Published in final edited form as:

J Biopharm Stat. 2016 ; 26(6): 1007–1019. doi:10.1080/10543406.2016.1226331.

A conversation with Dr. Greg Campbell

Aiyi Liu^a and Lilly Q. Yue^b

^aBiostatistics and Bioinformatics Branch, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland, USA;

^bDivision of Biostatistics, Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, Maryland, USA

1. Introduction

Dr. Gregory Campbell received his PhD in Mathematical Statistics from the Florida State University in 1976. After his completion of graduate studies, Dr. Campbell took a faculty position as an assistant professor in the Department of Statistics at Purdue University, and later joined the intramural program of the National Institutes of Health, where he was promoted to a senior investigator with tenure, and he served as the Acting Chief for the Laboratory of Statistical and Mathematical Methodology, and then Chief of the Analytical Biometrics Section.

Dr. Campbell joined the Food and Drug Administration (FDA) in 1995, and served as the Director of the Division of Biostatistics in the Office of Surveillance and Biometrics of the Center for Devices and Radiological Health (CDRH) at the FDA until his retirement from the federal government at the end of June 2015. His extraordinary leadership has resulted in the growth of the Division from 18 mathematical statisticians to 62 in five branches. He has made exceptional contributions to public health, the FDA, and medical device industry by leading the Division to provide statistical support to CDRH as a whole and, in particular, the statistical reviews of FDA's pre-market device submissions, bringing in numerous important and innovative medical devices to the market place.

During his nearly 33 years of federal service, Dr. Campbell has sustained high levels of professionalism and dedication. Because of his distinguished contributions as a driving force and a great leader in statistical innovations in the medical device world, the course of medical device evaluation and regulatory science has been greatly shaped. His pioneer work and sustained efforts in adapting and advancing the implementation in a regulatory environment of Bayesian statistics, causal inference, adaptive design, missing data, and diagnostic test methodology have made phenomenal, far reaching, and long-lasting impacts to public health and resulted in the increase of the integrity of clinical study design and the creditability of study results as well as the consistency, transparency, effectiveness, and

predictability of regulatory decision-making for the FDA and industry. Dr. Campbell has been recognized as an internationally renowned statistical leader in regulatory science.

Dr. Campbell served in leadership positions for the Eastern North American Region (ENAR) of the International Biometric Society, the Society for Clinical Trials, and the FDA Statistical Association (FDASA). He played a distinguished leadership role in the establishment of the statistical community of medical devices and diagnostics, and was instrumental in the formation of the American Statistical Association (ASA)'s Section on Medical Devices and Diagnostics (MDD).

Dr. Campbell was a member for over 10 years of the prestigious Senior Biomedical Research Service (SBRS) in the Department of Health and Human Services, and also served for 15 years on the SBRS credentialing committee. He has been a Fellow of the American Statistical Association since 1998, and recently become a Fellow of the Society for Clinical Trials. He has received numerous awards, including the FDA Distinguished Career Service Award, the FDA Commendable Service Award and the Merit and Outstanding Service Award, the CDRH Diversity Award and the Outstanding Scientific Award for Excellence in Analytical Science, the MDD's first Distinguished Achievement Award, and the FDASA Award for long-term support and significant contributions.

The interview presented below was conducted by Lilly Yue of FDA and Aiyi Liu of NIH on June 25, 2015, in Lilly Yue's office in Silver Spring, Maryland.

2. Education and early career in statistics

Aiyi and Lilly What led you to become a statistician?

Dr. Campbell When I was in high school, I gave no thought of a career as a statistician. In high school I was interested in mathematics. I had some very good teachers in high school, and took Advanced Placement (AP) calculus which was very new at that time. I did pretty well so I got college credit for the AP classes in calculus. However when I attended the University of Dayton as an undergraduate studying mathematics, I also took some probability and statistics courses. There were a number of good instructors there and I learned statistics from a book like the one by Hogg and Craig, a good mathematical statistics course for undergraduates. But even then I wasn't sure I wanted to be a statistician and decided instead to go to graduate school in mathematics. So I went to the Michigan State University which had an NSF traineeship for me so I could take a lot of courses and had no teaching or research duties. I was there for two years and took lots of mathematics courses and also a fair number of statistics courses. There was a separate Department of Statistics and Probability at Michigan State. I was in the PhD program at Michigan State so I took the qualifiers in mathematical analysis and I passed them; not a lot of people passed, certainly not on the first attempt, but I did. So I thought I would probably want to write a thesis in analysis. So I started to think about it but then realized that the analysis topics really were not ones that I thought would be interesting as a career. So I started to talk to the people in the statistics department there. I had taken an applied nonparametric course and thought that area was pretty interesting. So I could've stayed at Michigan State and transferred over to

the Statistics and Probability Department but I decided that I would look around and apply to other places. I ended up applying to Florida State University because that was a place where there were a fair number of nonparametric statisticians. Ralph Bradley, who is famous for the Bradley–Terry model, had done great work in nonparametric statistics. Myles Hollander was there as well and I knew he had written an excellent book on applied nonparametric statistics, coauthored with Doug Wolf. Also there was J. Sethuraman, a strong theoretician.

So I went to Florida State and that's how I really developed a strong interest in statistics. I had already all the mathematics I needed from MSU and so at Florida State I took a lot of statistics courses. In addition the department had this idea that some people would be interested in learning another topic besides statistics. I was interested in demography which is the study of populations, so I took a number of courses in demography and eventually wrote a paper on the stable equivalent population with Professor Tom Espenshade. But what I really did was to finish the coursework in statistics, pass the department qualifying exams, and then write a doctoral dissertation under Myles Hollander on rank-order estimation for Dirichlet process priors. These priors were something Thomas Ferguson from UCLA was famous for, and he actually taught a week-long NSF summer institute workshop at Ohio State. I went there, and Myles Hollander too, for the week, and that was really good. I finished my degree at Florida State. At the end of my time in Tallahassee, I was actually appointed as a lecturer in the department. I was also interested in reliability theory. There was a summer visitor Professor Richard Barlow from Berkeley who had an office right across from mine. I was the grader for the course he taught in the summer on reliability. He's very well-known and had written a book with Frank Proschan who was at Florida State as well and was on my committee. It's interesting to note that Richard Barlow was in the process of becoming a Bayesian. He had written this book with Frank that was a frequentist approach to reliability, but in fact he was slowly becoming a Bayesian.

When I finished my PhD, I thought an academic career would be interesting, so I applied and was appointed an assistant professor in the Department of Statistics at Purdue. There were lots of good statisticians there, from Shanti Gupta, who was the Head of the Department of Statistics, and Herman Rubin, to Maryellen Bock and Jim Berger who were up and coming at that time. I learned a lot and wrote some papers based on my thesis, did some work with Steve Samuels on optimal stopping problems, like the famous secretary problem that is useful in hiring strategies. I also became interested in survival estimation and had the idea to extend Kaplan–Meier to two dimensions, a bivariate survival estimation problem. I worked on that a while and wrote a paper with Antonia Foldes, a visiting fellow from Hungary, on large-sample properties for the estimators we developed for bivariate estimation. That was actually a lot of fun and I enjoyed that a lot. Later I went to a nonparametric conference in Hungary to present the work. I met David Cox there which is kind of cool.

3. Life at NIH

Aiyi and Lilly Many statisticians may not know that you actually worked at NIH for quite a few years. Can you talk about your experience at NIH and how you ended up at FDA?

Dr. Campbell After six years at Purdue my wife, who had come with me to Purdue as a visiting assistant professor in mathematics education, did various things when that temporary position ended but none of them became a permanent faculty position. At one point she was advising undergraduate computer science majors. So we decided to leave Purdue. She applied for a faculty position at the University of Maryland at College Park and she beat out hundreds of people to get the job offer and then it was up to me to see if I could find a job in the area. But I knew the Washington area is probably the densest place in the United States, and maybe in the world, for statisticians, so I looked for various positions. I applied for academic positions and for a position at Johns Hopkins Applied Physics Lab. I got a few offers and decided to accept a Staff Fellow position at NIH, in what was then called the Division of Computer Research and Technology (DCRT), and now is known as the Center for Information Technology (CIT). The group I was in was the Laboratory of Statistical and Mathematical Methodology, and it was headed by James E. Mosimann, who was a very good statistician. He actually developed the Dirichlet-multinomial distribution, the multivariate extension of the beta-binomial. He was very interested in distribution theory which was great. So we went through all the distributions in four volume set by Johnson and Kotz. He was also interested in size and shape so he and I wrote several papers on size and shape and built a model where you model the shape using the size as a covariate. For example, we were interested in human growth models. What happens is when babies are born, they have really large heads compared to the rest of their bodies. Then as they get older, the rest of the bodies catch up. We worked on that problem and developed a model to do that. The lab was an interesting place because it had mathematicians, computer scientists, and statisticians. I was in the section on statistical methodology (and eventually became its chief). There was also a section on statistical software that supported all statistical packages at NIH. So I enjoyed working there for 10 years, and it was interesting because a lot of institutes had their own statisticians, so when researchers in those institutes wanted statistical help they would usually go to the statisticians in their Institute. But some institutes either did not have statisticians or they couldn't find them so researchers would come to us with all kinds of consulting problems, so all statisticians in the lab did some consulting and collaboration work. I wrote a number of papers with researchers on a lot of different topics which was a good experience. But we were expected to do our own research as well. At that time I was starting to get interested in Receiver Operating Characteristic (ROC) analysis. Someone had come in with a problem about breeding horses, and I even didn't know exactly what that had to do with NIH. But the problem is related to studying for performance of diagnostic tests. This was in the early 1990s when ROC was not on anybody's radar, but I was getting quite interested in that and wrote a paper on bootstrapping ROCs and associated confidence regions. There was a symposium in 1993 organized by senior statisticians at NIH and I was invited to give a talk on that, which eventually ended up with a paper in *Statistics in Medicine*.

After a few years Jim Mosimann became very interested in fraud at NIH. There were some highprofile fraud cases which had come to the NIH's attention, and NIH needed to do something about it and some of the issues were very statistical. So he decided to leave the lab so he stepped down as lab director to serve as the statistician in the Office of Scientific Integrity and I became Acting Lab Chief for two years and that was my introduction to

management. The lab probably had about 15 people so I did all the administration for the lab. But I decided that I didn't want to stay there. There was a new director of DCRT who had come before Jim Mosimann had left and I decided at that time that it wasn't for me a good fit. So I looked around and Jonas Ellenberg decided to hire me. He was the Chief of the Biometry and Field Studies Branch at the National Institute of Neurological Disorders and Stroke (NINDS). I was there for about three years. I would like to let you know that Shanti Gupta was a great person to work for, and very supportive, Jim Mosimann was fabulous, and it was a pleasure to work in Jonas's branch as well. Jonas brought me in as the head of the Analytical Biometrics Section (which basically was the statistical analysis of brain images) within the branch. Nick Lange, a statistician from Harvard who was interested in neuroimaging, was in the Analytical Biometrics Section and Paul Albert and Lisa McShane were Staff Fellows in the Branch. Nick and I worked on really interesting problems in neuroimaging. One of the challenges is it is really easy to create these pretty pictures of the brain where you have different colors and you say red is where all the activity is going on but the reality is that's just very descriptive; there's no statistical machinery behind it. So Nick and I worked on all kinds of problems. I worked with a Japanese researcher who was interested in trying to figure out what areas of the brain were associated with the tapping of fingers. So they had functional MRI machine in the basement of the NIH Clinical Center, at that time a really powerful magnet that was four Tesla. At that time that was big, now they have seven Tesla magnets. The researchers did studies with finger tap experiments and then watched what areas of the brain would light up and then wanted to know where the real activity is as opposed to just seeing pretty pictures. That turned out to be really interesting. I did a lot of collaborative work there but also wrote statistical papers. I became interested in Receiver Operating Characteristic (ROC) methodology so I wrote a number of ROC papers with an engineer on what are called fuzzy ROC, and that uses fuzzy membership ideas and is nonparametric. It turned out there are some interesting statistical problems there. I wrote a review paper on ROC with a clinical chemist named Mark Zweig in NIH's Clinical Center; he was the head of the Clinical Chemistry Department there. This ROC review paper we wrote for the journal *Clinical Chemistry* in 1993 at one point was the most cited paper in the journal.

What happened at NINDS later was Jonas Ellenberg left NIH for Westat and eventually went to University of Pennsylvania. There was a change in the Institute Director in NINDS so things were kind of up in the air and it wasn't clear how valued the field of statistics was in the Institute. One of the questions was whether to stay there trying to build up the group again or to leave, and then I heard about this job at FDA and thought that could be kind of interesting. It had diagnostic component to it, something I was interested in. At that time I really didn't know a lot about clinical trials. The collaborative work I have done at NIH was with biomedical researchers who were more in the lab than doing large NIH studies. One of the nice things about NIH is that it is such a great place to work; there are all of these opportunities, interacting with people and going to lectures, which I did on all different kinds of topics. I learned all about gene sequencing since the computer division was interested in software associated with genomics and came close to hiring people who had written programs to do that, and there were some interesting statistical problems there. This was before NIH really made the decision to move in a concerted way toward the genomic

work they are doing today. But it was an exciting time to talk about sequencing the human genome; all these famous people such as Eric Lander would come to NIH and I was on campus so it was really easy for me to go to the lectures and even when I worked in NINDS that was in downtown Bethesda, within a good walk of the NIH campus. So there were a lot of opportunities to interact, and to be exposed to lots of ideas and learn a lot and so I took full advantage of that.

Aiyi and Lilly Is there any advice you might want to give to junior statisticians from your NIH experience?

Dr. Campbell I think it's really important that you work for people that you like and respect. I had that privilege so many times in my career with Shanti Gupta, Jim Mosimann, and Jonas Ellenberg. But when both Jim and Jonas moved on, it wasn't quite as good. In situations where it doesn't seem like things are working out, one can think about leaving. One of the things that I have learned over time is you shouldn't stop working or working hard even if the job might seem a little frustrating that's not the reason to give up.

It is also important to keep looking for places where you know statistics is important. One thing I should say around this time is that at FDA statistics is really important. In places like NIH this very much depends on the collaborators you work with. There are collaborators who show up at your office because they sent this paper away and they got a referees report and they need some statistical help. That's not a true collaboration as you can do better collaboration work when you're with the project from the very beginning and you're working on it together. Coming to FDA I realized that it's not like that here; for the most part everybody understands how important statistics is, many of the decisions that FDA made hinge on the data and how you analyze and interpret that data. That makes statistics much more important.

The other thing is that there is the tradition of pre-specification, that is, companies need to come and tell us exactly what they're going to do when they design study and how they're going to analyze the data. A lot of biomedical research, even good research, is not done that way. In many cases there's a general idea of why the experiment is being done, but exactly how to analyze the data may not be decided upon until the data actually show up. Scientifically that is not ideal. So coming to FDA in that regard was quite refreshing, because for all of the significant risk products, companies had to pre-specify in their protocol exactly how they're going to analyze the data with the data still unseen. So that turned out to be pretty nice.

4. Leadership

Aiyi and Lilly We have witnessed a great deal about your wonderful leadership. Can you share with us your experience in that regard?

Dr. Campbell I have thought a lot about leadership in the past month or two, getting ready for the 2015 Graybill Conference joint with the International Chinese Statistical Association (ICSA) Applied Statistics Symposium in Fort Collins, Colorado. There was a Leadership Forum Panel at the conference and I was on the panel with Janet Wittes and Xiao-Li Meng.

One of the things that I think is the most important about leadership is that you have to have a vision, have to be able to see the future that you want to be able to try to make. Imagine that future and then work toward realizing it and enlist other people to help you with that vision. Leadership is getting others to work toward a common goal, taking people to a place where originally they didn't even think they wanted to go but once they saw the vision then they realize that they wanted to go there too. So when I came to FDA my boss was Dr. Larry Kessler. He came to FDA from NIH in 1994; he was quite a character, full of energy. What I found out shortly after I got here was that the Center Director, Bruce Burlington, had come from CDER. He was influenced by Carl Peck who was used to be the Center Director for CDER but left for Georgetown. Carl was a pharmacologist by training and was very interested in Bayesian ideas but he couldn't get anyone in CDER to work with him. When Bruce came to CDRH from CDER to be the Center Director, there was a report called the Temple Report that FDA had written about a couple years before which was highly critical of the way devices were evaluated in CDRH. This was a report by a committee that had a lot of people on it but the one person on it was Bob Temple, the Chief Medical Officer in CDER. One interesting thing was that there were no criticism of statisticians in the report; the criticisms were all about the way the engineers evaluated the medical devices and there was very clinical expertise at all. So, one of Bruce Burlington's jobs when he arrived in CDRH was to hire a medical officer named Susan Alpert to lead the Office of Device Evaluation. So what happened as a result of the Temple report was that the biostatistics group became more important within CDRH because these statistical reports were no longer ignored. People would then pay attention to what we're doing because the Temple report looked at the memos that the biostatistics group had written and they liked these memos but then decisions wouldn't reflect the statistics opinions in the memo.

Shortly after I got here within a year or so, Larry Kessler said "What do you think about using Bayesian statistics in FDA?" and I said I hadn't really thought about it much, since when I was in NIH I didn't do anything Bayesian. I had this thesis which was Bayesian and when I was at Purdue, I interacted with the Bayesians at that time and Jim Berger put together a semester-long Bayesian course which I attended. So, I told Larry we could try to do that. Now the problem was that at that the time no one else in the Division of Biostatistics knew much about Bayesian statistics and it wasn't clear if they wanted anything to do with it either. So this is where the leadership part comes in. How do you get people to see your vision that they don't necessarily see or want to do? So we formed this reading group called "B-Team", "B" for Bayes. We read papers and brought in some speakers to do short courses and hired a man from NCHS, Don Malik, who had done some Bayesian work in sample surveys. So he came to CDRH on an Interagency Personal Agreement and stayed for about one and half years. He helped us to think through how you would do things. We were trying to figure out how to use prior information and one of the ways that I thought could work was to build hierarchical Bayesian models and Don had some experience on that and he wrote some software programs. So he came here and helped us a lot. The idea basically is you would borrow data from another source and you would let the Bayesian hierarchical model figure out how much to borrow. And if the data is really far away the model realizes it and it borrows nothing. If the data is really close, you borrow a fair amount, and when it is farther away you borrow less. There's actually a point at which when you're just far enough away

you don't borrow anything and actually cost you some data to realize that the model is wrong and that you shouldn't be borrowing anything at all. That actually appealed to me at that time because if a company gets it wrong by too much they do not borrow anything that didn't bother me; if they get it right then good for them. But if they get just wrong enough they need to know there is a penalty that they might have to pay. This is an example of leadership getting the group to think about Bayesian ideas. Early on we also hired two Bayesians, Telba Irony and Gene Pennello. One of the real problems was how much should we worry about the operating characteristics, how much should we worry about Type I error. Controlling Type I error is important to the FDA because when the FDA makes a Type I error, it approves a product that does not work or is unsafe.

The other thing that I would say is that it is helpful for leaders to think strategically. You want to prevail in a long run, and whether you lose a battle along the way isn't really as important as keeping in sight the long-range vision. One of the things that I learned about leadership is that people work for themselves; they don't work for other people. In the final analysis everybody works for himself or herself and that is so true. What is your job as a leader to do is not to boss people around and tell them what to do, it's to influence them and to try to get down to see the advantage from their point of view of doing certain things and not doing other things.

Another important thing is always to set an example in everything you do. Work hard. It doesn't hurt to be the hardest worker in the group. People see that and that will help. The other thing is to strive for excellence, don't just settle for getting something done; if you can do something, do it right, and do it well. The other thing about leadership is that leadership is service but a lot of people don't get this; a lot of people who are bosses don't get this. But the best leaders are those who look out for other people. And, the last point is that it is a skill you can learn to be a good leader; it helps to have good role models but you can learn it anywhere. At FDA and I think NIH as well, there are courses you can take on leadership skills, and it's not just management. You can be a leader without being a manager which is helpful to know. And you can be a manager without being a leader.

In terms of statistical leadership, the advice that I have for people is if you're going to be a statistician doing applied work, what you want to do is not just to be a consultant, instead, you want to be a collaborator. To be a good collaborator, you have to understand more than just the statistical aspects. To be a collaborator, you need to understand what the scientific issues are, which may require you to learn some biochemistry or medicine. For whatever the subject matter is, you need to understand the science. Part of that is that you have to be involved early; you have to be involved at the time when people begin formulating how they want to attack a problem.

5. Contribution to regulatory science

Aiyi and Lilly FDA focuses heavily on regulatory science. Can you tell us the contributions you have made in FDA regulatory science?

Dr. Campbell I can certainly talk about the guidance documents, but you know regulatory science is much broader than that. It took us a long time to write and revise the Bayesian guidance document. It came out in draft form around 2005, and was finalized in 2010. It was a group effort; there were number of people who were on the writing team besides myself.

Writing FDA guidance is uniquely regulatory; it's just guidance and not rule making. Nonetheless it's an important thing to do. The challenge is to make it understandable to people who are not statisticians; there are some fairly complicated Bayesian hierarchical ideas having to do with the exchangeability which are hard to describe without mathematics but that's what we were trying to do in the Bayesian guidance document.

There were other guidance documents. I was involved in the FDA Data Monitoring Committee guidance document. This was an effort led by Susan Ellenberg, and I was the CDRH statistician on the committee. That took between one and two years to write and then another two to three years to finalize. The final came out in 2006 so that was a good experience because Susan Ellenberg had experience with Data Monitoring Committees from her time at NIH and most of other people had experience working with companies that had Data Monitoring Committees. We worked through and there were a fair number of statistical issues that needed to be taken care of.

And, the latest statistical guidance is the Adaptive Designs for Medical Devices draft guidance which came out in April 2015. We had a group mostly of statisticians but also involved a few clinicians and engineers as well. I also led the effort for another guidance document, one on pivotal clinical study designs for medical devices. That was an effort to try and write down all the different kinds of designs that were possible in the area of medical devices and diagnostic tests. That turned out to be very interesting, because it's a much broader horizon than you would see in the drug world where you talk about two-arm, placebo-controlled trials.

Thinking about the regulatory research, I was involved in this project with microarrays, called the Microarray Quality Control (MAQC) project. It was led by a chemist at FDA, a man named Leming Shi, at FDA's National Center for Toxicological Research in Arkansas. He had led the project in the first phase and then in the second phase (MAQC-II) they were interested in prediction models; they wanted to see whether they could use microarray data to make predictions. They used a Federal Register Notice to encourage researchers to donate microarrays to the project. They actually got a fair number of data from clinical trials in cancer where they had the microarray data. Part of the effort was quality control; they wanted to make sure they had good quality microarrays. This was the early technology before the next-generation sequencing. So there was a biostatistical regulatory group which I headed. The group wrote a white paper by all of the people in the project, which would provide advice about how to build good statistical models that were predictive. I talked about bootstrapping and cross validation and all the techniques that you would use to build a model and try to estimate what its performance would be like if you have a second dataset. The interesting thing about this project was that we had two datasets, a training set and a test set, and everyone got the training set. There was a call worldwide for groups who would want to participate in the project. A total of 35 groups from all different countries, such as

China, Germany, Sweden, and the US, had access to the clinical datasets. So, we wrote this document and the biostatistical regulatory group that I headed had people from FDA, industry, statisticians, and non-statisticians. People in various international groups used all different techniques such as machine learning and Bayesian classifiers, all kinds of things. It was an interesting project, because everyone built the models and then we said “OK, lock in your model and now we are going to give you the test dataset for the prediction but we’re not going to tell you who’s in which group”. In addition to the microarray data, for the clinical datasets, there were covariates too. So one of the questions is not just “Can you build a model that would show that the microarray data does better than no microarray data”, but the real question is “Can you build a model where the microarray data add to what you have already known clinically from the clinical baseline covariates?” So there were people on the project who didn’t understand why that was important. I pushed pretty hard to let people realize that it’s not enough just to show that you can do just as well using the microarray as using just simple covariates such as how old the patients are or how long they have had cancer.

So the interesting thing about this was that people built their models, they then got the test datasets without the labels, and then they sent them in. Eventually they revealed the true patient classifications (cancerous or not, recurrent or not depending on the study), so you got to see how good the models were and how good the internal validation might have been from the bootstrapping and jackknifing, and then how well it was in the second dataset. So we learned a lot, and I think people in the center learned a lot, about building models, about diagnostic medicine, about the fact that when you do cross validation, you do random splits in the data that doesn’t always help. Some people did very well; they built good models. Some people thought their models performed pretty well on the training dataset but not so well on the test dataset. In general what happened was they didn’t follow the white paper that we had written and they were a little careless about keeping track of when and how to do the cross validation. There was a main paper in the *journal Nature Biotechnology* and I was the second author. That’s a great example of regulatory research.

6. Statistical innovations

Aiyi and Lilly We all know that you have made outstanding innovative contributions in medical devices and biostatistical methodology. Would you like to tell us in some details about your statistical innovations?

Dr. Campbell I can say a little bit about statistical innovations in the world of medical devices. Lilly Yue and I have recently written article for the *Journal of Biopharmaceutical Statistics*. In the article we identified areas that we thought were opportunities to really be innovative in the world of statistics and the world of medical devices. The first was the Bayesian example of two really different kinds of approaches. One is using prior information with the Bayesian hierarchical model, and the second is the Bayesian adaptive design where you may not have prior information but would use accumulating information from the trial to change the trial in some way, and we’ve seen lots of examples of both.

A second innovation (Lilly knows a lot about this too) is the area of propensity scores and causal inference. About more than 10 years ago, we saw that companies would come to FDA and say “we have a historical control but we don’t have a concurrent control. So we will start a single arm trial and compare our device to the historical control which usually could be literature or something the company has done.” But they are not quite so sure whether these two populations are similar. So an idea was why we don’t make this a whole lot more statistical. So, one way to do that of course is to do it by putting in the covariates and trying to adjust for the covariates in a direct fashion. There was some work done by Don Rubin and Paul Rosenbaum on using propensity scores and asking what would have happened if someone has gotten the therapy in a two-arm study, a counterfactual idea. So there was this propensity score machinery and you could use that in different ways such as clinical trials with missing data. The missing data basically destroy the experiment and the propensity score model can be used to try and bring it back to alignment. You can also use it for nonrandomized studies and this is what our colleagues in epidemiology have done for a long time in observational studies to figure out how comparable two groups are. So, Lilly was involved in the first advisory panel in 2004; she gave a presentation of a propensity score model and did a tutorial too. And then, we thought that was pretty interesting and so we started to ask the companies to do that. And whenever there was an advisory panel where some company used an historical control, we would encourage them to build the propensity score model. In some cases they refused to do so, but we would present the propensity score model. The beauty of the model is that, if you build a propensity score model, you can look at the propensity scores and see if you had the kind of balance that would suggest the two groups are comparable. You don’t have that in the other methods. So this was a way that we could check to see if the two groups are really comparable or not. That’s really a successful story. More recently Lilly has done some really fine work where we want companies to pre-specify how they’re going to build the model and to do it without seeing the outcomes. One of the criticisms I think to our epidemiology friends is that they build models where they have access to the outcome data and they may not be careful of about separating the outcomes from the baseline covariates that you would use in building the propensity scores. The propensity score is, of course, the probability that a person would have been, based on his/her baseline covariates, assigned to one treatment instead of the other. So we’re now having a lot of success getting companies to pre-specify that, which is just a much stronger scientific approach.

The third innovation example is the adaptive designs for medical devices and we have done a lot of thinking in the last couple years about that. We brought in some people to do short courses, and now have written a draft guidance, which I think is going to be innovative. It certainly encourages companies to do the simulations as a way of understanding the operating characteristics. The fact that we had the Bayesian experience helps us enormously, because we had thought about many of these problems already. In order to do these Bayesian trials and do them right, you have to do simulations to understand the operating characteristics of the same design, and the same thing is true in the adaptive world.

A fourth innovation is missing data which is a problem in general for FDA, because there is a lot of missing data. In the device world, there are probably even more, particularly in a longitudinal study, because patients tend to drift away. If you have a permanent implant, you

may want to monitor people with that implant for a year or two years. We are well aware of that and there are some very sophisticated statistical models for missing data. But then, it seems like a statistical exercise and one of the problems with those models is that you make assumptions about the models and no one knows if those assumptions are true or not, so the models can only get you so far. The other problem with missing data at the FDA is that we want companies to pre-specify what they are going to do (with the missing data) and that's really hard. It's another example appeared in the literature, where people had a lot of time to think about the missing data but they hadn't prespecified it; they just continued to build models until they found a model they liked. That's not necessarily good science. So I came up with this term "tipping point". The idea is that suppose you have a number m of people who are missing in the control group and a number n of people who are missing in the treatment group. So, you create a grid of $m \times n$. In the case you had a binary endpoint, you could imagine a number of people who are successes and a number of people who are failures in each of the two groups. Each block in that $m \times n$ grid represents the number of successes in the control group and the number of successes in the treatment group among people who are missing. Then the question is that if you had that and you have done your inference, would it change? For every point in the grid, you can figure out whether it changes or not. If it changes, then that point is a tipping point, so you can describe the grid in terms of points that would cause you change your inference or not. So you would get a picture that has a tipping boundary associated with it so that on one side of the boundary you don't change inference and on the other side you do. The advantage of it is that you don't need any modeling and it's really easy to show it to the clinicians. You can point to a point in the grid, and say that here is the average success rate among people who had complete data, and here's where the boundary is, the tipping boundary, so you can see whether it changes or not, and if it does change how far you would have to wander away before there's a change. That turned out to be very helpful, so a number of CDRH statisticians wrote a paper in the *Journal of Biopharmaceutical Statistics* about tipping point and we have been using it ever since to encourage companies to present that. If they don't present it, we present it as a way of getting the FDA Advisory Panel to think about missing data and how important it is to worry about that.

The fifth innovation is the whole area of diagnostic statistics. It's an area which when I came here to FDA was not very well-developed. People knew about sensitivity and specificity, but at that time ROC was not used much at FDA. When I came, there was a physicist, named Bob Wagner, in the CDRH's Office of Science and Technology (now OSEL), and he was very interested in ROC. He knew about my work from NIH, and so we ended up writing a lot of papers together. He was instrumental in evaluating the first digital mammography system approved by FDA, where the decision about whether to approve that or not was based on the ROC machinery. One of the problems then was that FDA had written a guidance document on digital mammography that came out basically in 1995 around the time I came here but I had no input into it. The guidance basically said if you want to do a digital mammography submission, you need to show that your mammography agrees with analog mammography. Companies tried but couldn't do it. The reason is that there's always variability. Once you have a mammogram, whether it is digital or analog, you have to have a reader to interpret the images and there's variability associated with the readers such that

you may not be able to show the equivalence of analog to itself. So companies failed. The work Bob and I did was helpful in guiding companies to do reader studies where you have more than one reader and you would model the variabilities associated with the readers, as well as everything else. The whole area of diagnostic statistics was not well understood, and the paper I had written with Mark Zweig, the clinical chemist at NIH, proved to be very helpful. He was not coming from the radiologic tradition where most of the ROC work has been done, instead, he was coming from clinical chemistry. At FDA at that time, there was a pretty big group that evaluated *in vitro* diagnostic products and they approved a lot of lab tests. A lot of lab tests needed ROC type of methods, so we ended up using them there as well. And then, there're whole host of interesting statistical problems that group has dealt with since then. What happened if you don't know the true disease status of some people? Kristen Meier and I co-authored a paper on that. Anyway, it's a very fertile area, it's really developed a lot in the last 20 years, and there's a lot of work that still needs to be done. But I think FDA has been a leader in this area. Kristen Meier in CDRH helped write a guidance document on reporting the results from studies evaluating diagnostic tests in 2007.

7. Other important contributions

Aiyi and Lilly What other important contributions do you have that you feel proud of?

Dr. Campbell First and foremost, I am very happy with the really talented statisticians that I've managed over the years to hire in the Division of Biostatistics in CDRH. In the 20 years at FDA from 1995 to 2015, I hired a total of 80 statisticians into CDRH. Not all of them stayed and there were reasons why some might have left, for family reasons and so on. But I am happy that uniformly excellent statisticians were enticed to come to FDA under my leadership. In the time that I've been there, another gratifying situation has been that a number of these statisticians have developed into fine leaders within the statistical community. That includes leadership in terms of research, leadership in terms of ascension to the Team Leader and Branch Chief positions, and the two Deputy Directors for the Division, Lilly Yue and Gerry Gray. I'm very excited about the future for the division because it's in such good hands and the people are all first-rate.

Another thing that I'm proud of on a more personal level is that I've been a Fellow of the American Statistical Association for more than 15 years. In addition, shortly after I arrived at FDA, I was nominated for the Senior Biomedical Research Service (SBRS) which is a very prestigious honor within the federal employment sector; within NIH and FDA, there are relatively very few researchers who are singled out for this accolade. For a period of more than 15 years, I served in SBRS. Not only that, I also served on what's called the Credentialing Committee for the SBRS that evaluates candidates. I am happy to report that there are now a number of statisticians at FDA in SBRS: Jim Chen from the National Center for Toxicological Research, more recently Lilly Yue in CDRH, and Jim Hung and Sue Jane Wang in Center for Drug Evaluation and Research (CDER), have joined the ranks as well. I think that's a real tribute to the statistical talent that FDA has managed to attract and develop.

One of the other things I am very proud of has been the development of the statistical community specific to medical devices and diagnostic products. This began in 2004 when Joint Statistical Meetings (JSM) was held in Minneapolis, Minnesota. I thought that it would be a great opportunity to create Topic Contributed Sessions at the JSM in order to grow the number of statisticians who would come to JSM with an interest in medical devices and diagnostics. Up until that time there were very few sessions that would be unique to the medical device field so this was an effort to try and change that.

Every year since 2004, I would send out an e-mail to a group of people who were statisticians with an interest in medical devices and encourage them to send their title and abstracts in advance to us. Then Lilly Yue and I would put together Topic Contributed Sessions and we did that for a number of years. We had upwards of 10 or 12 or more Topic Contributed sessions each year. We formed an interest group in the ASA expressly for statisticians interested in medical devices and diagnostics. We called that SIGMEDD which stands for Statistical Interest Group for Medical Devices and Diagnostics. That functioned for a number of years; it was relatively informal. We would meet every year at the JSM for a business meeting. Eventually, the topic arose that maybe we should form a section and so a number of us began to take steps in that direction. In 2014, the new section for Medical Devices and Diagnostics was officially launched so that is another very satisfying accomplishment.

Another effort in trying to build up the statistical community was an effort that FDA had with the AdvaMed which is a trade group for medical devices and diagnostics. I approached them about eight years ago and asked if they would be interested in co-sponsoring a statistical workshop with FDA. They said yes and so we now have I think eight successful years of statistical workshops. Every year we have the annual statistics workshop cosponsored by FDA and AdvaMed where the focus is on statistics in medical devices and diagnostics. The last one of these drew around 180 participants and had on the second day of the two-day event two tracks, one for therapeutic medical devices and the other for diagnostic devices

8. Leadership in the broader community

Aiyi and Lilly What about your leadership in the broader statistical community and internationally?

Dr. Greg Campbell I have enjoyed very much the opportunities to help to lead the development of statistics in the broader community. The Eastern North American Region of the International Biometric Society, ENAR for short, is one group that meets every year in the spring and topics in biometry and biostatistics are part of that conference. I served on the Regional Advisory Board for ENAR for three years from 1999 to 2001, and then became the Chair of this Regional Advisory Board for a two-year period right after that. It was fun to observe how leadership can really make a difference in guiding societies such as ENAR. I also served on the Board of Directors for the Society for Clinical Trials for a stint and was also a member of the Executive Committee for the Biopharmaceutical Section of the American Statistical Association for a three-year period. It's been an interesting effort to

think about leadership not just in the medical device community but in the broader areas of biopharmaceutical statistics, clinical trials, biostatistics, and biometry that the societies such as ENAR, SCT, and the ASA's Biopharmaceutical Section provide to the broader statistical community.

So the second part of broader statistical community concerns the international aspect. Over my career, I have been lucky enough to be able to travel and give talks in Hungary and Belgium in the 1980s, in Germany and Scotland in the 1990s, and in my role at FDA, I have had the luxury of being able to travel to Canada, Australia, Japan, and Saudi Arabia, and talk to the FDA counterparts in those countries about the importance of statistics in the regulatory environment. In addition, I have also had a chance to give talks recently in Ireland and in 2015 a keynote to China Region of the International Chinese Statistical Association in Shanghai. I have also visited in the past year Japan a couple times and Belgium, and talked about statistical issues. It is another aspect of influence in the broader statistical community.

9. Advice for younger statisticians

Aiyi and Lilly From your experience as a statistician and a Director of a statistical group, what advice would you give to young statisticians in order to have a successful career?

Dr. Campbell I have a number of pieces of advice which I'd like to give, seven in all.

The first is I think what's really important is to have a passion for statistics (or to develop one). It's a great time to be a statistician. We have a great computing power at our fingertips and digital journals at our desk computers. It is the era of big data. These next 10 years from 2010 on have been the decade of the statistician according to Hal Varian, the chief economist at Google, who said that the sexy job in this 10-year period is that of the statistician. We need to make sure that people understand and develop this passion.

The second point is to encourage our younger folks to develop good communication skills, to learn to speak well, to write well, and to listen well. "You can observe a lot by watching" as Yogi Berra has said. Truly listening can help you to be a very good consultant. This reminds me of a quote by John Tukey "It's far better to have an approximate answer to the right question which is often vague than an exact answer to the wrong question". So a really good skill for consultants is to be able to listen carefully and understand exactly what's being asked, to get beyond the simple question which might be "What size sample do I need for my study?" and understand in a much more detailed way exactly what the researcher is trying to accomplish.

I would encourage as a third point for younger statisticians to take the opportunity to continue to learn every day, to educate yourself by availing yourself of short courses, go to conferences and listen to talks, take courses at your institution if they're available and not limit yourself just to statistical learning but to learn as much as you can about the areas of application of statistics because these will help you do a much better job.

A fourth piece of advice is to jump in enthusiastically with both your feet into a problem, to volunteer for statistical as well as non-statistical activities, to get to know your institution,

the people in your institution, to keep active professionally, to be willing to ask for help. This is all part of jumping in with both feet.

A fifth point is that it is helpful that you find a mentor. If you have that opportunity where you work, you should avail yourself of it and also find other people you can talk to about problems that might come up. The Gallup organization has indicated that the best single predictor of job satisfaction is having a best friend at work. So you might keep that in mind as you think about your work experience and try to cultivate that so you can have some good satisfaction in your job.

A sixth piece of advice is to keep a healthy work–life balance. It’s very easy to become caught up in the work, in all your enthusiasm and passion, but it’s also important to keep in mind that there is life outside of work, there is family and I have been blessed with a great family, my wife, two children, and now four grandchildren. Friends are another great way to help balance your work–life. The third thing in this regard is to think about vacations as a down time, a time to take time out. It is not a real vacation if you are reading e-mail or calling in for messages. So when I go on vacation, I plan to do little work if any.

The final piece of advice for young statisticians is to plan ahead and prepare for success. There’s a great quote that I really like by Seneca: “Luck is what happens when preparation meets opportunity.” Let me finish this question for younger statisticians with this advice, that is, those folks who plan ahead are best prepared to succeed and that is in fact a secret of success.

Aiyi and Lilly Thank you so much for this great interview opportunity.