

10/31/75

Sister Rosalie Bertell, Ph.D. **705965**
Carmelite Monastery
Beckley Hill
Barre, Vermont 05641
August 25, 1975

Mr. Toby J. Mitchell
Research Department
Computer Sciences Division
Union Carbide Corporation
P.O. Box Y
Oak Ridge, Tennessee 37830

REPOSITORY DOE-FORRESTAL
COLLECTION MARKEY FILES
BOX No. 506
2.38 SISTER BERTELL
FOLDER CORRESP.

Dear Mr. Mitchell:

Thank you for your letter of August 14, which was forwarded to me at my new residence in Vermont. In the future, writing to the above address will be more direct. I am still acting as senior research consultant at Roswell, but have adopted a more contemplative life-style.

I am glad you have faith in the process of professional dialogue, and also that you have not immediately written me off as "the enemy". This encourages my honest reply to your comments.

The paper: An Alternate Method for Calculating an Odds Ratio, represents my first expression of disenchantment with biomathematical application in medicine - and dates around 1970-71. Believe it or not, the paper was accepted by the Journal and then misfiled by the secretary. By the time I became suspicious about its loss, and the Journal located the paper, many new developments of thought had taken place. I regret its publication date, out of sequence with the thought development, but feel that it represents the kernel of the change in approach. Were I writing it today I would certainly develop a different emphasis.

Let me say first, that I have two basic objections to using only the standard statistical approaches. Naturally I do use them at times. First, because of their generality, they fail to take advantage of the information present in the data but not usually present for the larger class of studies to which this study belongs. For example, in the Tri-State Leukemia Survey, all known leukemia cases were interviewed, and a true random sample of controls was obtained. When considering individual types of leukemia, this control sample was relatively large compared with the case series. Generalized tests which were based on the assumption that both samples were random from the respective populations,

1011054

and that the samples were of approximately the same size, were wasteful of information which was available to us. My method was designed to take advantage of the data at hand so as to use all of the information available. I did not sacrifice reality to the requirements of a generalized statistical procedure.

The second objection which I have to the standardized approach is that it assumes a common, though unknown, risk for the whole population. This then gives a neat statistic r , which estimates this unknown risk and gives one a feeling of having solved the problem. In real life we note that the risk attendant upon exposure to an environmental hazard is anything but common for all persons in the population. It often differs dramatically for males vs. females, the young vs. the aged, the "weak" vs. the "strong". The whole theory of immunization presupposes the ability to change risk. Differential public health laws protecting women, children, etc., assume that the risk or vulnerability differs. By assuming that such a common underlying risk exists and can be calculated, we are cut off from locating the susceptible persons within the population which should be the focus of public health measures. We have "averaged" them out of existence.

If you will reflect on the example you posed to discredit my procedure (and I could provide still more alarming examples), you will note that you chose a constant underlying risk, 4, and matching sample sizes for both cases vs. controls, and males vs. females. That was rather sharp of you, because any sensible investigator knows that this is precisely where the traditional theory is best applied. Obviously, I would also use this technique in this situation. Its occurrence in practice however, is most rare. Your example could not occur in the Tri-State Survey data without the disease incidence rates for the no-exposure group differing for males and females. It is essentially a two disease situation.

Although your example fails to take advantage of the many marginal summary estimates of risk possible with large diversified samples, it will satisfy the demands of the computer program. It totally misses all the advantages of the technique, primarily designed to locate susceptible sub-groups. However, I will run it on the program and send you the output.

A fine craftsman has many tools at his disposal. He is familiar with all their advantages and their drawbacks. He chooses the tool according to the job to be done - doesn't throw them out when they are not "best" for everything.

1011055

I have no quarrel with the academic value of a generalized test. However in today's society we are asking very particular questions, have immense possibilities for collecting precise data, and have all the advantages of computer processing. It is a time that demands much more flexibility in the mathematical models and statistical techniques used.

The results of this approach to bio-medical research are not apparent in the small example given in that first methodological paper. I would call your attention to more recent articles which have come from the Department of Biostatistics at Roswell - papers by Dr. Irwin Bross, N. Natarajan, and myself. We have spent five years trying to unravel the complex interaction of variables involving exposures, host defense systems, and clinical manifestations of breakdown, as reported in the Tri-State Survey. The existence of susceptible sub-populations, of persons with markedly different risks in clinically identifiable groups, is now well documented. Traditional statistics would have dead-ended our approach long ago, and not yielded the fruitful unlocking of the radiation-aging-leukemia relationship which is now serving as a unifier of so many unexplainable details.

On page 4 you mentioned my assumption of the Poisson distribution. It is a special case, and I should have mentioned that. It is an aside - another example of using the data at hand.

Your final comment that my work should be presented in a statistical journal is rather disturbing and makes me wonder what your statistical background actually is. Woolf and Haldane both published in the Annals of Human Genetics, while Cornfield, Mantel and Haenszel all published in the Journal of the National Cancer Institute. Morton Levin's important statistic, the portion of the cases attributable to the exposure, is found in *Unio Internationalis Contra Cancrum: Acta* Volume IX, 1953, and it is relative to this statistic that the "vague" summary risk is important.

Please tell me to whom you delivered this report, and whether or not the persons requesting it are serious enough to desire dialogue. I am also curious about your department, especially since your department title seems to include computer science.

The department of Biostatistics at Roswell consists of 27 persons - 7 with Ph.D.'s in mathematics, 2 Ph.D. candidates, and others with masters degrees and/or civil service statistical clerk status. The computer services are under a different department which provides service and hardware. All computer software for our department is developed by the mathematicians of the department. We believe that we have an unusually strong bio-mathematical department, probably unique in the country. We are very happy to share our resources and talents with those who are seriously interested and capable of understanding what we are doing.

1011056

Thank you again for your straightforward approach to your assignment. I hope that your colleagues will take the same approach.

Sincerely,

Sister Rosalie Bertell

Sister Rosalie Bertell, Ph.D.

SRB/smk

cc: Donald A. Gardiner
David G. Gosslee

1011057