Epidemiology 168

In-class exercise on Information Bias

Instructors' Guide

1. Two automated blood cell counters are tested four times with a prepared suspension of leukocytes containing 8,000 cells per cubic millimeter ($/mm^3$). The cell counts by Device A for the four tests are: 8,300/mm^3, 8,250/mm^3, 8,200/mm^3, and 8,150/mm^3. Device B's counts are 8,200/mm^3, 8,050/mm^3, 7,900/mm^3 and 7,850/mm^3.

   a. Which device gives leukocyte counts with greater validity?

      Device B: device A has a mean of 8,225 / mm^3; device B has mean of 8,000/mm^3; the true value.

   b. Which device gives leukocyte counts with greater reliability?

      Device A: it has a smaller standard deviation (55.90, compared to 136.93 for device B), therefore higher repeatability.

   c. What considerations come into choosing between these two devices for various purposes, e.g.,

      1. Evaluating the effectiveness of treatment for a particular patient's infection by monitoring the trend in the patient's daily leukocyte count?
      2. Screening a group of workers exposed to a home marrow-depressing chemical, to detect those with leukocyte counts less than 4,000/mm^3 for further clinical evaluation?
      3. Determining the mean leukocyte count in a study group consisting of 100 sets of identical twins?

      Since device A has higher reliability, it will be superior for examining change over time – even if the actual readings are too high, the differences between measurements will be more reliably measured. On the other hand device A will tend to miss persons with a borderline low leukocyte count (e.g., 3,900/mm^3). But if device A consistently reads 225/mm^3 above the correct value then its measurements can be corrected by that amount. Since the mean from a large number of measurements with device B should be accurate it could be used to calibrate device A. Another possibility would be to use device B but plan to run multiple (e.g., 9) analyses of each specimen.

2. In a study of precursors of stroke carried out by Friedman et al. (1968), elderly patients with and without stroke were compared with respect to the frequency of previous cardiovascular

conditions. Medical charts in a large clinic were reviewed by two abstractors. The presence of physicians' diagnoses such as angina pectoris and congestive heart failure (before any stroke had occurred) was noted and recorded on special forms for later tabulation and analysis. What are some potential sources of error in this data collection procedure? Try to think of three or four.

Potential sources of error in the study data include:

- **inconsistency by the same physician over time or across patients; that is, the same physician may make a different diagnosis even when the symptoms in patient A are similar to symptoms in patient B, or;**
- **inconsistency among the physicians in the clinic in their elicitation and interpretation of information from patients and in the diagnostic criteria on which they rely;**
- **difficulties in reading and interpreting physicians' notes on charts;**
- **intra-and inter-abstractor differences in information derived and recorded for a chart.**

If these errors and differences were not equally frequent in the data for the two groups being compared (persons with and without stroke), the study could have been biased away from the null. Bias could also be introduced if the abstractors were aware of the hypotheses under study and could distinguish the patients with stroke from those without stroke. Fortunately, there were some additional, more objective, sources of information about cardiac status, such as the chest x-ray to determine heart size and the electrocardiogram. However, even these are subject to errors and inconsistencies in measurement and interpretation.

3. At the present time, would the underlying cause of death be recorded accurately on a larger proportion of death certificates of 45-year olds or of 85-year olds? Explain your answer.

   The underlying cause of death would probably be recorded accurately for a larger proportion of 45-year-olds. It is more unexpected for younger people to die, so greater effort is expended to find out the cause. Also, elderly persons often have multiple diseases present at the time of death, making it difficult to select a primary cause. Also, if an elderly patient is senile, there may be less effort or ability to arrive at an accurate diagnosis. On the other hand, the elderly get more care, so their conditions should be better known and therefore better reported.

4. In an epidemiological study of about 84,000 persons (Klatsky et al., 1977), alcohol consumption was assessed by questionnaire. Compared with people who drank <u>less than three drinks per day</u>, those drinking <u>three or more drinks per day</u> had higher mean blood pressure and greater prevalence of clinically significant hypertension.

(a) What bias would you expect in the measurement of alcohol consumption by questionnaire?

Underreporting of alcohol intake, especially by persons consuming moderate or large amounts, is commonly observed.

(b) A critic of the study suggested that this bias would invalidate the findings that consumption of three or more drinks per day was associated with higher blood pressure. Do you agree? State the reason for your answer.

Underreporting in itself, if comparable for persons with and without elevated blood pressure, could not produce the reported association. Suppose that blood pressure is not related to alcohol intake. Therefore people in the high-alcohol and low-alcohol groups would actually have the same average blood pressure. Underreporting of alcohol consumption would misclassify some who belong in the high-alcohol group into the low-alcohol group. But if the misclassification is unrelated to blood pressure then moving some people from one group to the other, in either direction, does not change the average blood pressure in either group from what it would be with correct classification.

If the reported association between alcohol consumption and blood pressure <u>does</u> exist, then misclassification of some high-alcohol consumers into the low-alcohol group would elevate the average blood pressure of the low group, making it closer to that of the high group. Thus, given the misclassification that probably occurred, the true difference in average blood pressure between the groups was probably grater than was apparent in the data. For underreporting of alcohol intake to create the appearance of a positive association betwween alcohol intake and blood pressure when in fact there is none would require that people with lower than average blood pressures underreport their alcohol intake to a greater extent than people with above average blood pressures. Under this scenario, people with low blood pressure appear to have lower alcohol intake relative to people with higher blood pressure. There is no obvious reason to believe that this was the case.

5. A psychologist has developed a test for school children that is supposed to identify children who are achieving below their potential. To validate the instrument, two psychology graduate students each administer it to the same 200 children. The data are shown below.

		Student B		
	Test result	**+**	**-**	**Total**
Student B	**+**	80	30	110
	-	20	70	90
	Total	100	100	200

(a) Assess the level of agreement between the two testers (use Cohen's kappa):

Agreement between the two raters appears in the cells of the main diagonal (i.e., the cells in the upper left and lower right). The observed agreement: (80 + 70) / 200 = 75%

Obtain the agreement expected by chance alone from the table margins (e.g., apply the column proportions to the row proportions:

Upper left cell	Lower right cell	Expected agreement
$\frac{100}{200} * \frac{110}{200} = 0.275$	$\frac{100}{200} * \frac{90}{200} = 0.225$	0.275 + 0.225 = 0.50

$$\kappa = \frac{\text{Observed - expected}}{\text{Perfect - expected}} = \frac{0.75 - 0.50}{1.0 - 0.50} = \mathbf{0.50}$$

(b) What is the advantage of Cohen's kappa over simple pecent agreement?

κ corrects for chance agreement.

References

Friedman GD, Loveland DB and Erlich SP. Relationship of stroke to other cardiovascular disease. *Circulation*, 38:533-541 (1968).

Klatsky AL, Friedman GD, Siegelaub AB, Gerard MJ. Alcohol consumption and blood pressure: Kaiser Permanente multiphasic health examination data. *New England Journal of Medicine*. 296:1194-1200(1977).