bias”)
    - Investigators
    - Outcome assessors (“observer bias”)
    - Data analysts
  - Key issue: how “hard” is the outcome variable?
    - Strong versus “soft” outcomes
    - Blinding is very important for soft outcomes

Vit C and common cold

Case studies of bias in real life epidemiologic studies

Bias File 5. How blind are the blind? The story of Vitamin C for common cold

Compiled by

Madhukar Pai, MD, PhD

Jay S Kaufman, PhD

‘Hard’ Vs. ‘Soft’ endpoints

- ‘Hard’ [blinding is usually not a concern]
  - Death
  - Procedure performed (e.g. surgery)
  - Duration of hospital stay
  - Disease events that can be diagnosed with great certainty (e.g. bone fracture)
  - Laboratory results (e.g. hemoglobin, cholesterol)
- ‘Soft’ [blinding is critical]
  - Pain, stress, fatigue, etc
  - Resolution of symptoms
  - Physical signs (e.g. joint stiffness)
  - Disease events that are difficult to diagnose (e.g. angina)
  - Quality of life (QOL) indicators
  - Some side effects of drugs (e.g. rash, nausea)

Should music auditions be blinded?
The case of Abbie Conant, Trombonist

Abbie Conant

was recognized as especially talented at an early age and received a scholarship to the Interlochen Arts Academy, where she received a diploma in 1973. In 1977 she received her Bachelor´s Degree (cum Laude) from Temple University where she studied with Dee Stewart of the Philadelphia Orchestra. In 1976 she studied at Yale University, and in 1979 she received her Master´s Degree from the Juilliard School in New York City where she studied with Per Brevig of the Metropolotian Opera. In that same year she was a finalist in the Young Artists Competition in New York City. In 1979 she studied with Vinko Globokar at the L´Accademia di Chigiana in Siena. In 1984 she received a diploma from the Meisterklasse of Branimir Slokar at the Staatliche Hochschule für Musik Köln.

In 1979-1980 she was solo trombonist of the Royal Opera of Turin. From 1980 to 1993 she was solo trombonist of the Munich Philharmonic.

Home | Bios | Music Theater | Recordings | Press | Performance Inquiries | Articles Index
E-mail | Links | Demo-Video | Performances | Works | Colleagues Comment | Wired Goddess

"You Sound Like A Ladies Orchestra"

A Case History of Sexism Against Abbie Conant
In the Munich Philharmonic

By William Osborne
(Published in 1994)
(This article has won a Best of the Web Award.)

http://www.osborne-conant.org/ladies.htm

- In 1980, Conant auditioned at the Munich Philharmonic Orchestra
- 33 candidates, each played behind a screen, making them invisible to the committee
- When Conant finished, the music director cried out "That's who we want!"
- But when he found that Conant was a woman, he tried everything possible to demote her.
- He is quoted to have said ""You know the problem, we need a man for the solo trombone."
- After prolonged court proceedings, she was reinstated as first trombone and got paid on par with her male colleagues

Also recommend the book "Blink" by Malcolm Gladwell

Information bias in cohort studies

- Sources:
  - Misclassification of exposure at baseline (not likely to be influenced by outcome status, because outcome has not occurred)
  - Changes in exposure status over time (time-dependent covariates; dynamic exposures)
  - Ascertainment of outcomes during follow-up (which can be influenced by knowledge of exposure status: “detection bias” or “outcome identification bias” or “diagnostic suspicion bias”)
    - Clinical example: pathologist more likely to use the term “alcoholic cirrhosis” when evaluating a borderline liver specimen if the pathologist knows the patient is alcoholic
    - Another example: nephrologists were sent simulated case histories in which the patient’s race was identified randomly as black or white.
      - The nephrologists were 2x more likely to make a diagnosis of hypertensive end-stage renal disease if the patient was identified as black in the history

Information bias in case-control studies

- Sources:
  - Poor recall of past exposures (poor memory; can happen with both cases and controls; so, non-differential)
  - Differential recall between cases and controls (“recall bias” or “exposure identification bias” or “exposure suspicion bias”)
    - Cases have a different recall than controls
  - Differential exposure ascertainment (influenced by knowledge of case status)
    - Interviewer/observer bias (cases are probed differently than controls)

Poor recall versus recall bias

Journal of Exposure Science and Environmental Epidemiology (2009) 19, 369–381

www.nature.com/jes

Recall bias in the assessment of exposure to mobile phones

MARTINE VRIJHEID[a,b], BRUCE K. ARMSTRONG[c], DANIEL BÉDARD[d], JULIANNE BROWN[c], ISABELLE DELTOUR[a], IVANO IAVARONE[e], DANIEL KREWSKI[d], SUSANNA LAGORIO[f], STEPHEN MOORE[c], LESLEY RICHARDSON[a], GRAHAM G. GILES[g], MARY MCBRIDE[h], MARIE-ELISE PARENT[i], JACK SIEMIATYCKI[j] AND ELISABETH CARDIS[a,b]

[a]*Radiation Group, International Agency for Research on Cancer, Lyon, France*
[b]*Centre for Research in Environmental Epidemiology (CREAL), Municipal Institute of Medical Research (IMIM), Barcelona, Spain*
[c]*Sydney Cancer Centre and School of Public Health, The University of Sydney, Sydney, New South Wales, Australia*
[d]*McLaughlin Centre for population Health Risk Assessment, University of Ottawa, Ottawa, Ontario, Canada*
[e]*Department of Environment & Primary Prevention, Istituto Superiore di Sanità, Rome, Italy*
[f]*National Centre for Epidemiology, Surveillance and Health Promotion, Istituto Superiore di Sanità, Rome, Italy*
[g]*Cancer Epidemiology Centre, The Cancer Council Victoria, Melbourne, Victoria, Australia*
[h]*B.C. Cancer Agency, Vancouver, British Columbia, Canada*
[i]*INRS-Institut Armand-Frappier, Université du Québec, Laval, Quebec, Canada*
[j]*University of Montreal, Montreal, Quebec, Canada*

Most studies of mobile phone use are case–control studies that rely on participants' reports of past phone use for their exposure assessment. Differential errors in recalled phone use are a major concern in such studies. INTERPHONE, a multinational case–control study of brain tumour risk and mobile phone use, included validation studies to quantify such errors and evaluate the potential for recall bias. Mobile phone records of 212 cases and 296 controls were collected from network operators in three INTERPHONE countries over an average of 2 years, and compared with mobile phone use reported at interview. The ratio of reported to recorded phone use was analysed as measure of agreement. Mean ratios were virtually the same for cases and controls: both underestimated number of calls by a factor of 0.81 and overestimated call duration by a factor of 1.4. For cases, but not controls, ratios increased with increasing time before the interview; however, these trends were based on few subjects with long-term data. Ratios increased by level of use. Random recall errors were large. In conclusion, there was little evidence for differential recall errors overall or in recent time periods. However, apparent overestimation by cases in more distant time periods could cause positive bias in estimates of disease risk associated with mobile phone use.
Journal of Exposure Science and Environmental Epidemiology (2009) **19,** 369–381; doi:10.1038/jes.2008.27; published online 21 May 2008

Information bias in case-control studies

Carbonated Soft Drinks and Risk of Esophageal Adenocarcinoma: A Population-Based Case–Control Study

Jesper Lagergren, Pernilla Viklund, Catarina Jansson

The increased intake of carbonated soft drinks parallels the incidence of esophageal adenocarcinoma. To determine whether an association exists between carbonated drink intake and esophageal and cardia adenocarcinoma, we analyzed data from a Swedish nationwide, population-based, case–control study. During data collection in 1995–1997, 189 patients with esophageal adenocarcinoma (88% of all eligible), 262 patients with cardia adenocarcinoma (84%), and 820 control subjects (73%) were interviewed in person. All cancers were histologically classified. We calculated odds ratios with 95% confidence intervals using conditional logistic regression and multivariable analyses. Frequency of intake of carbonated soft drinks was not associated with risk of esophageal adenocarcinoma; high consumers (intake more than six times weekly) were at a statistically nonsignificantly decreased risk compared with never users (odds ratio = 0.89, 95% confidence interval = 0.49 to 1.64). Consumption of carbonated low-alcohol beer and combined intake of carbonated drinks were not associated with risk of esophageal adenocarcinoma. No association between intake of carbonated soft drinks or low-alcohol beer and risk of cardia adenocarcinoma was observed. [J Natl Cancer Inst 2006;98:1158–61]

Exposure: “How often did you on average drink carbonated soft drinks 20 years ago?”

What do you think of this exposure measurement?

Is there likely to be misclassification?

Who is likely to have poor recall – cases or controls?

Is this poor recall or recall bias?

Direction of bias: non-differential misclassification

		Case	Control
Exposure	Yes	a	b
	No	c	d

OR = ad / bc

Example: cases <u>and</u> controls have trouble recalling soft drink consumption
OR will be biased toward the null

Sensitivity and specificity for exposure is not dependent on the disease status; therefore non-differential

In general, non-differential misclassification occurs if there is equal misclassification of exposure between diseased and non-diseased subjects, or if there is equal misclassification of disease between exposed and non-exposed subjects.

Information bias in case-control studies

Carbonated Soft Drinks and Risk of Esophageal Adenocarcinoma: A Population-Based Case–Control Study

Jesper Lagergren, Pernilla Viklund, Catarina Jansson

The increased intake of carbonated soft drinks parallels the incidence of esophageal adenocarcinoma. To determine whether an association exists between carbonated drink intake and esophageal and cardia adenocarcinoma, we analyzed data from a Swedish nationwide, population-based, case–control study. During data collection in 1995–1997, 189 patients with esophageal adenocarcinoma (88% of all eligible), 262 patients with cardia adenocarcinoma (84%), and 820 control subjects (73%) were interviewed in person. All cancers were histologically classified. We calculated odds ratios with 95% confidence intervals using conditional logistic regression and multivariable analyses. Frequency of intake of carbonated soft drinks was not associated with risk of esophageal adenocarcinoma; high consumers (intake more than six times weekly) were at a statistically nonsignificantly decreased risk compared with never users (odds ratio = 0.89, 95% confidence interval = 0.49 to 1.64). Consumption of carbonated low-alcohol beer and combined intake of carbonated drinks were not associated with risk of esophageal adenocarcinoma. No association between intake of carbonated soft drinks or low-alcohol beer and risk of cardia adenocarcinoma was observed. [J Natl Cancer Inst 2006;98:1158–61]

Authors considered the possibility that cancer patients will better recall their soft drink consumption than controls:

“Risk of recall bias was alleviated by the fact that the hypothesis that carbonated drinks potentially affect the risk of these tumors was not known to the study participants”

If recall bias occurred, what would be the possible direction of bias?

Direction of bias: differential misclassification

		Case	Control	
Exposure	Yes	a ⇑	b	
	No	c	d	OR = ad / bc

Example: cases report higher soft drink consumption because they have the disease
OR will be biased away from the null

Sensitivity and specificity for exposure is dependent on the disease status; or
Sensitivity and specificity for disease is dependent on exposure status; therefore differential

In general, differential misclassification occurs when misclassification of exposure is not equal between diseased and non-diseased subjects, or when misclassification of disease is not equal between exposed and non-exposed subjects

Detection or diagnostic surveillance bias

- Exogenous unopposed estrogen (i.e. without progestin) use is now known to substantially increase the risk of endometrial cancer.
- But in the 1970s and early 80s, this was a very contentious issue. Several case-control studies reported a strong association between estrogen use and endometrial cancer, especially in women taking estrogen regularly for a number of years.
- Most investigators were convinced that this was a causal association.
- However, a few investigators argued that estrogens were merely causing the cancers to be diagnosed rather than to occur (Horwitz & Feinstein, 1978).
- In other words, they argued that "detection bias" explained the strong associations that were found in these studies.
- Estrogens induce uterine bleeding, even in healthy women. Therefore, women who regularly took estrogen are probably more likely to seek medical attention because of bleeding, therefore more likely to be worked up by physicians, thus causing a variety of gynecological conditions (including sub-clinical, symptomless or occult endometrial cancer) to be detected earlier or in some cases detected when they otherwise would have remained undetected.
- This was referred to as detection or diagnostic surveillance bias.

B-File #4 has the full story

Case studies of bias in real life epidemiologic studies

Bias File 4. The early controversy over estrogen and endometrial cancer

Compiled by

Madhukar Pai, MD, PhD

Jay S Kaufman, PhD

Recall bias

- "Systematic error due to differences in accuracy or completeness of recall to memory of past events or experiences" [Porta M, Epi Dictionary, 2008]
- Ernst Wynder, a famous epidemiologist, called this "rumination bias."
- Examples of "recall bias"
  - Ability to recall a past exposure (E) is dependent on outcome status (D)
    - Example: mothers of healthy infants vs. mothers of children with leukemia recalling perinatal exposures to household chemicals
    - Example: MMR and autism
    - Example: recall bias in case-control studies of congentinal malformations

Recall bias: example

SHORT REPORT

Recall bias, MMR, and autism

N Andrews, E Miller, B Taylor, R Lingam, A Simmons, J Stowe, P Waight

Arch Dis Child 2002;**87**:493–494

Parents of autistic children with regressive symptoms who were diagnosed after the publicity alleging a link with measles, mumps, and rubella (MMR) vaccine tended to recall the onset as shortly after MMR more often than parents of similar children who were diagnosed prior to the publicity. This is consistent with the recall bias expected under such circumstances.

The self controlled case series method[6] uses conditional Poisson regression to enable estimation of the RI using only cases by comparison of the frequency of events within and outside specified post-immunisation risk periods. In these analyses the risk periods for autism onset considered were within 2, 4, 6, and 12 months of MMR. Age was adjusted for by stratification into one month groups. In the first analysis, cases were restricted to the subset of children with core or atypical autism in whom parents reported developmental regression, with onset defined

Recall bias: example

Bias File 6. Double whammy: recall and selection bias in case-control studies of congenital malformations

Information bias in cross-sectional surveys: example

AIDS and sexual behaviour in France

*ACSF investigators**

The results of a massive telephone survey of sexual lifestyles in France should provide a basis for prevention strategies for AIDS and sexually transmitted diseases.

WITH sexual transmission of the human immunodeficiency virus (HIV) now known to be widespread in northern Europe and North America[1], many countries in these areas have decided to undertake surveys of sexual behaviour in the general population. The aim of these studies is to achieve better-defined strategies for preventing sexually transmitted diseases (STDs) and AIDS, as well as to provide the basis for more

basis for comparisons with other countries, particularly with the British study, which deals with a similar sample in a comparable country[7].

Our work began in July 1989. Three pilot surveys were carried out in the first 2 years to test the questionnaire, decide on the method of investigation (telephone or face-to-face) and to see if sending out a notifying letter affected whether people would participate in the

tive, and that the survey was being conducted by researchers employed by public-health institutions. The theme of AIDS and sexual behaviour was deliberately not mentioned in the letter to avoid worrying people, to prevent refusals before selection of the interviewee and to prevent people from preparing answers in advance. (The results of our pilot

Massive telephone survey on sexual lifestyles in France, and involved more than 20,000 participants. After pilot research, the telephonic method was selected, and involved more than 100 interviewers.

TABLE 1 Results of 20,055 questionnaires (including 4,820 long questionnaires)

	MEN Age (years)				WOMEN Age (years)			
	18–24	25–34	35–44	45–69	18–24	25–34	35–44	45–69
Population	1,716	2,232	2,284	3,696	1,670	2,238	2,261	3,958
Age at first intercourse (years)	16.5	16.9	17.6	18.2	17.1	17.9	18.8	20.8
Multipartner — 1 year	27.6	14.1	11.5	8.3	12.1	6.8	5.9	2.9
Intercourse (no.) — 4 weeks (m)	7.6	9.6	9.8	6.8	8.3	8.9	8.5	5.8
Homosexuality — life (%)	2.4	4.2	4.3	4.5	1.2	3.8	2.8	2.4
Intercourse with prostitutes — 5 years	4.7	3.8	3.4	1.8	—	—	—	—
IV drugs — life (%)	0.4	1.2	0.6	0.0	0.2	0.5	0.2	0.0
Condoms — life (%)	73.1	58.1	57.8	48.1	54.7	48.9	49.6	33.9
Condoms — 1 year	58.8	32.3	30.5	17.2	40.8	26.7	23.0	11.1
Monopartner	50.5	26.0	26.5	14.6	37.8	24.1	22.3	10.2
Multipartner (heterosex.)	79.0	69.0	59.7	42.7	62.4	62.0	33.3	33.1

Not surprisingly, the proportion of participants who admitted to using IV drugs was very low. As the authors pointed out, "people who regularly use drugs are the most difficult to contact, and/or most often refuse to participate in any kind of survey or to acknowledge an illegal practice." Social desirability bias is always a concern in these situations. Social desirability bias is the tendency of respondents to reply in a manner that will be viewed favorably by others. This will lead to overreporting good behavior and/or underreporting bad behavior.

Nature 1992

Information bias in surveys: example

There is considerable evidence that interviewer-administered surveys elicit lower self-reports of sensitive behaviors. Self-administration reduces social desirability bias and also provides anonymity. Computerization and audio-assistance may reduce measurement error.

Adolescent Sexual Behavior, Drug Use, and Violence: Increased Reporting with Computer Survey Technology

C. F. Turner,* L. Ku, S. M. Rogers, L. D. Lindberg, J. H. Pleck, F. L. Sonenstein

Surveys of risk behaviors have been hobbled by their reliance on respondents to report accurately about engaging in behaviors that are highly sensitive and may be illegal. An audio computer-assisted self-interviewing (audio-CASI) technology for measuring those behaviors was tested with 1690 respondents in the 1995 National Survey of Adolescent Males. The respondents were randomly assigned to answer questions using either audio-CASI or a more traditional self-administered questionnaire. Estimates of the prevalence of male-male sex, injection drug use, and sexual contact with intravenous drug users were higher by factors of 3 or more when audio-CASI was used. Increased reporting was also found for several other risk behaviors.

Science 1998

Table 3. Alternate estimates of prevalence of drug use, per se, and drug use during sex derived by using different methods of questioning. Results are from the 1995 NSAM.

Measurement	Estimated prevalence (per 100)		Crude OR	Adj. OR
	Paper SAQ	Audio-CASI		
Drug use				
Ever taken street drugs using a needle	1.4	5.2	3.85***	3.90*
Injected drugs within last year‡	0.0	0.8	–†	–†
Ever shared needle§	0.1	1.1	9.71**	9.56**
Smoked marijuana daily during last year‖	4.1	6.7	1.69*	2.03*
Used crack/cocaine within last year	3.3	6.0	1.89	1.96
Drank alcohol last year¶	65.9	69.2	1.16	1.29
Drank alcohol weekly last year#	15.0	19.4	1.34	1.56*
Ever smoked marijuana	41.2	43.3	1.09	1.30*
Drug use and sex (among those having sex)††				
Ever had sex with someone who shoots drugs	0.2	2.8	13.84**	17.06**
You/your partner drunk or high at last heterosexual intercourse	15.3	34.8	2.95***	3.04*
Always/often drunk or high during heterosexual intercourse last year	2.2	10.8	5.52***	5.69***
You/your partner had been drinking at time of last heterosexual intercourse	13.9	25.4	2.10***	2.14***
You/your partner used drugs at time of last heterosexual intercourse	9.7	15.8	1.74*	1.89*

For an in-depth analysis of this case study, see B-File #8

Bias File 8. Don't call my number, anymore! Bias in surveys of sexual behavior

Summary

- Non-differential misclassification of disease:
  - Sensitivity and Specificity for misclassifying disease do not differ by exposure
- Non-differential misclassification of exposure:
  - Sensitivity and Specificity for misclassifying exposure do not differ by disease
- Non-differential misclassification of BOTH disease and exposure leads to:
  - Bias towards the null

General rule:

Non-Differential Misclassification
of Both Exposure and Disease

⇓

Bias is always toward the null.

(provided no missclassification of control variables)

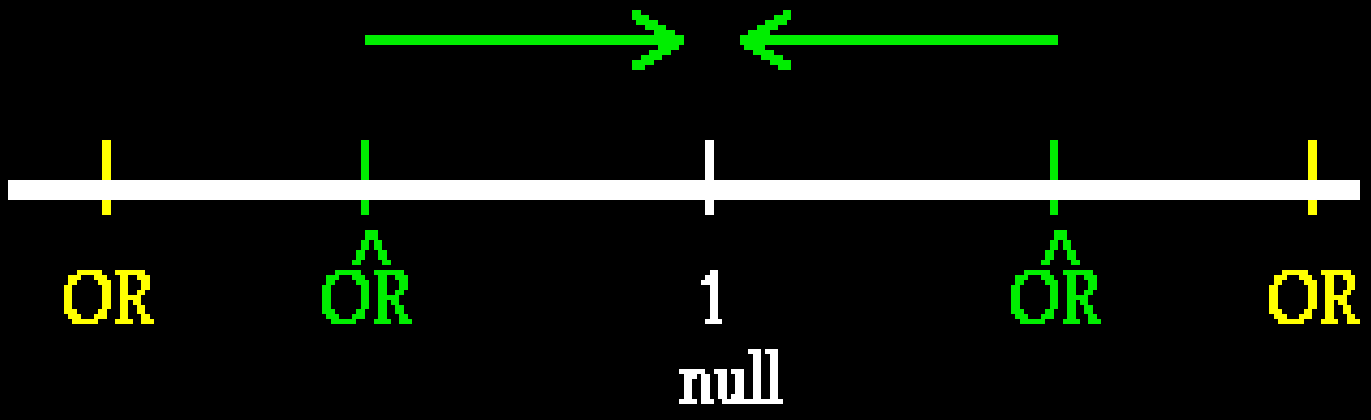


Exhibit 4–2 Hypothetical Example of the Effect of Nondifferential Misclassification of Two Categories of Exposure, with 30% of Both Exposed Cases and Exposed Controls Misclassified as Unexposed

No Misclassification

Exposure	Cases	Controls
Yes	50	20
No	50	80

$$OR = \frac{\left(\frac{50}{50}\right)}{\left(\frac{20}{80}\right)} = 4.0$$

30% Exposure Misclassification in Each Group

Exposure	Cases	Controls
Yes	50 − **15** = 35	20 − **6** = 14
No	50 + **15** = **65**	80 + **6** = *86*

$$OR = \frac{\left(\frac{35}{65}\right)}{\left(\frac{14}{86}\right)} = 3.3$$

Effect of nondifferential misclassification with two exposure categories: to bias the OR toward the null value of 1.0. (It "dilutes" the association.)

Note: Bold numbers represent misclassified individuals

Szklo & Nieto, 2007

Likely magnitude of non-differential misclassification bias

Table 13–3. Effect on the Odds Ratio of Nondifferential Error in the Measurement of a Binary Exposure Variable[a]

Sensitivity	Specificity	Prevalence of Exposure	True Odds Ratio 1.5	2.0	5.0	10.0
0.60	0.60	0.01	1.00	1.01	1.03	1.07
		0.50	1.08	1.14	1.31	1.39
		0.99	1.00	1.00	1.01	1.01
0.60	0.90	0.01	1.03	1.05	1.21	1.46
		0.50	1.24	1.42	1.99	2.31
		0.99	1.01	1.01	1.02	1.02
0.60	0.99	0.01	1.19	1.37	2.47	4.24
		0.50	1.30	1.54	2.29	2.74
		0.99	1.01	1.01	1.02	1.02
0.90	0.60	0.01	1.01	1.02	1.08	1.18
		0.50	1.26	1.48	2.40	3.16
		0.99	1.02	1.03	1.04	1.05
0.90	0.90	0.01	1.04	1.08	1.33	1.73
		0.50	1.38	1.72	3.29	4.79
		0.99	1.03	1.04	1.07	1.08
0.90	0.99	0.01	1.24	1.47	2.89	5.24
		0.50	1.43	1.82	3.63	5.42
		0.99	1.03	1.05	1.08	1.09
0.99	0.60	0.01	1.01	1.02	1.10	1.22
		0.50	1.35	1.68	3.61	6.46
		0.99	1.14	1.23	1.43	1.51
0.99	0.90	0.01	1.05	1.09	1.36	1.82
		0.50	1.45	1.89	4.44	8.35
		0.99	1.19	1.31	1.61	1.75
0.99	0.99	0.01	1.25	1.50	3.00	5.50
		0.50	1.49	1.97	4.77	9.09
		0.99	1.20	1.33	1.67	1.82

[a]The entries in the body of the table are the attenuated values of the odds ratio resulting from the effects of the nondifferential error in measuring exposure. Classification in terms of disease status is assumed to be error free.

Kelsey et al.
Methods in Observational Epidemiology
1996, Oxford Univ Press

Differential misclassification bias

- With differential misclassification, either:
  - Sensitivity and specificity for misclassifying disease differs by exposure status

  Or

  - Sensitivity and specificity for misclassifying exposure differs by disease status

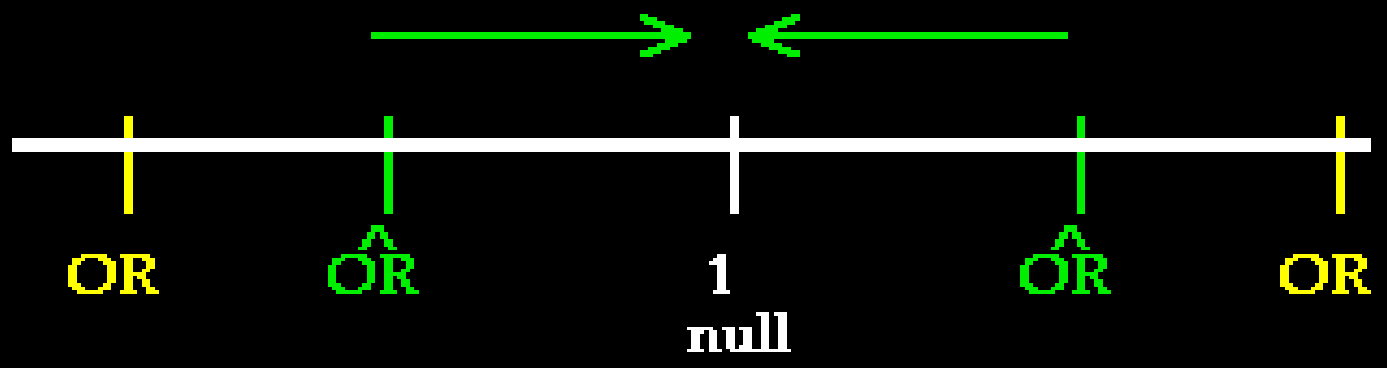
  - Differential misclassification of either disease or exposure can lead to bias either towards the null or away from the null

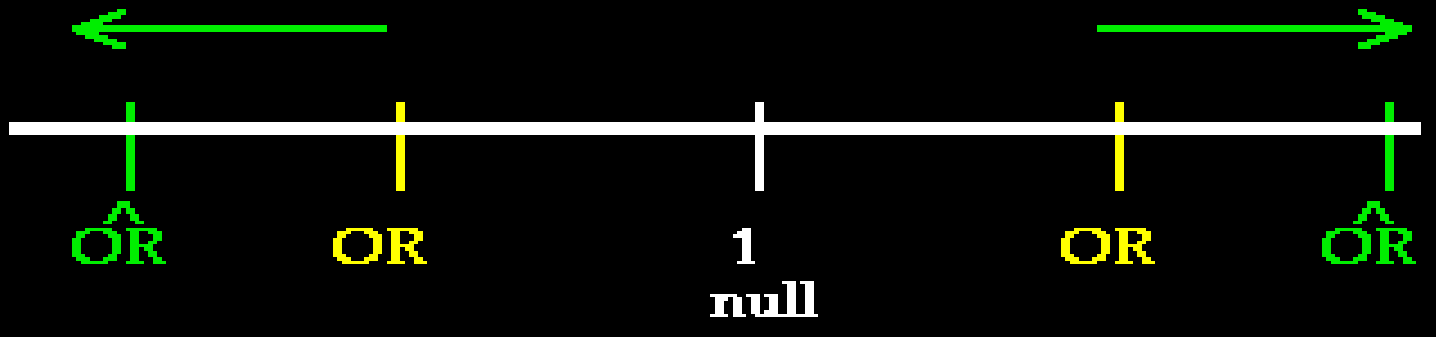
General rule:

Differential Misclassification
of Either Exposure or Disease

⇓

Bias can be Either toward the null.
or away from the null.





Reducing information bias

- Use the best possible tool to measure exposure and outcomes
- Use objective (“hard”) measures as much as possible
- Use blinding as often as possible, especially for soft outcomes
- Train interviewers and perform standardization (pilot) exercises
- Verify information using multiple sources (cross-check)
- Use the same procedures for collecting exposure information from cases and controls [case-control study]
- Use the same procedures to diagnose disease outcomes in exposed and unexposed [cohort study and RCTs]

Reducing information bias

- Collect data on sensitivity and specificity of the measurement tool (i.e. validation sub-studies)
- Collect data on reliability of measures (e.g. inter-rater agreement)
- Use a stronger study design: e.g. RCT, cohort and nested case-control where exposures are measured before disease occurs
- Correct for misclassification by “adjusting” for imperfect sensitivity and specificity of the tool (see Kleinbaum* for an excellent overview of the adjustment process)
- Perform sensitivity analysis: range of plausible estimates given misclassification (example on smoking and pneumococcal disease)

*Kleinbaum D et al. ActivEpi Companion Textbook. Springer Verlag, 2003

Correcting for misclassification

Observed (i.e., misclassified) Data

	E'	Not E'
D'	a	b
Not D'	c	d

$$\hat{OR} = \frac{ad}{bc}$$

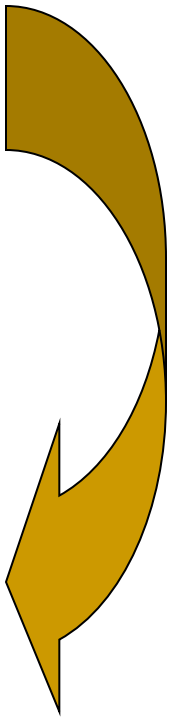
$$\hat{RR} = \frac{a/(a+c)}{b/(b+d)}$$

Corrected (i.e., adjusted) Data

	E	Not E
D	A	B
Not D	C	D

$$\hat{OR}_{adj} = \frac{AD}{BC}$$

$$\hat{RR}_{adj} = \frac{A/(A+C)}{B/(B+D)}$$



Software programs for bias analysis (sensitivity analysis)

The Stata Journal (2008)
8, Number 1, pp. 29–48

A tool for deterministic and probabilistic sensitivity analysis of epidemiologic studies

Nicola Orsini
Division of Nutritional Epidemiology
Institute of Environmental Medicine
Karolinska Institutet
Stockholm, Sweden
nicola.orsini@ki.se

Rino Bellocco
Department of Statistics
University of Milano-Bicocca
Milano, Italy

Matteo Bottai
Department of Epidemiology and Biostatistics
Arnold School of Public Health
University of South Carolina
Columbia, SC

Alicja Wolk
Division of Nutritional Epidemiology
Institute of Environmental Medicine
Karolinska Institutet
Stockholm, Sweden

Sander Greenland
Departments of Epidemiology and Statistics
University of California, Los Angeles
Los Angeles, CA

Correcting for Nondifferential Misclassification of Disease by Computer.

A computer program is available in **DataDesk**, the software package that is accessible using ActivEpi.

This is a general program that allows for differential misclassification of both exposure and disease, so that nondifferential misclassification of disease is a special case of the program in which the sensitivities and specificities for exposure given disease status are all specified to equal unity, corresponding sensitivities for disease given exposure status are specified as equal, and corresponding specificities for disease given exposure status are also specified as equal. That is, the input into the program are:

(1) Observed cell frequencies **a, b, c,** and **d**

(2) Sensitivities and specificities for exposure given disease:

$$Se_{E|D} = Se_{E|not\,D} \,(= Se_E) = 1$$

$$Sp_{E|D} = Sp_{E|not\,D} \,(= Sp_E) = 1$$

(3) Equal sensitivities for disease given exposure:

$$Se_{D|E} = Se_{D|not\,E} \,(= Se_D)$$

(4) Equal specificities for disease given exposure:

$$Sp_{D|E} = Sp_{D|not\,E} \,(= Sp_D)$$

You are not responsible for this material for the exams!

European Journal of Epidemiology (2006) 21:871–876
DOI 10.1007/s10654-006-9083-0

METHODS

Exposure-measurement error is frequently ignored when interpreting epidemiologic study results

Anne M. Jurek[1], George Maldonado[2], Sander Greenland[3] & Timothy R. Church[2]
[1]*Department of Pediatrics, University of Minnesota, Mayo Mail Code 715, 420 Delaware St. SE, Minneapolis, MN, 55455, USA;* [2]*Division of Environmental Health Sciences, School of Public Health, University of Minnesota, Minneapolis, MN, USA;* [3]*Department of Epidemiology and Department of Statistics, University of California, Los Angeles, CA, USA*

Accepted in revised form 7 November 2006

Abstract. *Introduction:* One important source of error in study results is error in measuring exposures. When interpreting study results, one should consider the impact that exposure-measurement error (EME) might have had on study results. *Methods:* To assess how often this consideration is made and the form it takes, journal articles were randomly sampled from original articles appearing in the *American Journal of Epidemiology* and *Epidemiology* in 2001, and the *International Journal of Epidemiology* between December 2000 and October 2001. *Results:* Twenty-two (39%) of the 57 articles surveyed mentioned nothing about EME. Of the 35 articles that mentioned something about EME, 16 articles described qualitatively the effect EME could have had on study results. Only one study quantified the impact of EME on study results; the investigators used a sensitivity analysis. Few authors discussed the measurement error in their study in any detail. *Conclusions:* Overall, the potential impact of EME on error in epidemiologic study results appears to be ignored frequently in practice.

Sensitivity analysis: incorporating uncertainty and exploring its effect on study results

Sensitivity Analysis of Misclassification: A Graphical and a Bayesian Approach

HAITAO CHU, MD, PHD, ZHAOJIE WANG, MS, STEPHEN R. COLE, PHD, AND SANDER GREENLAND, PHD

PURPOSE: Misclassification can produce bias in measures of association. Sensitivity analyses have been suggested to explore the impact of such bias, but do not supply formally justified interval estimates.
METHODS: To account for exposure misclassification, recently developed Bayesian approaches were extended to incorporate prior uncertainty and correlation of sensitivity and specificity. Under nondifferential misclassification, a contour plot is used to depict relations among the corrected odds ratio, sensitivity, and specificity.
RESULTS: Methods are illustrated by application to a case–control study of cigarette smoking and invasive pneumococcal disease while varying the distributional assumptions about sensitivity and specificity. Results are compared with those of conventional methods, which do not account for misclassification, and a sensitivity analysis, which assumes fixed sensitivity and specificity.
CONCLUSION: By using Bayesian methods, investigators can incorporate uncertainty about misclassification into probabilistic inferences.
Ann Epidemiol 2006;16:834–841.

Ann Epidemiol 2006;16:834–841

You are not responsible for this material for the exams!

TABLE 2. Data for the case–control study of cigarette smoking and invasive pneumococcal disease (22)

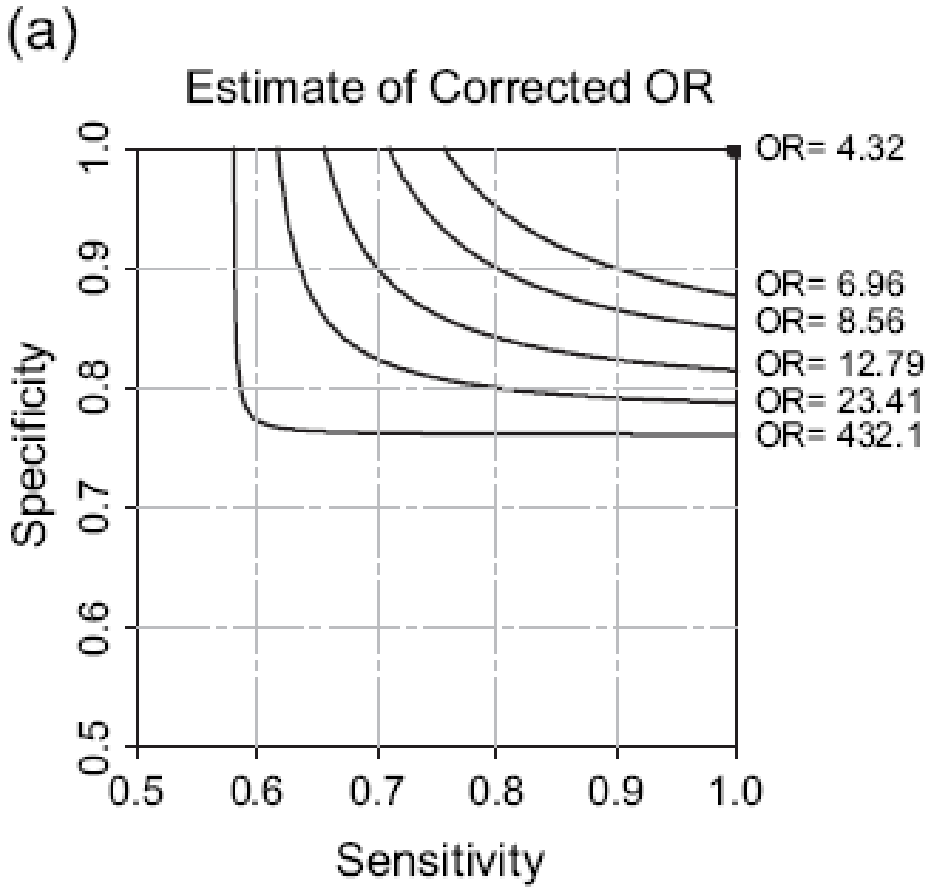
	Current cigarette smoking		
Invasive pneumococcal disease	E = 1	E = 0	Total
Case	126	92	218
Control	71	224	295

Uncorrected odds ratio estimate is 4.32 (95% confidence limits, 2.96–6.31).

- Cigarette smoking was dichotomized as current smokers (E = 1) versus nonsmokers (E = 0) based on a telephone interview.
- The uncorrected OR is 4.32 with 95% CI of 2.96 and 6.31.
- However, some subjects may erroneously report smoking status in the telephone interview.
- Based on studies using the superior cotinine validation methods, the sensitivity of self-reported smoking status ranged from 0.82 to 1.00, and specificity ranged from 0.91 to 1.00.

Ann Epidemiol 2006;16:834–841

You are not responsible for this material for the exams!



- Contour plot for the nondifferential sensitivity analysis of cigarette smoking and invasive pneumococcal disease.
- Uncorrected OR is at the upper right corner in the absence of misclassification.
- As sensitivity (Se) and specificity (Sp) decrease, the corrected OR and misclassification bias increase.
- When Se and Sp are small enough (i.e., when Se < 0.6 or Sp < 0.8), even a tiny decrease in values for Se and Sp would increase the bias greatly.
- The asymmetric shape of the contours indicates that Sp impacts on misclassification bias more strongly than Se in this example

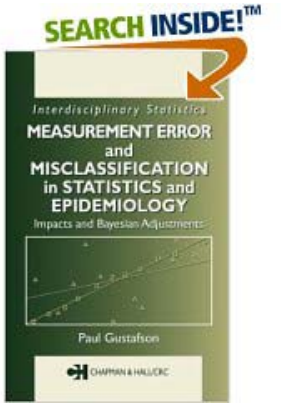
Ann Epidemiol 2006;16:834–841

You are not responsible for this material for the exams!

Good resources on exposure measurement and bias analysis

- **Principles of Exposure Measurement in Epidemiology. Second Edition. *Emily White, Bruce K Armstrong and Rodolfo Saracci***
  Oxford University Press, 2008

- **Measurement Error and Misclassification in Statistics and Epidemiology**: Impacts and Bayesian Adjustments. Paul Gustafson. Chapman&Hall/CRC (2003)



Applying Quantitative Bias Analysis to Epidemiologic Data
Springer, 2009
Lash, Timothy L., **Fox**, Matthew P., **Fink**, Aliza K.

Readings

- Gordis text:
  - Chapter 15: More on Causal Inferences: Bias, Confounding, and Interaction
- Rothman text:
  - Chapter 5: Biases in study design
- Article:
  - ERIC Notebook handout on Information Bias. UNC.

WRITING YOUR THESIS OUTLINE

NOTHING SAYS "I'M ALMOST DONE" TO YOUR ADVISOR/ SPOUSE/PARENTS LIKE PRETENDING YOU HAVE A PLAN

STEP 1 Aim for a respectable number of chapters:

THESIS OUTLINE

1.
2.
3.
4.
5.
6.
7.

← chapter #'s

5 = "That's IT??"
6-7 = "Not bad"
8+ = "Are you crazy??"

STEP 2 Fill in the "freebies":

THESIS OUTLINE

1. INTRODUCTION
2. LIT REVIEW
3. METHODOLOGY
4.
5.
6.
7. CONCLUSIONS

You're half way done!

STEP 3 Make up titles for the "meat" chapters:

6. LIT REVIEW
3. METHODOLOGY
4. (THAT STUFF YOU DID YOUR FIRST YEAR)
5. (STUFF YOU'RE SUPPOSED TO BE DOING NOW)
6. (MAKE STUFF UP)
7. CONCLUSIONS

(It'll be years before you actually have to work on that later chapter, and by then your thesis topic will have changed anyway)

STEP 4 Voilá! You just bought yourself another two years

JORGE CHAM © 2006

www.phdcomics.com