

Interview with John Karon, '63

1. Can you tell us what sticks in your head about your time at Carleton? Are there people --- faculty, staff, students --- of whom you have particularly fond, or vivid, memories? Are there courses or other experiences that define Carleton for you? What was your comps experience?

I entered Carleton in the fall of 1959. Ken May was math department chair. Faculty offices and almost all classes were in Goodsell. Ken shaved his head; math was known as “The Department of the Three Domes”. At that time, no (or almost no) high school taught calculus. Fall semester we took precalculus (from a book Ken had written?). There were well over 100 in the class; Ken taught it in Skinner Chapel, using either blackboards or overheads. Ken also taught my calculus 3 class and then differential equations second semester of my sophomore year. That class was TuTh (and maybe Saturday) at 0800, if I recall correctly (yes, Saturday morning). On one of those -30° January or February mornings, Ken walked in, stripped off his parka, and said, “When I went out in my pajamas to start the jeep this morning, I nearly froze!”

I took most of the math classes that were offered, including number theory from John Dyer-Bennett and linear algebra, also I think from JD-B. Roger Kirchner joined the faculty from graduate school my senior year and taught the complex variables course from a text written by his dissertation advisor, the famous complex analyst Lars Ahlfors. I think we used the same text in graduate school; Roger’s course was challenging!

Carleton got its first computer, an IBM 610, my junior year. You stored your program on punched paper tape; it had 80 internal storage registers and a plug-board with the basic mathematical operators (and I think logical comparisons) that you could wire to write a subroutine. Printed output was on an electric typewriter. I wrote a subroutine to compute one or several trig functions from the infinite series expansion; debugging the code was not easy! My senior year we got an IBM 1620, which used punch cards, had a reasonable (for the time) number of internal storage locations, and line printer.

Steve Stigler is undoubtedly the best-known of the math majors from my year. Gary Kampen was probably regarded as the most able student. As I recall, he went to Michigan hoping to work with a number theorist. That didn’t work out. I believe he left with an M.S. and became a computer programmer.

Comps were a three- or six-hour written exam covering the basic math courses, with an one-hour oral exam covering anything the faculty wished to ask. I did the oral exam wearing a madras plaid sport coat. Ken May had decided not to cover the evaluation of definite integrals by substitutions in calculus 2 or 3, so I had to learn that on my own during spring break of my senior year in order to prepare for comps.

I was fortunate to win an NSF fellowship for graduate school. Luckily for me, no research proposal was required (competition was based on GREs, faculty recommendations, and I presume a transcript), as I would not have had a research proposal. I told Ken May that I was interested in applied math; he suggested applying to Stanford and NYU; I also applied to

Wisconsin as a “safety”. I chose Stanford.

2. Can you tell us about your post-Carleton education? I know you have a PhD from Stanford in mathematics, what was your specialty? Who was your thesis advisor? Did Carleton prepare you for that experience? Was Polya in the Stanford math department in those days? How about Szego, Bergman, or Royden? Any particular memories of being in that department full of world-class mathematicians?

Stanford had picked up 5 world class German mathematicians: George Polya, Gabor Szego (the coauthors of the “problems” book), Stefan Bergman, Max Schiffer (analysis, mathematical physics), and Charles Loewner. If I remember correctly, Schiffer and Loewner (independently) solved the Bierberbach conjecture for different values of n (3 and 4?). Polya and Szego were retired, but Szego had an office in the math department. The faculty emphasized analysis (few faculty in algebra or logic).

I believe that I was one of at least 50 first-year grad students in 1963. Apparently the department did not have the reputation required to attract enough students who were likely to complete a PhD to be more selective. Some students (especially from Ivy League schools) had already had the equivalent of first-year grad courses. I was not as well prepared. Ralph Phillips taught the analysis course, rather directly from Royden’s text (Royden was on the faculty). I got about half A’s and half B’s in the basic graduate courses (not a sterling performance).

Qualifying exams were 6 hours of real variables, 6 hours of complex variables, and 4 hours of your choice between applied analysis and algebra. I chose the former. Previous years had resulted in about 1/3 pass, 1/3 take again, and 1/3 fail. I was not very confident but passed on my first attempt.

I wrote a dissertation in analysis under Samuel Karlin. He was a very hard worker (“I take one evening off per week to talk to my wife”) with a good reputation in analysis (at the time, specifically total positivity and splines) and stochastic processes, and I think he had had a joint appointment in statistics. I also took courses in statistics. The analysis of variance course required that we do the calculations on Monroe calculators. I took a theory course from Brad Efron, who had just joined the faculty; he is now one of the most distinguished American statisticians doing work that can be applied to data analysis (developer of the bootstrap, in particular).

3. After a few years of teaching and research in mathematics at Syracuse and Colorado College you took a post-doctoral position in biostatistics at North Carolina. Biostatistics must have, at that time, just been coming into existence as a discipline (is this right?); what possessed you to pursue it? Special foresight, intrinsic interest, or happenstance?

I went to Syracuse as a result of there being a fine approximation theorist on the faculty (Lorenz). He left for Texas before I got there. After 3 semesters, I went back to Stanford as a sort of post-doc for a year and a half, then to Colorado College (rejecting offers from Carleton and St. Olaf; I felt that Colorado had a younger and more vibrant faculty; Roger Kirchner was very dismayed).

I was denied tenure (too demanding a teacher; I was probably doing more research – splines and mathematical modeling – than any other math faculty member). I had been interested in mathematical biology, hence applied for and was selected for a 3-year NIH-funded post-doc at UNC-Chapel Hill.

Biostatistics was a well-established discipline. The department at UNC-CH was in the School of Public Health and was regarded as one of the better departments in the country. According to its website, it was founded in 1949. The statistics department also had a distinguished faculty. I audited many of the basic graduate courses, including a few theory courses, learned to code in SAS (which was young at the time), and worked on a theoretical evaluation of the analysis of epidemiologic data with one of the senior faculty members.

4. Tell us about your work on cardiac disease prevention. What did the study you were involved in uncover? What mathematical and statistical tools did you use? How did you learn the relevant biology and medicine?

After completing the post-doc, I chose to stay on as a research-track faculty member in the clinical trials unit. Their project at that time was a heart-disease prevention trial. They were the data management and statistical coordinating center for a twelve-center epidemiologic prevalence study and a seven-year, ten-center study of the effect of lowering LDL cholesterol in men, with 3800 participants who took either cholestyramine (described to me as resembling sand, mixed into orange juice or something similar) or an equally foul placebo. I joined the staff two or three years before the end of the study.

I was responsible for laboratory quality control and the analysis of exercise ECG data. Lipid analyses were done at the individual centers, with supervision and reference materials from the Centers for Disease Control (CDC). Other clinical chemistries were done at a commercial lab in southern California. We used standard quality control ideas to evaluate the quality of the laboratory analysis. I supervised lower level staff who did the analyses and drafted the reports.

As I recall, no one had looked at the exercise ECG data before I joined the staff. Statisticians did not have access to the data files; we specified analyses for the programming staff, who used SAS or perhaps FORTRAN. I used standard statistical methods, including survival analysis (a positive test was a secondary endpoint for the trial).

We had excellent epidemiologists on the staff; they supplied the necessary medical knowledge. We had weekly working seminars at which staff presented summaries of what they were doing. I did an exercise ECG to understand how the data were collected and went to the ECG reading center to understand how the data were coded.

We were delighted to learn at the completion of the trial that the incidence of cardiac disease (fatal and non-fatal myocardial infarction) was significantly lower in the cholesterol-lowering group than in the placebo group based on a one-sided hypothesis test (this was controversial). The results were published as a lead article in JAMA in 1984. The treatment group had a 19% reduction in risk compared to the placebo group. Secondary endpoints (including a positive ECG exercise test) showed similar reductions.

5. After the cardiac disease work you worked for fourteen years on HIV/AIDS forecasting and modeling. Was this for the CDC? Please tell us about that work. Many of us remember that the US government was in denial about HIV/AIDS for much of the 1980s (cf. *And the Band Played On*). Did politics intrude on your work?

As a result of my contacts with laboratory staff at CDC related to laboratory quality control, I was recruited to work in a laboratory group there in 1984. After one year, I moved to a group doing a study of the effects of exposure to Agent Orange in Vietnam enlisted army veterans. I specified and did statistical analyses, most using basic statistical methods, and supervised the statistical staff. I coordinated a study that developed an exposure index (based on military records) in these men. The laboratory staff developed a very sophisticated assay for dioxin based on serum with a sensitivity of a few parts per trillion. We found no association between our exposure index and the dioxin probability distribution, and indeed the distributions were the same in the Vietnam veterans as in those who had not been in Vietnam. As a result, there was no exposure study.

I then moved (in early 1987) to the HIV/AIDS group as the senior statistician there. CDC had identified what came to be known as AIDS in 1981. Quite soon AIDS became a reportable disease in every state; CDC collected the surveillance data (which included demographic information, exposure risk, and death). The test for HIV was developed in 1985; some states began reporting persons with a positive HIV test then.

I was responsible (with the help of able statistical staff) for the analysis of the AIDS and HIV surveillance data, including forecasting future AIDS cases and deaths. Initially we used regression extrapolation (this worked because distribution of the time from infection to AIDS is long and variable). Another method (back-calculation) which estimated HIV incidence was developed by first-rate statisticians at NIH and Johns Hopkins; this requires solving an integral equation of the first kind (this is an ill-posed problem). I coordinated the synthesis of results from a variety of methods, as well as doing evaluations of data to check the plausibility of estimates. When the first anti-retroviral drug (AZT) was introduced about 1989, the distribution of the time from infection to AIDS changed (increased). In addition, CDC broadened the case definition so that most of those infected could be classified as AIDS earlier in the course of infection. This required modification of our methods; I did some of the work that resulted in quite accurate short-term forecasts.

So far as I know, there was no political effect on our statistical work at CDC. We did what needed to be done, with no expectation of particular results. My recollection is that CDC was not doing very much prevention work (compared to now) in the late 1980s and early 1990s. I enjoyed reading *And The Band Played On*. I worked with many of the key CDC scientists who appear in the book.

6. Could you tell us some details of that work? What do the epidemiological models look like? What did you learn about AIDS prevention and transmission? Do you keep up with that? What, if anything, should the world public health community be doing about AIDS in Africa that we are not doing?

CDC uses statistical models, rather than mathematical models, to analyze surveillance, epidemiologic, and behavioral data. Some sophisticated simulation transmissions models have been developed for HIV. CDC funded some of that work (by Herb Hethcote, a mathematical modeler at Iowa). These models generally require many parameters, both demographic and behavioral parameters. The latter are very hard to estimate and change over time; past transmission depends on past values, and even current values are hard to estimate. Some of Hethcote's models were remarkably unsuccessful. Thus, CDC has not participated in the transmission modeling industry. Sophisticated models, with sophisticated graphics, have been developed for Africa, and I collaborated in supplying data for some of that effort.

One of CDC's recent important efforts has been the estimation of HIV incidence (numbers of infections per year). Usually such an estimate requires the dates of at least two tests, a negative and a positive test. Negative HIV tests are not reported, and a substantial proportion of those infected never had a negative test. The CDC lab developed an assay that is practical in the United States that requires only one blood draw (it uses a test that becomes positive a mean time of six months after infection; the sensitive test used to determine infection becomes positive within a few weeks). One of "my" statisticians estimated the probability distribution of the time from infection based on the sensitive test to a positive result on the new test. I realized that we could use this to estimate the number infected in a manner similar to estimating the size of a population from sample survey data. Luckily, "I" had hired a very able younger statistician. We collaborated with an excellent statistician from Johns Hopkins and an operation research faculty member from Yale to develop the necessary theory and evaluate the performance of our estimator, using simulation. The results were published in *Statistics in Medicine* last year (I am first author). CDC developed the necessary procedures to collect the samples and do the alternative test. This work was the basis for the new estimates of HIV incidence published in *JAMA* last year, using our statistical procedure.

I also was the primary statistician for clinical trials of preventing mother-to-child HIV transmission in Thailand and in Cote d'Ivoire, using a very short course of AZT to the mother shortly before birth (efficacy of a long course had been shown in the US). I did data analyses and worked with the monitoring committee. We were delighted when the Thai trial showed a reduction in transmission to about 10% in the AZT group compared to about 25% in the placebo group. The Cote d'Ivoire trial was still in progress, showed somewhat similar results (Thai women are incredibly compliant with instructions), and was stopped.

7. What do you do for fun? Hobbies? Family? (Where, by the way, do you live?)

I "retired" from CDC in the spring of 2002 and we moved to Albuquerque. I continued to work for CDC as contractor, particularly on the incidence estimation project, but also on other analyses. In particular, I was the statistician for a Phase II trial of a vaginal microbicide done in Thailand, designed to prevent HIV (there were no infections in any group; we evaluated safety and acceptability). This trial was done jointly by CDC and The Population Council, a New York City non-profit that has a long history of contraceptive research. I am continuing to work with them, primarily on contraceptive research trials of devices suitable for use in less-developed areas. I have also been working on the evaluation of statistical methods for analyzing

cervicovaginal HIV data, for which a large proportion of the results are below the quantitation limit.

My wife, Kate Killebrew (a somewhat remote relative of Harmon, but without his athletic ability) is a radiologist. She works about one week per month at an Indian Health Service hospital on the Navajo reservation in the northwest corner of the state. We have two grown daughters. Our older daughter has a DVM and MPH from Madison, completed a 2-year CDC epidemiologic training program, and now works for the Wisconsin Department of Health (she has primary responsibility for H1N1). Our younger daughter completed an MFA at the Iowa Writer's Workshop and has now started the Masters program in journalism in Madison. We have granddogs and grandcats but no grandchildren.

I have no trouble keeping busy. We have dogs; both are registered therapy dogs, and "we" go to rehab hospitals (besides long daily walks). I grow vegetables and have about fifteen fruit trees. I provide advice to the gardening public and give talks on seeds and plants to third graders as a master gardener. I participate in local archaeology trips and am part of a rock art recording group (with weekly recording work). I play over-the-shoulder B-flat cornet in a Civil War re-enactment style band (yes, the bell points backwards). I do local hikes, an annual five or six-day Sierra backpack trip (as much off trail as possible), and cross-country ski in winter. I was a runner at Carleton and continued until the spring of 2003 (including fifteen marathons), when hip arthritis forced me to quit; after a replacement, I have resumed easy jogging. New Mexico has fascinating history; I read what I can and attend lectures. Albuquerque has a remarkably active jazz scene, and we enjoy attending concerts. No statistical work related to archaeology yet, but I am threatened with being asked to develop a method to compute confidence intervals from three-dimensional data that estimate the direction of the magnetic north pole at a specified date; the idea is to use directional data from a new sample to estimate the date of a site.

8. Do you have any parting words of advice or wisdom for current Carl math majors or other alums?

Current majors should know that there are lots of ways to use mathematical ability besides being a "traditional" mathematician. Statistics is a very rewarding field. As one talented academic statistician has said, it is a wonderful experience to be able to play in other people's sandboxes. The ability to think and reason quantitatively is invaluable.

It is not necessary to do excellent (or even good) research in order to make important contributions. My statistical research contributions are very modest, but I was very pleasantly surprised at how well I was regarded at both UNC-CH and at CDC. I believe my strengths are that I can think clearly (thinking about assumptions, potential biases, and data limitations are really important for a statistician), work well with other people, write clearly, give clear and well organized talks (that was a learning experience), develop complete analyses for presented data, and manage staff well. I have kept up reasonably well with the development of new statistical methods and think about how these methods can be applied to problems. These are important skills, but you don't necessarily learn them in your academic work.

It may be better to be an “average” student in a graduate program with an excellent reputation than a “star” student in a graduate program with a lesser reputation. I was hardly one of the best PhD students at Stanford, but its reputation opened some doors for me. I still had to earn admission on my own merits, of course.

Learn from working in an organization. Observe how good managers and technical staff approach and solve problems.

For some important decisions in life (my choice between Colorado College and Carleton was such a decision), you can never obtain all the information you need to be sure that you made the optimal decision. Made a decision and go on; don't worry about the result (unless you can learn something about making a future decision). New opportunities will arise from your current choice. Had I accepted the Carleton offer, I might well have obtained tenure; I would not have become a statistician; working with talented young students would be a kick, but I likely made a greater contribution to society as a result of my actual career path.

If you are a manager, hire the best people you can, provided that they will contribute properly. A person with a difficult personality may be worth hiring, provided that the benefits are strong enough. Hire people who have better technical ability in an area than you do. Assemble a team with a variety of skills; don't expect one person to do everything (a great strength of my statistical group at CDC was that skills and interests were complementary). Give staff both positive and constructive feedback. If there is a personnel problem, solve it early, before it can become so difficult that it is hard to solve.

Know your limitations. Recognize when you need a solution or advice and know where to find someone who can help. This is really important! As you approach the end of your career, recognize when you need to let others take over.

Most of us need to develop other interests in recognition that our professional careers will end. Really talented scientists are even busier after they “retire” as a result of requests to be consultants or serve on advisory committees. Many retirees have said or told me that they are even busier after retiring. Plan ahead! There are opportunities to help nonprofit organizations. You will be surprised at how many board members can't or don't think quantitatively and can't use a spreadsheet. Basic knowledge can go a long way!