Interview with **Professor Geert Molenberghs**

Toshimitsu Hamasaki and Scott Evans

Geert Molenberghs.

CHANCE invited Dr. Geert Molenberghs, professor of biostatistics at the Universiteit Hasselt and KU Leuven in Belgium, to talk with editor Toshimitsu Hamasaki and executive editor Scott Evans.

Molenberghs is internationally recognized as a leading biostatistician. He has made important contributions to numerous areas of statistics, in particular methodological developments on surrogate markers in clinical trials, categorical data, longitudinal data analysis, and the analysis of non-response in clinical and epidemiological studies.

Molenberghs received his BS degree in mathematics and PhD in biostatistics from the Universiteit Antwerpen. He is founding director of the Center for Statistics at Hasselt University and currently director of the Interuniversity Institute for Biostatistics and Statistical Bioinformatics (I-BioStat), a joint initiative of the Hasselt and Leuven universities.

Molenberghs has served as joint editor of *Applied Statistics* (2001–2004), co-editor of *Biometrics* (2007–2009), co-editor for *Biostatistics* (2010–2015), and president of the International Biometric Society (2004–2005).

Molenberghs was elected a Fellow of the American Statistical Association in 2003, and became an elected member of the International Statistical Institute in 2007. He has received numerous awards, including the Guy Medal in Bronze (2002) from the Royal Statistical Society; Outstanding Contribution to the Society Development by the International Biometric Society (2010); Best Contributed Paper Award from the Biopharmaceutical Section of the ASA (2005); Donald E. Francke Award for Overall Excellence in Journal Publishing from *Drug Information Journal* (2009); and Excellence in Continuing Education Award from the ASA in 2003, 2005, 2006, 2009, and 2011.

With Professor Geert Verbeke at KU Leuven, Molenberghs edited and authored several books on longitudinal data analysis (Springer Lecture Notes 1997, Springer Series in Statistics 2000, Springer Series in Statistics 2005, Chapman Hall/CRC 2007). He has taught well over 100 short courses at universities and in industry in Europe, North America, Latin America, Australia, and Japan.

Hamasaki and Evans talked to Molenberghs about his remarkable career, and specifically his research and educational activities in biostatistics.

With an educational background in mathematics at Universiteit Antwerpen, when and how did you first become aware of statistics as a discipline?

Like in so many things in life, there was a good deal of chance involved. After graduating as a mathematician, I started working on a PhD in algebra at the University of Antwerp. In the late 1980s, military service was still compulsory in Belgium, but I opted for the alternative "civil service." At the time, it meant 20 months in an organization or institution of societal relevance, shortened to 16 months if you were a medical caregiver. I ended up in the biostatistics department of the Leuven academic hospital, under the guidance of Emmanuel Lesaffre. At the end of these 16 months, I was converted to biostatistics and decided to stay on that path.

Why did you become a statistician working primarily in medical research? What was it that inspired you to pursue a career in statistics or biostatistics?

While my initial exposure to biostatistics was an event of chance, it was love at first sight, different from anything I had seen in my very theoretically oriented mathematics training. I was thoroughly charmed by the combination of mathematics, modeling, programming, and applications.

> "…**new times bring new challenges** that require flexibility and bring the need for adaptation. But this does not imply that we should let go of our foundations and principles…"

That it is possible to make sense out of data through models and its parameters, using standard software or one's own code, was magical to me then, and it still is now. Over the years, I have had great joy in trying to understand subtleties in the underpinnings of statistics, as well as in the development and use of modeling trials.

You are a co-founder of the Center for Statistics (CenStat) at Hasselt University, which was established in 1998. During the last two decades, CenStat quickly became internationally renowned, for both its theoretical and its applied component. What is the biggest challenge in growing the organization?

When the institute was founded, it was thoroughly rooted in a solid tradition of mathematical statistical research on the one hand and the then nearly decade-old master of science in biostatistics program on the other. These were no small achievements of the research group's founding fathers: Herman Callaert, Noël Veraverbeke, and Paul Janssen.

In the late 1990s, components of biostatistics research and consulting were emerging and, to increase visibility and consolidate activities, the center was founded. It has been growing ever since. This in itself, perhaps, is the greatest challenge: an expanding organization by definition does not find itself in a steady state. It continually brings administrative and practical headaches.

Ten years later, the interuniversity umbrella I-BioStat (Interuniversity Institute for Biostatistics and Statistical Bioinformatics) was formed, a joint venture bringing together CenStat at UHasselt and L-BioStat at KU Leuven. The strongest asset of the institute as a whole is a fine group of dedicated faculty, all with their own strong and often internationally recognized research lines, but at the same time, a lot of within-institute, as well as national and international, collaboration.

You served as president of the International Biometric Society (IBS) in 2004 and 2005. The IBS, is one of the largest statistical communities and is promoting the development and application of statistical and mathematical theory and methods in the broader biosciences, including agriculture, biomedical science and public health, ecology, environmental sciences, forestry, and allied disciplines. How did you lead the IBS and coordinate all of the IBS activities during your time as president?

My IBS presidency was a coincidence in its own right. I served as general secretary on the Executive Committee when vice president Rob Kempton (from Biomathematics and Statistics Scotland) untimely passed away. Ad hoc elections had to be organized to replace Rob for the remainder of 2003. I then became president in 2004–2005, followed by the customary year of outgoing vice presidency in 2006. The new president-elect was going to have much less time to travel through the inevitable learning curve than usual. This is why several people looked in my direction: I was serving on the Executive Committee already, and benefited from prior experience with the society as former editor of the *Biometric Bulletin*.

The most important assets during my presidency term were the preceding and succeeding presidents: in 2004, Norm Breslow was outgoing vice president, followed by Tom Louis in 2005. Both are giants in the field, scientifically as well as in terms of the leadership and vision they displayed. I learned a lot from them.

The IBS is a confederal society: Regions and Networks have a lot of autonomy and there is room for local culture and focus. Some entities place more emphasis on statistics in agriculture, for example, while others focus more on medical statistics. Some Regions are very small, with a few dozen members, while the largest ones—the German Region and the Eastern North American Region—have more than 1,000 and more than 2,000 members, respectively.

I started a visiting program where Executive Committee members, in particular the president, were encouraged and enabled to visit a number of Regions during their terms. It took me, a European from a small country, to Asia, Africa, South, Central, and North America. It was a very rich experience, perhaps for the Regions visited, but definitely for the president. A beneficial effect is that one gets to know personally a much wider circle of colleagues at large. This is helpful when soliciting volunteers to serve on the organization's committees and boards.

You have served as the editor of several internationally recognized, high-quality statistical journals over the years, including *Applied Statistics, Biometrics,* and *Biostatistics*. How did you promote and optimize the quality of these journals?

Each of these journals is a little different in the way it is managed. *Applied Statistics*, at the time one

of the four (now one of the three) scientific outlets of the Royal Statistical Society, is very closely managed by the society's officers and staff, with executive editor Martin Owen undoubtedly being the most important person in the journal's success, for many decades now.

Biometrics is organized differently. While it is also a society-owned journal, the International Biometric Society this time, the pivotal people are Marie Davidian (North Carolina State University), executive editor, and Ann Hanhart, editorial manager. Both of the journals are published by Wiley.

Biostatistics is different still, since it is not society-owned, but rather an initiative of the Biometrika trust, within Oxford University Press.

Incidentally, all three journals work with multiple editors. The RSS journals have two so-called joint editors each, serving four-year terms, with one of them changing over every two years. *Biometrics* currently has a system of three co-editors, serving three-year terms each (one from North America, one from Europe, one from "Rest of World"), with a changeover every year. Thus, the incoming editor can benefit from the experience and good advice from the senior colleague.

Biostatistics is different in the sense that currently only the third pair of editors is serving. The journal was founded in 2000 and the founding editors, Peter Diggle (Lancaster University, UK) and Scott Zeger (Johns Hopkins), served for a full decade. Then, Butch Tsiatis (North Carolina State University) and I took over for two three-year terms. We had agreed to serve for six years and that is what we did. Working with Butch Tsiatis was both very pleasant and an educational experience for me.

Geert Molenberghs, ASA Past President Sally Morton, and Geert Verbeke (left to right) at the President's Invited Address reception at the Joint Statistical Meetings in Washington, DC, 2009. Photo courtesy of Eric Sampson.

The current and third editor team of *Biostatistics* consists of Dimitris Rizopoulos (Erasmus University Rotterdam, the Netherlands) and Jeff Leek (Johns Hopkins).

Throughout these varied experiences, I have learned that if one favors quality, everyone wins. A key to the success of a journal is the team of associate editors. Their role is essential—they are the experts in the various areas and should combine scientific excellence with administrative efficiency, to ensure not only high quality but also timely turnaround. Given that, there is no need for a revolution in the way the journals are run ("If it ain't broke, don't fix it"), but not for a standstill, either.

Good management of a journal, I believe, involves a slowly varying "improvement function" over time. Editors should not merely follow their intuition, but listen carefully to the associate editors, referees, authors, executive editors, staff members at the publishing houses, etc.

Are there people or events that have been influential in your career? Who have been your mentors? Is there something in particular they did to help you?

Perhaps it is fair to say that "we are who we encounter and work with." That definitely applies in my case. I already mentioned my transition from algebra to biostatistics. This would not have happened without the contagious enthusiasm of Emmanuel Lesaffre and his guidance and leadership. When I started working in statistics, on December 4, 1989, I encountered Geert Verbeke and now, over a quarter of a century later, we are still actively working together on numerous statistics and sometimes administrative projects.

During my time in Leuven, Emmanuel introduced me to Mike Kenward, to the current day a fine collaborator as well.

After graduating, in 1993, I was appointed by Herman Callaert and his colleagues at

Hasselt University and dispatched to the Harvard School of Public Health. There, the late Steve Lagakos, Louise Ryan, Stuart Lipsitz, and Garrett Fitzmaurice had a defining impact on my career and research interests.

At Hasselt University, I started working with Els Goetghebeur and Marc Aerts. Marc and I are still at the same institute and our scientific and administrative collaboration continues to this day. At that time, in 1994, I had a conversation with Marc Buyse about some technical issues in surrogate marker evaluation. The conversation had consequences lasting to this day: More than 20 years later, we and a group of colleagues published a second book on the topic.

I already mentioned Marie Davidian, Butch Tsiatis, Norm Breslow, and Tom Louis, who have had a very important influence on me as well; this is also true for Clarice Demétrio, another former IBS president.

Looking back on the colleagues I was fortunate enough to find crossing my path and what I have been working on over the years, I have to admit that it is amazingly concurrent: categorical data, multivariate data, and classification with Emmanuel; mixed models and longitudinal data with Geert; missing data and sensitivity analysis with Mike and Els; clustered data and dose-response modeling with Louise; semi-parametric methods through Stu and Garrett, etc.

Of course, there is more than scientific and administrative collaboration. Many of the people listed I consider very close friends, too. It is also fair to say that Mike Kenward has always been there as a mentor as well.

What skills and qualifications are most important for a statistician working in medical research? What's the potential for career development?

I sometimes compare the work of a medical statistician to that of a surgeon. Ideally, the surgeon is well-trained and skilled, and has vast knowledge and practical experience. But, that said, every patient is different and, when in the operating theater, ad hoc or, shall we say, patient-specific decisions have to be taken. It is arts and science, and so is medical statistics.

No matter how experienced we are and how much formal and informal training we receive, every study, every design, and—for that matter—every data set is different. Open the data set, look at the data, and you may see peculiar features.

Over the years, I have seen many, especially theoretically trained students, become frustrated over this aspect. Arguably, this is why teaching statistics can be perceived as very difficult, even to the mathematics teacher. Some people prefer the "cleanliness" of pure mathematics research, while others enjoy the "organic" nature of applied statistics, whether geared toward the medical field or elsewhere.

What are the first steps in entering biomedical research as a statistician?

In view of the above, I think it is best for a statistician working in biomedical research to be well aware of (empirical) research in at least one area. There is an irreplaceable cross-fertilization between consulting and research. The methodologist working in biostatistics and a related field should, therefore, know what matters in the medical or clinical practice. The most successful research CVs are a blend of contributions to the medical literature and to the methodologically oriented statistical literature.

With this in mind, "exposure to the data" cannot start early enough, and arguably during master's-level training. Fortunately, many graduate programs make room for this.

You are best known for your methodological work on surrogate markers, incomplete data, and longitudinal data analysis, all of which are important and hot topics in clinical trials. What led you to this research?

It will be evident from the above that the choice of research topics has to do with people, first and foremost. Mike Kenward added missing data to my research agenda when I was working on my PhD dissertation; with Geert Verbeke and several others, I have been working on longitudinal data, and I would not have been working on surrogate marker evaluation without Marc Buyse. The collaborators from day one were then joined by PhD students and postdoctoral collaborators, at which point one's research themes start to solidify, attract further collaborations, etc.

What do you feel have been your most important contributions to biostatistics and medical society? What achievements are you most proud of? What has been the most exciting development?

Perhaps we aren't well-placed enough to judge our own contributions in terms of importance but, at the same time, it is true that some work gives a sense of pleasure and achievement.

This is especially true when, for example, research and teaching come together.

For example, as an outgrowth of workshops organized by Paul Janssen in 1994 and 1995, Geert Verbeke and I published our 2000 and 2005 books on linear mixed models and discrete longitudinal data, respectively. The 2000 book, with its 1997 predecessor, has been cited more than 7,000 times. It brings together a general synthesis of the field with some of our own research, and has been a tool for our and others' shorter and longer courses. Teaching the material is a very rewarding experience, as it brings one in close contact with colleagues and students from around the globe—it engenders new research ideas, and sometimes consulting and funding opportunities.

The same is true for several other book projects on longitudinal data, missing data, and surrogate markers.

While I was one of many authors of the 2010 National Academy of Science report on missing data in clinical trials, commissioned by the Food and Drug Administration, it was a source of pleasure and pride to be able to work with Rod Little and so many fine colleagues and, in the process, to have a little bit of impact on clinical-trial practice. In the same sense, I am grateful for the ability to have been able to contribute to handbooks on longitudinal and missing data, jointly with Garrett Fitzmaurice, Geert Verbeke, Marie Davidian, Mike Kenward, and Butch Tsiatis.

Some very satisfying work is done far away from the spotlights, such as serving on Data Monitoring Committees in clinical trials, especially when the products under investigation are in severe, sometimes life-threatening conditions.

Committee of Presidents of Statistical Societies 50th reception in 2013 (left to right): David Banks, Geert Molenberghs, and Jasper Molenberghs.
Photo courtesy of Eric Sampson.

You have taught statistics or biostatistics at many universities and institutes over the years, and have received the Excellence in Continuing Education Award of the American Statistical Association many times. What are the keys to teaching statistics or biostatistics while adapting to meet the changing needs of students?

A lot has been said and written about teaching statistics, and I don't consider myself an expert in the matter, except maybe that I have done it a lot.

In our Belgian system, it is not uncommon for professors to teach between five and 10 courses, sometimes as the sole instructor; at other occasions, as a member of a teaching team. In addition, as already mentioned, I have had the good fortune of teaching abroad, in Europe and the rest of the world, oftentimes with Geert Verbeke on longitudinal and missing data.

Audiences are varied, not only geographically and culturally, but also in terms of background. Teaching statistics for mathematicians, master's students in statistics, or biomedical students is profoundly different. Over the years, I have come to understand the importance of gauging the level of abstraction the audience is comfortable with. Pitching a course at the wrong level means that the fundamental concept or the modeling principle you are trying to get across is completely lost.

Also, it is not necessarily a good idea to give the audience what they most want. Indeed, solid principles, which they can rely on at times when you as a teacher aren't available, are so much more important than cookbook statistics.

Geert Molenberghs discussing Models for Discrete Repeated Measures at JSM 2011.
Photo courtesy of Eric Sampson.

Finally, the best way to raise interest is by showing one's own enthusiasm. When compatible with one's personality, the teacher can use a bit of humor to make sure that not only the joke but, in particular, the statistical concept attached to it is remembered.

What do you think the future of teaching statistics or biostatistics will be? What do you think will be the upcoming challenges in engaging students?

Never before did we have so many communication tools at our disposal. Chalk and blackboard became whiteboard, overhead transparencies, and then PowerPoint and PDF presentations, to further explode into a plethora of multimedia opportunities. They should be used, however, sparingly. Choose a communication tool that supports your educational goal, not the other way around! Sometimes, the flashiness of a PowerPoint presentation is inversely proportional to its content.

Further, attention spans are getting shorter because social media and other communication channels compete for attention. The statistics teacher needs to be aware of that and work around it. I have been rather pleased, in various teaching environments, with web lectures and screencasts. Not only do these allow (the professor) to experiment with the so-called flipped classroom, but even when conventional teaching takes place, it is a rich additional resource that many students like.

Statistics teaching may encompass both very technical components, like mathematical theorems and proofs, model formulation, numerical analysis, software development and use, etc., and more applied elements like case studies, data analysis, the existence, nature, and consequences of a regulatory framework, etc. All of these should find their proper place in the classroom. The challenge for teachers is to be comfortable with all of these aspects. Both the differential geometry used in influence analysis and the clinical aspects of a trial in schizophrenia, for instance, should be given proper attention.

Do you have general advice for university or graduate school students who are interested in careers as statisticians in medical research?

The breadth of applied statistics in general, and medical statistics in particular, is fascinating. Mathematics, computer science, logical thinking flow together with pharmacological, biomedical, clinical, epidemiological, genetic, and related sciences. Looking back at our profession's history, we can identify a number of times at which science and our capabilities leapt forward, like rapids in a mountainous river. When this happens, it is tempting to think that the field is being entirely redefined or, as some think, replaced by another field.

An early example is the advent of epidemiology next to experimental research in agriculture and clinical trials. It looked for a while as if a totally new scientific field was emerging, but I believe what essentially happened is a broadening of focus and expansion of the toolkit; to put it simply, the then-dominating linear models were supplemented with logistic regression.

We somehow never seem good at absorbing such changes. Looking around the world, it is not totally clear whether epidemiology should be under the same organizational umbrella as (bio)statistics or a

department of its own. We see all sorts of organizational models.

The same thing happened when molecular genetics and bioinformatics came around the corner, and it is happening again in these days of Big Data, data science, etc.

The distinction between, for example, statistics and data science is a bit of a false one. Yes, new times bring new challenges that require flexibility and bring the need for adaptation. But this does not imply that we should let go of our foundations and principles, which are essentially geared toward extracting information from data in a trustworthy fashion, i.e., the art of separating signal from noise.

Today's, yesterday's, and tomorrow's students should be aware of this, recognize challenges but also opportunities, and be prepared to adapt but continue to make contributions, with flexibility to adjust to an ever-changing environment, whilst maintaining rigidity on scientific principles.

The Committee on International Relations in Statistics of the ASA, where you are the chair, has developed a unique educational program, the "Educational Ambassador Program." It aims to recruit a number of educational ambassadors from foreign countries to advance lasting collaborations between the ASA and other international statistical societies for permanent exchanges of knowledge. What is the rationale for this program and what is your lesson learned from the program?

The program was founded in 2005 by Martha Aliaga, ASA director of education at the time. Yearly, an educational ambassador is selected from a country around the globe. He or she attends the Joint Statistical Meetings, primarily takes a number of courses in the ASA Continuing Education program, and uses the so-acquired knowledge in teaching and other dissemination activities back home. Former Educational Ambassadors have come from Argentina, Ethiopia, Vietnam, Morocco, Armenia, Costa Rica, Botswana, Colombia, Bangladesh, and Nigeria.

A well-qualified Educational Ambassador who is supported by the local community can bring a lot to both the community and the ASA. Many of them maintain good contacts with colleagues they encounter during JSM, in particular the short-course instructors. In some cases, it has even meant that a short-course instructor has traveled to the ambassador's country and taught a course there.

What are challenges and opportunities for statisticians in medical research during the next decade?

Let me return to an example that I gave earlier. When observational data emerged, statisticians had to leave the comfort zone of rigidly designed experiments. Incorporating confounding and effect modification was the answer. The framework extended without becoming a completely separate one, even though some insist that this is what happened.

Another important example brings us back to the early days of AIDS clinical trials. Conventional clinical trial machinery and principles simply did not work: Patients refused to enroll in a trial and be denied access to newly emerging medication in other trials. As a result, co-enrollment became standard practice and … led to HAART, a highly active antiretrovirus therapy, aka cocktail therapy. The rigidity of academic, industrial, and regulatory procedures alike was successfully countered by influential patient communities. This is a very powerful and sobering lesson.

Currently, with the seemingly endless opportunities of individual data, clinical trials are evolving toward individual, rather than population-based, matters. Research in the field of personalized medicine, dynamic treatment allocation, etc., should be seen against this background. Such developments will require us to rethink, once more, how proper inferences will be drawn.

We need to make sure that we are ready and have the necessary equipment, flexibility, and arguing power to ensure that the larger research community understands the need for proper statistical inferences in this context.

About the Authors

Toshimitsu Hamasaki is chief of the Office of Biostatistics and Data Management at the National Cerebral and Cardiovascular Center in Japan. He has served as editor-in-chief of the *Journal of the Japanese Society of Computational Statistics* and is an associate editor the *Japanese Journal of Applied Statistics, Journal of the Japanese Society of Computational Statistics,* and *Statistics in Biopharmaceutical Research*. Hamasaki is an elected member of the nternational Statistical Institute and was awarded the Distinguished Article Award from the Japanese Society of Computational Statistics and Hida-Mizuno Prize from the Behaviormetric Society of Japan.

Scott Evans is a senior research scientist at Harvard University, where he is the director of the Statistical and Data Management Center for the Antibacterial Resistance Leadership Group and teaches clinical trials. He has received the Robert Zackin Distinguished Collaborative Statistician Award for significant statistical contributions to HIV research and a recognition award from the Harvard School of Public Health IRB. He is also a Fellow of the American Statistical Association. Evans is a member of an FDA advisory committee and has served on and chaired numerous data monitoring committees and scientific advisory committees. He is the past president of the ASA Boston Chapter, the ASA Development Committee, the ASA Teaching Statistics in the Health Sciences and Statistics in Sports sections; and member of the board of directors for Mu Sigma Rho, the National Honorary Society for Statistics.