ERIC Notebook

January / February 2001 Issue 16

UNC SCHOOL OF PUBLIC HEALTH DEPARTMENT OF EPIDEMIOLOGY

Michel Ibrahim, MD, PhD
Director of Education Program

Lorraine Alexander, DrPH

Carl Shy, MD, DrPH

Sherry Farr, Graduate Research Assistant

Sandra Deming, MPH
Graduate Research Asst.

www.sph.unc.edu/courses/eric

VA ERIC EPIDEMIOLOGIC RESEARCH & INFORMATION CENTER Durham, North Carolina

Ron Horner, PhD
ERIC Director

http://hsrd.durham.med.va.gov/ERIC/

The ERIC Notebook is funded by the Department of Veterans Affairs (DVA), Veterans Health Administration (VHA), Cooperative Studies Program (CSP), to promote the strategic growth of the epidemiologic capacity of the DVA.

Information Bias

Bias occurs when an estimated association (RR, OR, difference in means, etc) deviates from the true measure of association. The source of error can be random or systematic, but the consequence of the bias is systematic error in the RR or OR estimate. Bias may be introduced at the design or analysis phase of a study.

Major types of bias

Selection bias
Information bias
Confounding bias

Information bias is a distortion in the measure of association caused by a lack of accurate measurements of exposure or disease status. Information bias, also called Measurement Bias, arises when study variables (exposure, disease or confounders) are inaccurately measured or classified. Bias in the RR or OR can be produced even if measured errors are equal between exposed and non-exposed, or between diseased and non-diseased study participants.

Non-differential Misclassification

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. If exposure or disease is dichotomous, then non-differential misclassification causes a bias of the RR or OR towards the null.

Non-differential misclassification of exposure status

Non-differential misclassification of exposure status in a case control study occurs when exposure status is equally misclassified among cases and controls. Non-differential misclassification in a cohort study occurs when exposure status is equally misclassified among persons who develop and persons who do not develop disease.

Non-differential misclassification of disease status occurs in a case-control study when disease status is equally misclassified among exposed and non-exposed subjects. Non-differential misclassification of disease status occurs in a cohort study, when a study subject who becomes diseased is equally misclassified among exposed and non-exposed cohorts.

Effect of non-differential misclassification of exposure

Non-differential misclassification always biases the RR or OR towards the null if exposure classification is dichotomous, i.e., either exposed or non-exposed. If exposure is classified into 3 or more categories, intermediate exposure groups may be biased away from the null, but the overall exposure-response trend will usually be biased towards the null.

Effect of non-differential misclassification of disease

In most cases, non-differential misclassification of disease will produce bias toward the null, i.e. the RR or OR will be biased towards 1.0. If errors in detecting presence of disease are equal between exposed and non-exposed subjects (sensitivity is less than 100%), but no errors are made in the classification of disease status (specificity is 100%), the RR in a cohort study will not be biased, but the risk difference will be biased towards the null.

Effect of non-differential misclassification of disease status

If no errors are made in detecting presence of disease (100%sensitivity), but equal errors are made among exposed and non-exposed in the classification of disease status (specificity less than 100%), both the RR and risk difference will be biased towards the null.

Combined errors in both sensitivity and specificity further increase the bias towards the null, but specificity errors produce larger biases overall.

Differential misclassification

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.

Differential misclassification causes a bias in the RR or OR either towards or away from the null, depending on the proportions of subjects misclassified.

Effect of differential misclassification of exposure or disease

Differential misclassification of exposure or disease can lead to bias the RR or OR in either towards or away from the null. The direction of bias is towards the null if fewer cases are considered to be exposed or if fewer exposed are considered to be diseased. The direction of bias is away from null if more cases are considered to be exposed or if more exposed are considered to be diseased.

Effect of differential misclassification of exposure or disease can lead to bias the RR or OR in either direction. The direction of bias is towards null if fewer cases are considered to be exposed or if fewer exposed subjects are considered to be diseased. The direction of bias is away from the null if more cases are considered to be exposed or if more exposed subjects are considered to be diseased.

Interviewer bias

Interviewer bias is a form of information bias due to:

1) lack of equal probing for exposure history between cases and controls (exposure suspicion bias); or
2) lack of equal measurement of disease status between exposed and non-exposed (diagnostic suspicion bias)

Solutions:
1) blind data collectors regarding exposure or disease status
2) develop well standardized data collection protocols
3) train interviewers to obtain data in a standardized manner
4) seek same information about exposure from two different sources, e.g. index subject and spouse in case-control study

Recall or reporting bias

Recall or reporting bias is another form of information bias, due to differences in accuracy of recall between cases and non-cases or of reporting of disease between exposed and non-exposed.

Cases may have greater incentive, due to their concern, to recall past exposures. Exposed persons, in a cohort study, may be concerned about their exposure and may over-report or more accurately report the occurrence of symptoms or disease.

Solutions:
1) add a case group unlikely to be related to exposure
2) add measures of symptoms or disease unlikely to be related to exposure

Complications in predicting direction of misclassification bias

Misclassification of confounders results in unpredictable direction of bias. Non-differential misclassification of a polychotomous exposure variable (3 or more categories) may result in bias away from null, though this is less likely than bias towards the null.

Non-differential misclassification of disease limited to a loss of sensitivity of detecting disease, without any loss in specificity, does not bias toward null, whereas a loss of specificity always biases toward the null.

Conclusions

Some inaccuracies of measurement of exposure and disease occur in all studies.

If a positive exposure-disease association is found and non-differential measurement errors are more likely than differential ones, measurement error itself cannot account for the positive finding since non-differential error nearly always biases towards the null.

Strive to reduce errors in measurement:

- develop well standardized protocols
- train interviewers and technicians well
- perform pilot studies to identify problems with questionnaires and measuring instruments
- Attempt to assess the direction of bias by considering likelihood of non-differential or differential misclassification.

Self-Evaluation

Q1: Indicate which of the following statements are true or false:

a) Non-differential misclassification of either the exposure or outcome will, in general, bias the association between the two toward the null.

b) Differential misclassification of either the exposure or outcome will, in general, bias the association between the two away from the null.

Q2: You are doing a case-control study of the effect of folic acid intake on neural tube birth defects, and are concerned about recall bias affecting your estimate of association. Which of the following potential control groups would reduce the likelihood of recall bias?

a) Women with normal births.

b) Women who gave birth prematurely.

c) Women with newborns who had other types of birth defects without a known association to folic acid intake.

Answers:

Q1a) True, one can conclude that non-differential misclassification will generally bias the estimate of the association toward the null.

Q1b) False, one cannot determine how non-differential misclassification will change the estimate. The estimate of association may change in either direction. See the example below:

Hypothetical true classification of exposure and disease status: (Note: in many studies, you many not know the unbiased distribution of exposure and disease)

	diseased	non-diseased
exposed	100	110
not-exposed	140	300

OR = (100 X 300) / (140 X 110) = 1.90

If 20 people who had the disease were misclassified as no-exposed, and 10 people who had the disease were misclassified as being exposed, the resulting table would be:

	diseased	non-diseased
exposed	110	110
not-exposed	130	300

OR = (110 X 300) / (130 X 110) = 2.31
Non-differential misclassification of exposure status in this case caused a bias of the estimate away from the null.

If 20 people who had the exposure were misclassified as diseased, and 10 people who had the exposure were misclassified as being non-diseased, the resulting table would be:

	diseased	non-diseased
exposed	90	120
not-exposed	140	300

OR = (90 X 300) / (140 X 120) = 1.60
Non-differential misclassification of disease status caused a bias of the estimate toward the null in this case.

Q2: The correct answer is c. Women who have negative pregnancy outcomes may be more likely to recall events or exposures that occurred during their pregnancy due to concern or guilt. Women who give birth to children with birth defects unrelated to folic acid intake are likely to have similar incentive or concern (due to a negative pregnancy outcome) when recalling exposures during pregnancy as mothers who had children with neural tube defects. Thus recall bias is assumed to be similar between the two groups, which should minimize the effect on the measure of association.

Glossary:

Information bias: A distortion in the measure of association caused by a lack of accurate measurements of exposure or disease status which can results from poor interviewing techniques or differing levels of recall by participants.

Non-differential Misclassification: It is equal misclassification of exposure between diseased and non-diseased subjects, or equal misclassification of disease between exposed and non-exposed subjects.

Differential Misclassification: is unequal misclassification of exposure between diseased and non-diseased subjects, or unequal misclassification of disease between exposed and non-exposed subjects.

ERIC NOTEBOOK IS PRODUCED BY THE EDUCATIONAL ARM (MICHEL A. IBRAHIM, MD, PHD, DIRECTOR) OF THE EPIDEMIOLOGIC RESEARCH AND INFORMATION CENTER AT DURHAM, NC (RON HORNER, PHD, DIRECTOR)

If you would like to receive ERIC Notebook please fill out the form below:

Name: _____

Degree(s):_____

Address:_____

City, State, Zip:_____

Telephone Number:_____

Fax Number:_____

E-mail Address:_____

Please fax to: 919-416-5836 – Attn: Beth Armstrong **Or**

Mail to : Beth Armstrong, ERIC Program Manager, VA Medical Center (152), 508 Fulton Street, Durham, NC 27705

Upcoming Topics and Special Notes:

ERIC Notebook can be viewed online at

http://sph.unc.edu/courses/eric

or

http://hsrd.durham.med.va.gov/ERIC/Education/Education.html.

Please let Beth Armstrong know which topics are of special interest to you so that we can

BETH ARMSTRONG, ERIC PROGRAM MANAGER
VA MEDICAL CENTER (152)
508 FULTON STREET
DURHAM, NC 27705