Statistics and Biomedical Research

by Robert W. Miller*

Medical observations of leukemia or Hodgkin's disease clusters in time and space have led statisticians to invent techniques to evaluate the probability that these occurrences were due to chance. A difficult computation in human genetics was simplified by a matrix devised by a non-genetical statistician, and readily adapted to newly emerging computer-processing. These experiences exemplify the special opportunity in Japan to enhance research by merging mathematical talent with biomedical observations.

The topic originally assigned to me was: Unsolved Problems in Cancer Epidemiology: Can Biostatisticians Help? The answer is yes. Biostatisticians can help immensely in all fields of biology and medicine, not only in epidemiology, but too often biologists and physicians do not know they have problems susceptible to statistical solutions. There is a need to enhance consultations with creative statisticians on major issues. I will not review such major issues in epidemiology, because the entire meeting is devoted to that subject. To me the most important unsolved problem is how to foster interest between biomedical scientists and innovative statisticians.

An Anglo-American Example

In the early 1960's, virologists hoped to show that a virus causes human leukemia. Time-space clusters of leukemia were cited as evidence of its infectious transmissibility. The trouble was that boundaries were drawn tightly around the clusters, so the rates for leukemia had to exceed normal expectation. The medical issue raised by these observations induced biostatisticians in England and the United States to devise dispassionate tests to determine if rare events cluster excessively in time and space (1, 2). Leukemia does not, and these methods helped dispel a misbelief. A main benefit from the virus leukemia research program was, thus, not detection of an infectious agent, but the invention of useful statistical procedures.

Later, other human observations generated hypotheses that a cancer of the lymph glands

(Hodgkin's disease) was transmitted through a healthy intermediary (i.e., a carrier) (3). Again, statisticians came to the rescue and demonstrated errors in the methods, devised better ones, and found no evidence that supported the hypothesis (4). The original authors then claimed that closeness of the date of onset of Hodgkin's disease among family members supported the concept that an environmental agent, such as a virus, was involved (5). Dr. Blot played a major role in showing misconceptions in the analysis of the data (6, 7). Thus a flaw was revealed in the dictum that closeness of date of onset of a disease in a family favored an environmental cause, whereas closeness of age at onset favored a genetic cause.

Two Japanese Assets

It helps immensely if physicians or biologists can call the attention of statisticians to sophisticated mathematical problems in biomedical research — problems whose solution requires imagination, training and experience in mathematics. Japan has a particular asset in this regard, for it is a nation of natural-born epidemiologists. On the two occasions when I worked here in the 1950's, all ten of the Japanese pediatricians on my staff undertook epidemiologic studies of their own as extra dividends from the main studies. For people to think epidemiologically is not rare, but for physicians to do so, is — except in Japan. In other countries, such as the United States, physicians tend to select their careers because of an interest in the diagnosis and treatment of one patient at a time. In Japan, so strong is the national inclination to think about dispersion

---

*National Cancer Institute, Bethesda, Maryland 20205

about the mean, that such concepts even pervade medicine.

There are some well-established epidemiologists in Japan, designated as such. There are many more who are unrecognized as they go about their research — unrecognized as epidemiologists by others or by themselves. I have noticed also that distinguished Japanese medical laboratory scientists often have an advantage over their counterparts elsewhere in the world because they supplement their research thoughts with epidemiologic observations.

I understand that Japanese mathematicians are among the best in the world, but statistics and biology/medicine have gone their separate ways here. I have long believed that Japan and the world would have much to gain by the interaction of these two national scientific treasures: excellence in statistics and in natural aptitude of physicians to think epidemiologically.

A Japanese Example

We have seen some signs of the benefits that may come from such intermingling. In 1961, Dr. Kudo and I were in the Department of Human Genetics at the University of Michigan. He was asked to consider a problem: is there a way to simplify the calculation of the coefficient of inbreeding, to enable technicians to obtain the value easily and accurately. The established procedure was markedly error-prone. Looking at the problem as a mathematician without the traditional beliefs of geneticists, Dr. Kudo quickly developed a paper-and-pencil matrix that did the trick. He used a grid and listed the maternal ancestors along the left margin, and the paternal ancestors across the top. Ancestors common to both parental lines were marked in appropriate boxes of the grid and, by a simple calculation, the coefficient of inbreeding could be determined. The method was published (8) and readily adapted for use by computers (9, 10) as they replaced other methods soon after. The experience illustrated how

problems in biomedicine might yield to the power of Japanese mathematical statistics.

An Objective of This Conference

U.S. and Japanese epidemiologists meet one another often, but the same is not true of their statistical counterparts. Most of the U.S. statisticians present have not previously been in Japan. The others were here to work at the Radiation Effects Research Foundation. This conference thus brings together for the first time U.S. and Japanese statisticians to discuss biomedical applications of their discipline. As the concepts of each group are revealed to the other, new insights my develop.

RERF has had a long history — 30 years — of medical statistics, the depths of which have not been as fully explored as they might be. Additional benefits may accrue from this meeting, if new ideas emerge with respect to the analysis of data here.

REFERENCES

1. Mantel, N. The detection of disease clustering and a generalized regression approach, Cancer Res. 27: 209 (1967).
2. Caldwell, G. G., and Heath, C. W., Jr. Case clustering in cancer. South. Med. J. 69: 1598 (1976).
3. Vianna, N. J., Greenwald, P., and Davies, J. N. P. Extended epidemic of Hodgkin's disease in high-school students. Lancet 1: 1209 (1971).
4. Pike, M. C., and Smith, P. G. Clustering of cases of Hodgkin's disease and leukemia. Cancer (Suppl) 34: 1390 (1974).
5. Vianna, N. J., Davies, J. N. P., Polan, A. K., and Wolfgang, P. Familial Hodgkin's disease: an environmental and genetic disorder. Lancet 2: 854 (1974).
6. Blot, W. J. Letter: Familial Hodgkin's disease. Lancet 2: 1520 (1974).
7. Mantel, N., and Blot, W. J. Is Hodgkin's disease infectious? Discussion of an epidemiologic method used to impute that it is. J. Natl. Cancer Inst. 56: 413 (1976).
8. Kudo, A. A method for calculating the inbreeding coefficient. Am. J. Hum. Genet. 14: 426 (1962).
9. Mange, A. P. Wright's coefficient of inbreeding, F, for human pedigrees. In: Computer Applications in Genetics, N. E. Moton, Ed., University of Hawaii Press, Honolulu, 1969, pp. 72-78.
10. Elandt-Johnson, R. C. Probability Models and Statistical Methods in Genetics. Wiley, New York, 1971, pp. 206-208.