FIG. 1. *Nathan Mantel.*

A Conversation with Nathan Mantel

Mitchell H. Gail

Abstract. Nathan Mantel was born on February 16, 1919, in New York City. He received a B.S. degree in statistics from the City College of New York in 1939 and an M.A. degree from American University in 1956. He is a Fellow of the American Statistical Association and the Institute of Mathematical Statistics, has been elected Fellow of the Royal Statistical Society (RSS) and was recently made an Honorary Fellow of the RSS. He is also an elected member of the International Statistical Institute and a Fellow of the American Association for the Advancement of Science. He has been President of the Eastern North American Region of the Biometric Society and a member of the Council of the International Biometric Society. He has served on the editorial boards of *Risk Analysis*, *Biometrics*, *Journal of the National Cancer Institute* and *Cancer Research*. His tenure at the National Institutes of Health (NIH) included the years 1947–1974. This time was spent entirely as a mathematical statistician at the National Cancer Institute. While at NIH, he also held the position of Adjunct Professor of Biostatistics, Graduate School of Public Health, University of Pittsburgh. He was a recipient of the Superior Service Award, one of the highest civilian awards given by NIH. Since leaving NIH, he has been a Research Professor of Statistics at George Washington University and currently holds the title of Research Professor of Statistics at American University. Concurrently, from 1984 through 1990, he was Visiting Professor, Neuroepidemiology, at Temple University School of Medicine.

Gail: Could you tell us about your experiences before coming to NIH and about your training in statistics?

Mantel: I got my Bachelor's degree at the City College of New York in 1939. They didn't recognize it, but I said that my B.S. was in statistics. I took statistics courses in the Economics Department and in the Psychology Department. When I got my degree, things were not all that easy. During the Depression you were uncertain that you would ever get any work at all; and being able to get some work which actually involved mathematics, which was my bent, was unusual. After I got my degree, I left New York City to take a position as Assistant Messenger with the Securities and Exchange Com-

mission for $1,080 a year. I held various other jobs until I was drafted into the Army in late 1942. I had one of the highest ratings in my Army Division, but I was turned down for the Army Specialized Training Program and for aviation cadet in meteorology. Eventually I got transferred to the Statistical Control Office at Wright Field (later Wright Patterson Air Force Base) in Dayton, Ohio.

Gail: How was it that you finally came to NIH?

Mantel: Well, after I was separated from the Army, I worked for the War Production Board. Later, after applying for unemployment insurance under the Veterans Benefit Program, I came out to the National Cancer Institute for an interview with Harold Dorn and was recruited by him to work there. Dorn's group was not actually with the National Cancer Institute at that time; his group was detailed from the Division of Public Health Methods in downtown Washington to the National Cancer Institute. Eventually, all of us who were with this group under Dr. Dorn did get absorbed into the National Cancer Institute.

Mitchell H. Gail is Chief, Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, Maryland 20892.

When I arrived in November, 1947, there was one old-timer there, Jerry Cornfield. I imagined that he had been there for a long time, and it wasn't until many years later that I learned that he had arrived only a month before me. Jack Lieberman had arrived the preceding year, and in the succeeding year, in 1948, Marvin Schneiderman and Sam Greenhouse came on board.

Gail: I'd like to hear about how the group of you that arrived at that time worked together.

Mantel: We were there in order to provide help to anyone, whether in the Cancer Institute or anywhere else at the National Institutes of Health, who needed statistical assistance. So, the five of us were in one room, with one desk for each, and whoever answered the phone when a call came in ended up helping the client. We didn't automatically do things together. My first interaction with Cornfield was when he claimed to have found something incorrect in something R. A. Fisher had published, and he asked me to look at it. This problem was related to probit analysis, which was new to me. In fact, it was new to all of the statisticians. The investigators in the laboratories, who had the experience with it, were people like Ray Bryan or Harry Eagle. Cornfield was going to drop a line to R. A. Fisher, citing his error, and he asked me to look at it. I didn't understand what was going on, but I did see something wrong with what Cornfield was doing, or what I thought was wrong. Cornfield didn't agree that there was something awry, but he never did send that letter. However, we did in time publish the Cornfield–Mantel paper [1].

Gail: I was struck by a comment that I read in your very nice paper [17] called "A personal perspective on statistical techniques for quasi-experiments," which was published in 1976. In this paper you mentioned the fact that the statisticians that came to NIH needed to learn from the laboratory men; and in that paper you specifically mentioned Harry Eagle or Ray Bryan. I am quoting, "They were already using statistical procedures of which we were in ignorance." "We" (meaning the statisticians) "had to get our learning from them initially, even if we subsequently refined it." Can you tell me about some of the laboratory scientists that you worked with and about the projects in which you were involved?

Mantel: In the earliest work that I did with laboratory scientists, they would come around, drop off the data and discuss things with me. Then we might do an analysis of variance; but generally, we were just giving advice. One of these encounters did

turn into a continuing relationship, that with Abraham Goldin. Sam Greenhouse and I worked with Abe Goldin, but after a while Sam got dropped from that. The earliest work I did with Abe was related to probit analysis.

Gail: Some of your earliest work with Goldin was on antileukemic agents. What were the problems that you were facing in developing antileukemic agents?

Mantel: The question was how do you determine whether or not a certain chemical was good or effective? The practice had been to compare two agents on a milligram-per-kilogram basis. They would say, "Well, what will this amount of the chemical do?" I indicated that this was not a proper comparison; that is, you should judge that agent to be superior if it did more damage to the leukemia for a given amount of damage to the host. So this was quite a departure from the standard practice, and we developed a way of plotting the data to answer the question: for a given dose, what was the antileukemic effect as compared with what was the damage to the host?

There had been a practice of estimating the therapeutic index for an agent. But the therapeutic index did not relate to the effectiveness of the agent. They would, for a given preparation, compare the dosage which had an effect against the leukemia with the dosage that had a toxic effect on the host. The therapeutic index was the ratio of the dosage damaging to the host to the dosage damaging to the tumor. But how much damage there was to the host and how much damage there was to the tumor did not come into this. So the therapeutic index was really a measure of the safety of the compound. How likely were you to do damage to the host when you were trying to do damage to this tumor? And, what we did was to change that to ask "How much damage did it do to the tumor relative to the damage to the host?"

Gail: That's interesting. Were these experiments done primarily in mice, or did you do experiments in humans too?

Mantel: Mainly in mice, but there was a connection to work in humans. Dr. Goldin was aware of one of these compounds being used for treating patients with malignant tumors. We were concerned with the fact that you could protect against the antihost effect of a chemical by using folic acid or citrovorum factor to protect the host. The agents that we were using were aminopterin and, later, methotrexate. We tried using these agents to improve therapy, but, at first, instead of making things better,

we found it made things worse. Then we came up with the idea of giving the antifolate (methotrexate) first. After a delay (we assessed different periods of delay), we found that if you put in the antifolate first, it would do damage to the tumor while the host would survive on its own natural defenses. When you came in later with citrovorum factor, the host would be protected after the tumor had already been damaged by the chemical. And that got to be popular in clinical use; they would call it citrovorum rescue.

Gail: So the idea of timing the use of the two agents and of rescue after the initial agent had a chance to work, that hypothesis came from laboratory experimentation?

Mantel: Yes.

Gail: Did you do other work in combination therapies?

Mantel: Yes. In the very first work, where you gave the preparation just once, you could see what the damage was to the tumor for a certain amount of damage to the host. Later on, when we gave repeated treatments, we could no longer use that idea. Instead we titrated the chemical. At first, as you used more chemical, the survival time of the animals would increase up to a point at which the chemical was damaging, and then with further increases the survival time would go down. Thus we could compare two different preparations and two different dosage schedules by the maximum survival time.

The original idea was changed once we went on to multiple administrations of the same chemical. But then this notion was applied to drug combinations. We regarded a given proportion of the two chemicals as a new chemical in its own right. And so we would take various ratios of, say, agent A to agent B, and try it along different levels of agent B, and in that way examine the plane and see where the peak survival time was—what combination was giving us the best results. That is, was it in the plane or was it on one of the edges? We said that we would have therapeutic synergism if the best result was through a combination of two different drugs rather than a single one. While I was in the Army working with insecticides, under their usual definitions a given preparation could be its own best synergist, because giving a little more of the same insecticide could improve things. So, for the work with Goldin, I wanted a method which would not falsely find synergisms when it would be better to increase the dose of one agent.

Gail: I noticed in some of your earlier papers that not only did you work on developing antitu-

mor agents, but you were also interested in carcinogens.

Mantel: The work with the carcinogens began when there was a concern that aminothiazole in cranberries could have bad effects and produce cancers. I don't know if the Secretary for Health, Education and Welfare or someone at the Cancer Institute suggested a need for a protocol for testing these possible carcinogenic effects, but Beau Mider (G. Burroughs Mider), who was the Scientific Director at the Cancer Institute at the time, approached Ray Bryan and me to develop protocols for seeing whether or not an agent would produce cancers in humans. In response to his request we prepared a memorandum in which we brought out that you could never show that an agent was not carcinogenic, because no matter how big a study you had (for instance, even if you took 1,000 animals in the test group) and you found no tumors, you could only say, for example, it's likely that the risk is less than half a percent. We then came up with the idea that instead of trying to show something was absolutely safe, we had to be willing to take some level of risk. At first, we said suppose we want to allow a level of risk to be one in a hundred million. Could we come up with a dosage about which we could say it is unlikely that this dosage has a risk of more than one in a hundred million? Our procedure was, for each dosage level that we employed, whatever result it gave, to come out with an upper limit on the risk.

Taking chance into account, we would ask what is the upper limit on the risk? Then, we took an arbitrarily shallow slope of one probit per 10-fold dose increase, and we extrapolated backwards to the dosage level at which the risk would be less than one in a hundred million. This became known as the Mantel–Bryan procedure [19]. It was very popular for some time, but people were concerned that this was too conservative a procedure and that nothing would seem safe under it. Well, it was considered to be too conservative until some other group said no, this is far from conservative, this is radical. They were using a simple linear model through the origin, so that no matter what result you got the agent would seem unsafe. With this procedure, almost nothing could pass.

Gail: They would use a simple linear extrapolation of risk versus dose, instead of probit versus log dose?

Mantel: Yes. And this slope that we took of one probit per 10-fold dose increase we assumed was very shallow, since in actual experience slopes might

run 3, 4, 5, 10, 15, 20, 50, 60 probits for a 10-fold dose increase.

Gail: How did you come to be involved in retrospective studies, or case–control studies?

Mantel: Bill Haenszel wanted to write something on the use and analysis of retrospective data, and he had been much involved in the question of showing a connection between smoking and lung cancer. He wanted to prepare something that people could use in general. So, he asked me to help out. I had not been involved before with such epidemiologic data, but then I did write something. Apparently, Haenszel liked it so much that he made me the first author.

Gail: Are you referring to the famous paper in 1959 in the *Journal of the National Cancer Institute* [21]?

Mantel: Yes. Had we sent it to a statistical journal, it would no doubt have been rejected. However it came about, we seemed to have unknowingly anticipated a lot of different things. It turned out that the procedures in the paper could be extended so that they met perhaps 90 to 95% of the kinds of problems that people were encountering.

Gail: Are you referring to the stratification-type analyses?

Mantel: In part. There were three legs on which the Mantel–Haenszel procedure rested. One was the idea of stratification; one was the idea of coming up with a χ^2 procedure; and the third one was that of providing a summary odds ratio.

Gail: You said in that paper that the primary goal was to reach the same conclusions in a retrospective study as would have been obtained from a forward study, had one been done. It would seem to me that thinking in the prospective "forward way" was what we had been taught to do, and turning the problem around does seem very novel and counterintuitive. I wonder if there was a lot of controversy.

Mantel: Well, at the time we didn't have any choice. If we wanted to test whether smoking caused lung cancer, we could expose people to cigarette smoke and wait 40 years or so. But the retrospective study was done to save this time and get a quick consensus.

Gail: So, first it was a question of necessity. Did you ever become convinced of the validity of studies of this type, vis-à-vis the forward study?

Mantel: There are lots of problems you can get into when you use a retrospective method. I think Neyman pointed out one of these difficulties. Let's suppose you determine the patients who are observed with a certain condition and whether or not they had a certain exposure. Suppose that you had a situation in which the exposure increased the survival time of the people. You might then find the fact that the exposure prolonged survival could become misinterpreted as suggesting that the exposure causes disease because most of the unexposed individuals with the disease would have died before observation was made, leaving a disproportionate number of exposed individuals among the cases. So, you have to be careful about the way the cases and the controls are sampled and how they come to you.

Another interesting fact is that there is always the chance that some confounding factor that you did not consider led to an observed association. I think in the Mantel–Haenszel paper we had required that you should get a relative risk of at least 1.5 as a protection against any of these confounding factors. I think somewhere or other Cornfield had mentioned a relative risk of 2. I have been reviewing data on the passive smoking–lung cancer relationship. A metaanalysis of many epidemiologic investigations yields a relative risk of 1.2 with a lower limit on this at 1.04. That 1.04 could very easily be due to other confounding factors as, for instance, the fact that some of the women who claimed to be nonsmokers or never to have been smokers, actually were current or even ex-smokers. However, the advocates of this hypothesis could not think of any factor which should have been taken into account. My opposite position is that there could be a factor that you don't even know about. That was the point of the 1.5 relative risk requirement, to cover unanticipated factors.

Gail: Were there statisticians outside NIH who had an important influence on your work?

Mantel: Yes, David Blackwell. He had taught at the Graduate School at American University. Before my veteran's privileges on graduate education ran out, I signed up to take some course work. I got my Master's degree in 1956.

Gail: You published almost 400 articles and letters on an enormous range of topics, including topics outside the biological sciences. You published statistical methodology in 10 or 12 areas ranging from case–control theory to contingency tables to permutational methods for space–time clustering. How did you get involved in such a wide range of topics?

Mantel: Well, I generally don't generate ideas of my own. Someone has to come to me with a problem. And, apparently, I'm pretty good at coming up with solutions or ideas for solutions. Identifying problems is what is important—solutions just follow.

Gail: Most of your experience has been with observational data. Comparing that type of study with

experimentation, where the experimenter actually-controls and usually allocates a treatment by randomization, your papers seem to indicate that observational data have played a major role in medical research; and that sometimes the controlled experiments don't always go very well. In fact, you had a comment in your paper on "personal perspectives" [17] that said the simplicity of analysis of data from a designed study is largely illusory. I wondered if you would elaborate on these ideas.

Mantel: In that paper I said that mankind has been learning from observational data for the longest time. It is only recently that he has learned to do experiments to test things. But, if we are going to get results from observational studies, we need to recognize that there are a host of factors that might affect them. And, we must look at a wide range of measurements that we might take on the individual just in case they might have an effect. In contrast, when we do a designed experiment, we don't bother with looking at such covariates; we just randomize those out. In observational data, there is always someone who will say, "But, did you take this into account, or that into account?" But, in observational data we can't measure every conceivable covariate.

Gail: Do you think that even in a designed experiment there is an obligation at least to look at some of these other factors?

Mantel: Sometimes, even in a designed experiment imbalances can occur that may affect a result. For example, after looking at a data set, I might see that in one group there are an unusually large number of males. I would point out to the investigators that even though they had randomized the individuals to treatments, or claimed that they had, I could still see that there was something unbalanced. And the response that I would get was, "Well, we randomized and therefore we don't have to bother about it." But that isn't true. So, as long as the imbalance is in an important factor you should take it into account. Even though it is a designed experiment, in working with humans, you cannot count on just the fact that you randomized.

Gail: Whenever I do work with case–control studies these days I rely on two areas that you pioneered: use of stratification and matching techniques, and corresponding methods of analysis for controlling for confounding. There is also a proof that I am continually amazed at that was given in your *Biometrics* paper [16], in which you showed that if one had a prospective logistic risk model that would apply to the cohort or forward study, you could estimate that same model, apart from the in-tercept, from case–control data. I wonder how you developed these results on "synthetic retrospective studies."

Mantel: I had been having discussions on this with Max Halperin. But mainly the person involved was Tavia Gordon, who tried to use maximum likelihood procedures for a logistic model. But this was an iterative procedure, and it broke down with a large amount of data, because each cycle of iteration would require heavy computation. Thus I came up with the idea of a synthetic retrospective study to cut down on the computational burden.

Gail: The approach you took is to consider what a case–control study really is; it is the cases and a random sample of the noncases, ideally from a well-defined cohort. You developed this theory based on a deep understanding of what it was to be a case and a sampled control in the case–control subsample. And given the logistic risk model, and the sampling scheme, you showed that one could fit a prospective risk model, which is what we now do routinely.

Mantel: I also had another paper on the odds ratios of a 2×2 contingency table [22]. The different approaches to estimation of the odds ratios were based on whether you had nothing fixed, or the grand total fixed, or one set of marginals fixed, or both sets of marginals fixed. And I showed the different circumstances under which you would get the correct odds ratios. For a single 2×2 table, we showed that if you used AD over BC, then, however the study came about, you would get a proper estimate of the odds ratio. Well, many years later, Max Halperin brought to my attention a very complicated proof of this in the *Annals of Mathematical Statistics*. And Halperin expressed amazement that I had proved the very same thing very simply—just like rolling off a log!

Gail: I would like to get back to the use of covariate adjustments. You have expressed concern about the unthinking use of such adjustments. In a paper that you published in a National Cancer Institute monograph [18] in 1985, you write, "Blind faith in the suitability of mathematical models for analyzing data seems to be the rule, with investigators little aware of the strong assumptions implicit in the use of these models, and the possibility of obtaining misleading results if the assumptions are violated. My prescription would be to keep the covariates limited in number so as to include variables either known or suspected with good reason to play a role in the disease." I wonder if you have seen examples of people going astray with covariate adjustment.

Mantel: One example had to do with the study of factors relating to APGAR scores in newborn

children. The researchers seemed to find nothing significant. The reason was that whatever they tested, they adjusted for all the remaining variables. If those remaining variables either explained the variable of interest or explained the dependent variable, the current variable of interest would seem to be insignificant.

Gail: I have the impression that you prefer stratification to model-based adjustment on covariates, if you have a choice. Is that correct?

Mantel: Well if you use a model, then your results can be affected by the model that you use. It would usually be a strong assumption that the particular model chosen applied. The actual unknown models could be very complicated. So, the only way that you can be sure that you are adjusting properly is to stratify. But if you only have a limited amount of data, the sensible thing to do is to use a model. If you have a lot of data, don't use a model, use stratification. Try to stratify or adjust using covariate analysis on factors that are not independent of the exposure that you are investigating.

Gail: I was struck by some quotes that you gave in your paper on "personal perspectives" [17] about your role in helping the collaborator to produce a manuscript. And, one thing you wrote was, "I'm a packager. The end product of research is a manuscript, and so I try to visualize what kind of paper the investigator's work can be packaged into." And, another quote from a different paper [15] describes one aspect of how you like to package things. It says "For myself, I prefer to have experimental results recorded so there is a minimum obstruction between the reader and the data. It is so arranged as to permit rather full impact of the force of the data. For this purpose, the usual statistical significance tests can sometimes be just so much window dressing, and are frequently not so helpful as simple descriptive statistics derived from the data."

When you start talking to a client, especially a researcher, when do you start to think of a manuscript and how you are going to be presenting the results?

Mantel: Well, I don't think of a manuscript at all. I think about the data and how the data might be shown. Then later on, a manuscript may come about. But, I don't like the use of derivative data; I want to be as close as possible to the original data.

Gail: You mentioned earlier that there were many times, especially when you first came to NIH, where you simply did analyses and were not made an author on the resulting papers. Was there ever an issue at NIH of the professional status of the statistician in his relationship to other scientists at NIH, and when was it that statisticians began participating as coauthors on papers and as scientific peers?

Mantel: Initially we just helped out; then gradually we would learn that we could be of more help. Then it was up to the investigator if he wanted to acknowledge us. Eventually it tended to become more the rule than the exception. So actually, my first publication with an investigator was not until some years after I had been at NIH.

Gail: Did the statisticians talk about this sort of thing, or, rather, was it just a natural evolution of collaboration?

Mantel: Yes, a natural evolution.

Gail: Could you give us a few pointers on what Sam Greenhouse, in complimenting you [7], called the "art" of statistics? Are there a few tips that you would like to give us regarding your approach to data analysis, maybe things to avoid, maybe how you approached a problem? I know this is a very general question, and it is hard to teach artistry, but maybe you have a few thoughts for us.

Mantel: It was only that I just saw the things simply; I never saw complications. I always saw simplicity and so would tend to just see a way out. I think Max Halperin used to say that my success lay in my just seeing or saying the obvious. I never had a complicated thought.

Sometimes when I found that I wanted to do the obvious thing, it would turn out to be completely wrong because the investigator had not told me how the data arose. He would not tell me up front that these data had come from a certain kind of medical practice. And the kind of medical practice that it came from affected the result. And the reason that we found a very strong association had to do with where the data came from. But this was not told to me, and I came to realize later why the analysis that I would have come up with would have been completely wrong.

Gail: Are there any other topics that you would like to discuss today, thinking about NIH and your experiences here?

Mantel: Well, we have said nothing at all about the work on disease clustering. It was Fred Ederer who got me into this area, resulting in the Ederer–Myers–Mantel procedure [3]. Fred had the problem, and he came up with certain approaches. And each time, he would come up with the way he thought it should be handled. And I would say no, this is bad because of this, because of that and so on. And then I came up with this idea of looking at empirical clusters and summing them up. It was in this connection that I became aware of the clustering procedure that Knox [9] had evolved.

Gail: You are talking about how to detect clustering in terms of space and time, rather than just a clustering of geographical area or a cluster in time. Is that right?

Mantel: Yes. So there are two space–time clustering techniques with which I am associated: one is the Ederer–Myers–Mantel procedure, and the other one, which is related to the Knox procedure, takes all possible pairs of observations and sees how close they are together in time and how close they are together in space and then shows how closeness in space is related to closeness in time. So these are two distinct things that I got involved with, but, again, I did not generate this myself; it is because Fred asked me to help out.

Gail: I seem to recall that you defined a measure of closeness in space, a measure of closeness in time, and then took the product of the closeness in space and time, and summed over all possible pairs; and then developed the permutational moments of the sum of these products. Were you able to apply this to a particular disease?

Mantel: Yes, there were some diseases for which we were able to extend the use of the Ederer–Myers–Mantel procedure.

Gail: You also did a lot of work in familial aggregation and some work in genetics too, with C. C. Li at the University of Pittsburgh.

Mantel: That work actually evolved indirectly from another problem. I was working with people in tropical diseases. I think that the senior person was Dr. J. Bozicevich [23]. Their work had to do with the detection of amebiasis by stool examination. Well, if you examine a person's stool, it might be negative because there was no infection, or it could be negative because by chance you didn't happen to see any of the amoebas. We were interested in what proportion of the population had the condition (prevalence) and the efficiency of this test in detecting the condition. Each individual might have been examined a number of times, some of them twice, some three times, some four times. Then the question was how could you use this data to come out with estimates both of efficiency and of prevalence?

Well the notion that I enunciated at that time was the A over B principle [11]. For each examination you would ask (independent of the results of this examination), was the individual known to have the disease? If the answer to that was yes, you would score 1 in the denominator. Then according to whether the result of this particular examination was positive or negative you would score 1 or 0 in the numerator. You would then sum those up over all individuals.

Gail: So this is a way of estimating overall sensitivity of the test.

Mantel: Or, what we called the efficiency. Then I saw where this idea could be expanded to cover the case of batteries of tests, and it could be used to resolve fairly complicated situations. Also, for diagnostic tests, a concern would be whether or not the conduct of a particular test depended on what had gone on before. Because if that were the case, then you could not use the results of subsequent tests to provide confirmation. Thus, you needed careful judgment in using this procedure.

Well, it turned out that this idea would come into play in other situations such as genetics problems. I applied the diagnostic test idea from 15 years before to my work with C. C. Li on estimating the segregation ratio [10].

Steve Fienberg sent me a manuscript which got published in *Biometrika* [4] in which he considered a related situation involving observations in a multidimensional contingency table. You would know which animals were trapped on which occasions but not which animals were never trapped. The issue was to estimate the population size. Fienberg's idea was to assume that you were in a complex contingency table with one missing cell entry. However, his method assumed that all animals came from this same complex contingency table. That was one thing that I would have disagreed with. Maybe the risks were not really homogeneous from animal to animal. Anyway, this simple idea, which was the basis for my paper on evaluation of diagnostic tests, is related to the trapping problem and to genetics problems.

Gail: There are a couple of other topics I would like to ask you about. One is the general area of contingency table analysis, apart from your methods for combining information from multiple tables based on the Mantel–Haenszel procedure. You helped develop the polychotomous logistic regression model. Was there any particular problem that motivated this?

Mantel: It was Marvin Kastenbaum who had gotten me into this. Kastenbaum and Lamphiear had published a paper on modeling a $2 \times 3 \times 5$ contingency table [8]. I saw what he had done, and I recognized that this was the fitting of a logistic. The $2 \times 3 \times 5$ contingency table lent itself to logistic analysis because one factor was at two levels. I saw how I could adapt this idea to fit his data and come out with the same results he had. I sent my results to Kastenbaum, who was amazed that I had gotten those results so expeditiously, since he had had a battery of clerks working on it over an extensive period of time to come up with the re-

sults of his fitting. And here I seemed to get it just like that.

Then he put to me another question. Here I was taking advantage of the fact that this was a $2 \times 3 \times 5$ table and one of the axes had only two possibilities for it. But then he threw a problem at me, "What would you have done had you been in a more complex situation?"

Gail: Where all the variables had multiple levels?

Mantel: Yes. I could not take advantage of the fact that one of the variables had only two levels. One day I got this idea. I reviewed it with Greenhouse and with Cornfield and they both said, "Gee, this is a novel idea!"

Gail: And that became your 1966 paper in *Biometrics* [12]: "Models for complex contingency tables and polychotomous dosage response curves."

There are some other unifying works that you did in contingency table analyses having to do with relationships between marginal homogeneity, symmetry and the Mantel–Haenszel procedure.

Mantel: Yes. Dave Byar and I were having lunch one time and I could see immediately how the Mantel–Haenszel procedure could be adapted to cover a problem Dave had raised. But it was a lot harder than that to finally get it through to publication.

Gail: Yes, I recall it came out in *Communications and Statistics* [20]. I think it was Darroch who wrote on marginal homogeneity later and thought that this paper really tied together a number of important concepts.

Mantel: Yes, and other work that simplified Ed Gehan's method for the extension of the Wilcoxon procedure to allow for truncated observations also received some very complimentary remarks [14]. Gehan allowed for left-truncated or right-truncated observations or both [5, 6]. Each of Gehan's papers had about a hundred formulas. But in my work it was all reduced to a single formula one inch long.

Gail: It seems that in the late 60s you got very interested in analysis of censored data. Did you realize that your 1966 paper [13] in which you developed the log rank test would become (after Cox's paper in 1972 [2]) one of the most widely used procedures in survival analysis?

Mantel: Not at the time. It wasn't until a meeting of the International Statistical Institute, here in Washington, that Peter Armitage brought out the relationship between the Cox and Mantel methods.

Gail: In the special case of a two-sample problem, the score statistic that Cox recommended would have been, essentially, equivalent to yours. But, of course, Cox's came six years later.

Well, I know there are a lot of other topics we could be discussing today.

Mantel: I think we should have a follow-up session in a year from now, and you can generate additional questions.

Gail: Wonderful. Thank you very much.

REFERENCES

[1] CORNFIELD, J. and MANTEL, N. (1951). Some comments on "Estimates of LD50: a critique." *Biometrics* **7** 295–298.

[2] COX, D. R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. Ser. B* **34** 187–220.

[3] EDERER, F., MYERS, M. H. and MANTEL, N. (1964). A statistical problem in space and time: do leukemia cases come in clusters? *Biometrics* **20** 626–638.

[4] FIENBERG, S. E. (1972). The multiple recapture census for closed populations and incomplete 2^k contingency tables. *Biometrika* **59** 591–603.

[5] GEHAN, E. A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* **52** 203–223.

[6] GEHAN, E. A. (1965). A generalized two-sample Wilcoxon test for doubly-censored data. *Biometrika* **52** 650–652.

[7] GREENHOUSE, S. W. (1997). Some reflections on the beginnings and development of statistics in "Your father's NIH." *Statist. Sci.* **12** 81–85.

[8] KASTENBAUM, M. A. and LAMPHIEAR, D. E. (1959). Calculation of chi-square to test the no three-factor interaction hypothesis. *Biometrics* **15** 107–115.

[9] KNOX, E. G. (1964). Epidemiology of childhood leukemia in Northumberland and Durham. *British Journal of Preventative and Social Medicine* **18** 17–24.

[10] LI, C. C. and MANTEL, N. (1968). A simple method of estimating the segregation ratio under complete ascertainment. *American Journal of Human Genetics* **20** 61–81.

[11] MANTEL, N. (1951). Evaluation of a class of diagnostic tests. *Biometrics* **7** 240–246.

[12] MANTEL, N. (1966). Models for complex contingency tables and polychotomous dosage response curves. *Biometrics* **22** 83–95.

[13] MANTEL, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports* **50** 163–170.

[14] MANTEL, N. (1967). Ranking procedures for arbitrarily restricted observations. *Biometrics* **23** 65–78.

[15] MANTEL, N. (1969). Some statistical viewpoints in the study of carcinogenesis. *Progress in Experimental Tumor Research* **11** 431–443.

[16] MANTEL, N. (1973). Synthetic retrospective studies and related topics. *Biometrics* **29** 479–486.

[17] MANTEL, N. (1976). A personal perspective on statistical techniques for quasi-experiments. In *On the History of Statistics and Probability* (D. B. Owen, ed.) 103–129. Dekker, New York.

[18] MANTEL, N. (1985). Avoidance of bias in cohort studies. In *Selection, Follow-up and Analysis in Prospective Studies. National Cancer Institute Monograph* **67** 169–172. U.S. Government Printing Office, Washington, D.C.

[19] MANTEL, N. and BRYAN, W. R. (1961). "Safety" testing of carcinogenic agents. *Journal of the National Cancer Institute* **27** 455–470.

[20] MANTEL, N. and BYAR, D. P. (1978). Marginal homogeneity, symmetry, and independence. *Comm. Statist. A—Theory Methods* **7** 953–976.

[21] MANTEL, N. and HAENSZEL, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* **22** 719–748.

[22] MANTEL, N. and HANKEY, B. F. (1975). The odds ratios of a 2×2 contingency table. *Amer. Statist.* **29** 143–145.

[23] TOBIE, J. E., REARDON, L. V., BOZICEVICH, J., SHIH, B. C., MANTEL, N. and THOMAS, E. H. (1951). The efficiency of the zinc sulfate technique in the detection of intestinal protozoa by successive stool examinations. *American Journal of Tropical Medicine* **31** 552–560.