CENSTAT IN 2008

universiteit hasselt

CENTRUM VOOR STATISTIEK

D/2009/2451/??

Content

PREFACE

Beyond any doubt was 2008 an exceptional year for the Center for Statistics. CenStat joined in celebrating the 35th anniversary of Universiteit Hasselt, and contributed to the festive mood by adding the 20th anniversary of its Master of Science in Biostatistics Program, as well as the first decennium of the Center for Statistics. At the same time, CenStat looked ahead. Following a proposal of the Rectors of Universiteit Hasselt and Katholieke Universiteit Leuven, the activities in biostatistics and statistical bioinformatics of both universities were placed under the umbrella of the International Institute for Biostatistics and statistical Bioinformatics (I-BioStat). Also the UHasselt activities in mathematical statistics find a home in I-BioStat. Both rectors signed the agreement on September 10, 2008.

To befittingly honor the initiative and to stylishly celebrate the occasion, the two universities awarded a joint honorary doctoral degree to Professor Norman Breslow of the University of Washington at Seattle. Professor Breslow holds worldwide recognition in the fields of biostatistics and epidemiology. He has been instrumental in promoting the use of statistics in large medical-epidemiological studies, through which to date numerous new clinical insights have been gained.

The new institute I-BioStat ensures optimal coordination between the research, educational, and consultancy activities. With over sixty collaborators, it is the largest academic group within its area, providing a much wider knowledge and expertise base than its constituent entities could do individually. This, in turn, dramatically increases the potential to generate added value.

Both entities will continue their existence as such. This implies, in particular, that CenStat's dedication to research, education, and consultancy stays unaltered. In 2008, important contributions were made in all three areas. The focus was invariably placed on biostatistics, bioinformatics, and mathematical statistics. Existing collaborations were honored and continued, while new were initiated. In 2008, young researchers joined the team and several CenStat members defended their doctoral thesis.

In 2007, the first year of the restructured Master of Statistics was rolled out. This was followed in 2008 by the second year of the overhauled program. To top it of, the program underwent a so-called Vl.I.R. visitation. On July 3, 2008, the final report was formally presented to Vl.I.R. by Professor Albert, president of the visiting committee. Taken together, the program was evaluated on 21 facets. The four specialties Applied Statistics, Biostatistics, Biostatistics ICP, and Bioinformatics enjoyed four times, sixed times, seven times, and five times, respectively, the qualification 'Excellent.' The corresponding figures for 'Good' are nine, nine, eight, and ten. `Excellent' implies that the program can be seen, for this particular aspect, as an international role model. `Good' signifies that the program is consistently of high quality on this aspect, without there being any deficiencies.

The visiting committee attributed to the program, befittingly in the light of the above, an overall positive judgment. The committee concluded that "the program reaches its objectives in an excellent way and stands as a unanimously recognized reference for an international Statistics program that is focused on training students originating from all over the world."

Of course, everything is open to improvement; in this spirit, the recommendations of the visiting committee were immediately taken to heart. For example, a curricular modification was prepared, implying that from the next academic year onward the specialty "Applied Statistics" will be called by its new name "Epidemiology & Public Health Methodology."

In agreement with CenStat's vision and to follow up on past engagements, and finally and definitely not in the least thanks to the Vl.I.R. sponsored Internationaal Cursus Programma Biostatistiek, headed by Professor Paul Janssen, a lot of focus continues to be placed on students from the South. In reverse, CenStat is expanding its already extensive network of contacts in the South, with initiatives ranging over Cuba, Kenya, Ethiopia, Madagascar, and Mozambique.

At home, CenStat remains an active partner in the Federal Government sponsored Interuniversitaire Attractiepool (IUAP) in statistics, jointly with the universities of Leuven, Gent, and Louvain-la-Neuve. The project further includes partners from Grenoble, Utrecht, London, and Santiago de Compostella. In collaboration with the University of Antwerp on the one hand and research institutes at UHasselt on the other, CenStat prides itself in successfully securing so-called Methusalem funding, designed to stably support senior researchers; funding is provided by the Flemish Government. In particular, CenStat is the UHasselt partner in a Methusalem consortium with the VAXINFECTIO Institute of Universiteit Antwerpen, headed by Professor Herman Goosens (UAntwerpen) and Professor Geert Molenberghs (UHasselt). A first "Joint Scientific Meeting of the Methusalem Consortium" took place on September 4, 2008, in Antwerp.

Taken together, it will be unequivocally clear that we can proudly and with great satisfaction look back, on a productive, successful, and pivotal year 2008.

In this document we invite the reader to sample a few of CenStat's research, educational, and consultancy initiatives. Needless to add the selection presented does not pretend completeness.

Professor Marc Aerts
Director of the Center for Statistics

Professor Geert Molenberghs
Director of the International Institute
for Biostatistics and statistical Bioinformatics

Policy and strategy

Collaborators

Faculty and Visiting Faculty

Prof. Dr. Marc Aerts, Prof. Dr. Roel Braekers, Prof. Dr. Tomasz Burzykowski, Prof. Dr. Christel Faes, Prof. Dr. Niel Hens, Prof. Dr. Paul Janssen, Prof. Dr. Geert Molenberghs, Prof. Dr. Ziv Shkedy, Prof. Dr. Noël Veraverbeke, Prof. Dr. Luc Bijnens, Prof. Dr. Marc Buyse, Prof. Dr. Luc Duchateau, Prof. Dr. Helena Geys, Prof. Dr. Emmanuel Lesaffre, Prof. Dr. Herbert Thijs, Prof. Dr. Ingrid Van Keilegom, Prof. Dr. Emmy Van Kerkhove, Prof. Dr. Geert Verbeke, Prof. Dr. Joost Weyler, Prof. Dr. Herman Callaert, Prof. Dr. Clarice Demétrio, Prof. Dr. Jeannine Houwing-Duistermaat

Senior Researchers, Postdoctoral Researchers and Visiting Scientists

Ariel Alonso Abad, Jose Cortinas Abrahantes, Caroline Beunckens, Liesbeth Bruckers, Rizopoulos Dimitris, Ivy Jansen, Saskia Litière, James Lindsey, Didier Renard, Dirk Valkenborg, Thierry Van Effelterre, Kristel Van Steen, Tony Vangeneugden, Luc Wouters

Junior Researchers (doctoral students and research assistants)

Pryseley Assam Nkouibert, Teshome Ayele Birhanu, Kaatje Bollaerts, Castro Sanchez Yovanna Amparo, An Creemers, Emanuele Del Fava, Auguste Gaddah, Arthur Gitome, Nele Goeyvaerts, Philippe Haldermans, Tom Jacobs, Tetyana Kadankova, Adetayo Kasim, Annouschka Laenen, Dan Lin, Girma Minalu, John Maringwa, Goele Massonnet, Harriet Namata, Setia Pramana, Jan Serroyen, Cristina Sotto, Abel Tilahun Eshete, Suzy Van Sanden, Kristien Wouters, Qi Zhu, Hans Bekaert, Cecile Goethals, Lies Provoost, Roel Straetemans, Bart Van Rompaey, Marc Vancaudenberg, Bart Winters

Administrative and Technical Staff

Kris Callaert, Annemie Hermans, Martine Machiels, Sabine Maldoy, Viviane Mebis, Conny Wijnants, Hilde Wijnen

Strategic Plan and Mission Statement

CenStat's research, both theoretically oriented as well as applied, receives international recognition. The same is true for its educational and statistical consultancy efforts. Let us focus on each of these pillars in turn, where after we will switch to key instances of such projects.

Education

Since decades, CenStat has held responsibility over the statistics education in academic curricula, for students in biomedical sciences, biology, and chemistry. At the same time, CenStat holds responsibility over statistical education for economy, physics, and mathematics. Furthermore, CenStat is pleased to offer contract education for biopharmaceutical industries, the government, and other research institutes. Whenever necessary, CenStat gladly collaborates with academic and industrial partners. The cherry on the pie is the Master in Statistics program with four principal branches: biostatistics, biostatistics/ICP, applied statistics, and bioinformatics.

Research

CenStat's research ranges over mathematical and various applied areas of statistics, with particular focus on biostatistics and bioinformatics.

For the mathematical statistics area, key research lines encompass:

- Bootstrap methodology
- Smoothing techniques
- Censored observations
- Asymptotics

For applied statistics, research interests embrace:

- Clustered observations
- Repeated measurements
- Frailty models for survival analysis
- Incomplete data
- Bioinformatics
- Infectious diseases
- Environment and health
- Psychiatry and mental health
- Teaching statistics
- Microbial risk assessment
- Clinical trials, including the evaluation of surrogate response

Statistical Consultancy

CenStat offers scientific statistical consultancy for academic partners, with emphasis on biomedical, biological, and other life science partners. An important part of the activity is geared towards the biopharmaceutical industry, while not neglecting other industrial partners. In addition, scientific collaboration takes place with local, regional, and federal governmental agencies. For one and a half decade, CenStat has been involved with the Belgian Health Interview Survey, organized since 1997 by the Scientific Institute of Public Health, after a few years of preparatory efforts. This is but one in a long list of cases. In sections to come, some of these will be highlighted.

International Dimension

CenStat's scientific research is internationally renowned, both in the theoretical as well as in applied statistical fields. The same holds for education and scientific consultancy. Let us expand on each of these three pillars.

CenStat's international dimension, in general and specifically focused on the master level training, is extensive.
The Advisory Board, active since the inception of the program, is internationally composed. The visiting faculty stems from Flemish, Belgian, and foreign universities. As Section 3.2.1.4 shows, the list of visiting faculty members is extensive and comes from a variety of highly reputable locations. Also, CenStat's own faculty has a stable and far reaching network of international contacts.

The international flavor of the MScBiostat was further strengthened when UHasselt obtained recognition from the former ABOS, currently VLIR-UOS, to organize a so-called VLIR-International Course Program (ICP).

Due to the international faculty and student mix, the natural working language for the program and its accompanying administration is English. This has allowed the UHasselt to attract a truly global student population, to have got Advisory Board members who genuinely belong to the absolute top in the field, and to employ renowned experts as visiting faculty, irrespective of their origin and native language. Moreover, biostatistics and other fields of applied statistics are essentially Anglo-Saxon in nature, with the largest contingents located in Anglo-Saxon countries, such as the United States and the United Kingdom. This implies a vast body of terminology and technical terms is hardly translated in other languages.

CenStat members and faculty of the program, including visiting faculty, have been active for decades in formal and informal outreach programs towards the rest of the world, with a strong emphasis on but not restricted to economically developing countries. A large number of CenStat members and master program faculty has held international functions. The international activities have largely contributed to a slate of honors and awards that have been bestowed upon CenStat members. In 2004, the Herman Callaert Leadership Award in Biostatistical Education and Dissemination was established. A large contingent of our master students conducts their master thesis in locations abroad. Finally, and very importantly, over the years students have enjoyed mobility support from the Erasmus program and its forerunner Socrates. Around 15 students from primarily Estonia and the Czech Republic, supplemented with others, have enjoyed such grants.

Apart from national sources, grant funding has been secured from such international organizations as the European Union and NATO's Scientific Program. CenStat members collaborate with biopharmaceutical and other companies around the globe, with the same holding true for academic department, research institutes, and governmental agencies.

Research

In this chapter, we zoom in on a selection of research and consultancy projects. We genuinely view these two activity areas, as part of CenStat's vision, and as beneficially belonging together.

Key Competence 1: Environment and Health

In the nineties, there was substantial scientific and public debate regarding the relationship between environment and health, fueled in part by the dioxin crises in food on the one hand, and as a result of waste incinerator exhaust on the other hand. In addition, there was concern regarding the adverse impact of cell phone network antennas, etc. More recently, the press reported on the relation between cadmium in the environment and an increased risk of lung cancer, on the adverse health effects of small particles, on increased PCB values in the eel population in Flemish aqueous bodies, etc.

The health risk associated with environmental contamination have been abundantly documented. However, there is no unanimity as to the precise quantification of such risks. This is unsurprising, given the tremendous complexity of the relationship between environment and health. First, toxic components can enter the human body system in a variety of ways: air-borne, through nutrition, or via direct skin contact. Second, the health effects of various pollutants are usually far from additive; they can reinforce or rather dampen each other's effects. Third, time enters into the equation as well; in the human body, the health effects of a number of toxic components may become noticeable after important time lags only; these include cancer or DNA damage.

Effects of environmental factors are hard, if not impossible, to isolate because they result from various causes combined. For example, think of someone who manipulates glue in a hobby-related contexts, eats of lot of their own catch from local rivers, is exposed to toxic agents in the workplace, and/or smokes actively or passively. All of such factors, in addition to other environmental exposures shared with other members of the geographically closed population, can have negative effects on one's health.

To further insight in the complex relationship between environmental exposures and health, and address the public's concern, the Flemish government has been setting up large scientific projects, in which the Center for Statistics takes an active part.

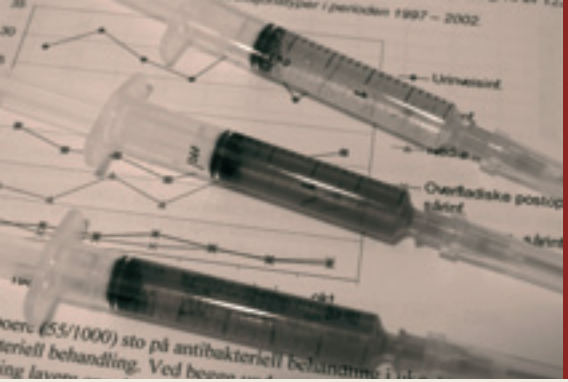
►► BONK: BevolkingsOnderzoek Noorder-Kempen (Population Research Noorder-Kempen)

The zinc factories in Balen, Overpelt, and Lommel have entered considerable amounts of cadmium into the environment, from the end of the 19th century until the seventies of the 20th century. Environmental regulations were not as strict in the past as they presently are. In the 1970s, industry switched to so-called electrolytic technology for zinc production. However, cadmium stayed behind in the soil and became a permanent source of contamination. An important exposure route is, not surprisingly, private-garden vegetable consumption. In addition, the effect of domestic dust inhalation ought not to be underestimated.

Research conducted in the 1990s establishes a relationship between the risk of lung cancer and cadmium exposure. Following up on these result, the Ministers responsible for Public Health and the Environment, Vervotte and Peeters, respectively, have initiated the "Action Plan Cadmium." One of the points of action is the design and conduct of a large population-based study, to establish the risk in the current populace of the area.

►► Project "Environment and Health"

In 2001, the Flemish Government established a so-called bio-monitoring network across Flanders. The combined exposure to pollutants is the sum of various exposures at home, in traffic, in the workshop, in open air, via food and drinking water intake, etc. Measurements at industrial chimneys alone are, therefore, falling short in quantifying the total exposure. Only measurements in human subjects, e.g., in blood, urine, hair, etc., i.e., bio-monitoring, has the potential to provide a reasonably accurate quantification of the various exposure routes.
Practically, three types of biomarkers are studied. First, exposure markers capture the presence of toxicants in the human body. Second, there are *biomarkers* measuring health effects. Principal attention is devoted to effect markers for growth and development, asthma, allergy, and effects that relate to development of cancer. Third, there are *markers for individual sensibility*. Indeed, the uptake or capacity for chemical compounds, or the potential to metabolize them, is individual-specific, owing to genetic factors, for example.
The study encompassed various age groups: babies, fifteen year old adolescents, and adults in the age range of 50-65 years. In a perinatal setting, researchers scrutinize exposure at the onset of life. For a number of fat-soluble, persistent compounds, infants may suffer from important exposures, in the prenatal and, via breast feeding, early postnatal phases. Such compounds result from cumulative exposures in the mother.
The phase of puberty is key because, by definition, essential processes take place in the body and the secondary sex characteristics (pubic and armpit hair, breast development) become discernible. It is important to study which chemical compounds



have an adverse effect on such aspects of development and maturation. In the younger age groups, the emphasis is on contemporary exposure. The adult group offers routes of information on the cumulative presence of pollutants in the human body resulting from decades of exposures. The first bio-monitoring campaign in Flanders (2002-2006) revealed that one's region of residence has a measurable impact on the pollutants in the body and their ensuing biological effects. The current program (2007-2011) is geared towards the study of point-source exposure. These include waste incinerators. It is thereby important to ascertain the impact of one's geographical proximity to such a source. Furthermore, the project wishes to study the effects of life style, age group, and socio-economic variables.

►► Genk-Zuid

"The environmental conditions at and near the industrial pole Genk-Zuid is poor." This was the conclusion of the department for Public Health Supervision, following on an assessment of the health risks based on measurements of the Flemish Environmental Societey (*Vlaamse Milieumaatschappij*, VMM). Concern relative to health is primarily focusing chrome (Cr) and nickel (Ni) in fine particles (PM10). The public shows a genuine concern regarding health effects. For this reason, last year the steering committee "Environmental Quality Genk-Zuid" was formed. One of their assignments is the conduct of a well-targeted Health Interview Survey that has the potential, apart from mapping the living and life-quality conditions, to picture environmental burden experienced by the local population. The study will undoubtedly provide useful information about the further steps to be taken by the government.

►► What is the Statistician's Role in These Studies?

Even though the precise research questions in the studies described are different, the substantive researchers have similar statistical and methodological questions.

► How many study subjects ought to be included?

The cost of large-scale empirical environmental studies is considerable. To estimate the amount of pollutants in the human subject, a number of blood and urine samples are collected. This renders necessary, in turn, the use of sophisticated and hence expensive laboratory analyses. It is therefore of the utmost importance to pre-determine the number of study subjects to be included, so as to reach conclusions with a sufficient amount of precision. At the same time, overly large sample sizes would be unethical and expensive, and would decelerate the study.

How ought participants to be selected?

A study's results are useless if they cannot be generalized from the sample to a target population of interest. For this reason, it is imperative to properly design a study, rather than leaving the choice of study subjects to chance. For example, it is unacceptable to work with spontaneous volunteers; these may have an interested in health in the first place, lead a healthier life, and come from more well-off social classes. Conversely, people with documented health problems may have a personal interest to volunteer. Evidently, all of this can bias the results in non-negligible ways. Summarizing, the establishment of a state-of-the-art sampling plan is crucial.

How to capture information?

Many studies are based on lengthy questionnaires. The relation between biomarkers and effect markers is generally more complex. A variety of factors influences this relationship. For example, smoking has an impact on the amount of cadmium accumulated in the body and hence on the chance of getting lung cancer. Therefore, the study between environmental cadmium exposure and lung cancer cannot be undertaken without factoring smoking into the equations. This explains why questionnaires ought to collect a large amount of auxiliary information, including food, life style, and working environment information. To ensure that such data are trustworthy, questionnaires need to be constructed following the highest professional standards. One of the golden rules is to restrict the collection of information to what shall be analyzed later.

How to analyze the data collected?

The ultimate goal is to construct models that provide an adequate explanation of the observed differences and effects. A paramount example of such a question is: *To what extent can the level of pesticides in the body be explained from the region of residence, the respondent's food habits, occupational setup, etc.?*
Such models are used, not only to explain but also to predict. For example: *What is the risk for a child to be allergic, given exposure to the environment and through nutrition?*

For each one of these question, the most appropriate method of analysis needs to be selected, so as to map out a potentially complex network of inter-relationships. The statistician is responsible for checking whether the boundary conditions are satisfied to reliably apply a particular method. In case such conditions would not be satisfied, it is important to assess how much deviations may affect the conclusions.
Neglecting one aspect when designing a study and/or collecting and analyzing data may render a study essentially useless and invalidate conclusions based there upon. This explains why statistics plays an essential part in large-scale policy-related surveys, similarly to the statistician's role in, say, clinical trials.

Key Competence 2: The World of Infectious Diseases

Did you know that approximately 80% of all people was infected at some point with the so-called *human papiloma virus* (HPV) which in women can lead to cervical cancer? Or did you know that annually still as much as 450,000 children die from measles? Salmonella is one of the most prevalent food infections worldwide and global estimates of the number of cases vary from 14 to 120 per 100,000 people.

The historic and epidemiological literatures abound of outbreak descriptions of a variety of infectious diseases, such as the pest epidemic in London in 1665-1666 and the Spanish and Asian flue epidemics in 1918-1919 and 1957, respectively. These and other epidemics led, and still lead, to millions of deaths. In setting intervention policies against imminent infectious diseases, such as SARS, and a potentially human variant of avian influenza, mathematical modeling of the spread dynamics of the disease plays a crucial role. Thanks to such mathematical modeling, a variety of outbreak scenarios can be simulated and the effects quantified.

Consider, for example, one of the most typical childhood diseases: e.g., chicken-pox, caused by the so-called *varicella zoster* virus. The virus can be passed on to the human by coughing or body contact, via inhalation. The incubation period is between 14 and 21 days. After a prodromal phase, characterized by mild fever and cold-like systems, usually taken a few days, the first red spots appear, which transform after roughly a day to small, liquid-filled bulbs set against a red background. They invade the entire body, often starting in the neck area. Typically, such bulbs are visible on the hair-covered part of the head. They can vary from less than 10 to several hundreds. After roughly a week they will have disappeared for the better part. The illness is extremely contagious and most people, more than 90%, attract it during childhood. The disease is dangerous only in people with a poorly functioning immunity system. In babies younger than about 6 months, chicken-pox often takes a very mild form, owing to the antibodies passed on from the mother. It is better to go through the disease during childhood, because the adult form generally is more severe.

A population is made up of a large number of individuals that can be portioned into compartments for any given infectious disease, such as, for example, the *varicella zoster virus*. One typically distinguishes between: *susceptible individuals* that have not yet had the disease; *latently infected individuals,* i.e., these without disease symptoms that either will become ill or appear to be recently cured; *infectious individuals,* that are contagious; and *vaccinated individuals,* that can no longer attract the disease.

The basic model, consisting of three compartments: susceptible – contagious and ill – cured, is represented in Figure 1. Further extensions of the basic model take into account, for example: *maternal immunity*: the natural protection provided by the antibodies passed on from the mother onto the child; the *latency period:* and *becoming susceptible again* for the disease. Such

models are almost out of necessity a simplification of the real disease dynamics. This notwithstanding, they routinely provide a very useful paradigm to represent the spread of a disease in a given population.

Figure 1: The basic model 'susceptible – contagious and ill – cured,' with λ representing the force of infection and α the cure rate.

One of the crucial parameters in the compartmental models is the so-called **force of infection** λ that represents the transition from "susceptible" to "contagious/infected." The parameter expresses the instantaneous chance to become infected, given that one is still fully at risk of getting the disease. From the analysis of empirical data it follows that the force of infection is age and time dependent. When the infection is in a state of equilibrium, the force of infection can be estimated based upon the presence of antibodies in the blood. In Figure 2, the age-dependent force of infection for the *varicella zoster virus* is depicted. Clearly, the force of infection is at its peak in 6-7 year olds, i.e., the age at which children enter primary school and hence are exposed to large numbers of potentially infective contacts with peers. A second maximum is discernible in the group of 20-27 year olds, i.e., the age at which parents attract the disease from their children who in turn got exposed in early grades of primary school.

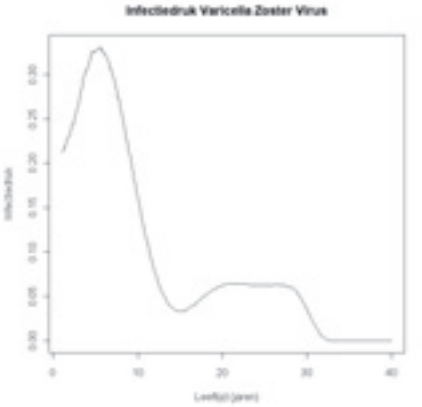


Figure 2: The time-dependent force of infection for the varicella zoster virus, following a flexible logistic regression model fitted to Belgian serological data.

The transmission of an infectious disease is gradually determined by, on the one hand, the intensity of contacts with individuals and, on the other hand, the infectiousness of the pathogen. Thanks to a recent study, contacted in collaboration with the Center for the Evaluation of Vaccination at Antwerp University, within the framework of the European "POLYMOD" project, data are available regarding the contact pattern in humans. These data allow for the estimation of *infectivity* of diseases for which transmission takes place via air, such as the *varicella zoster virus*. Figure 3 displays the intensity of the daily mean number of contacts between people of varying ages. Clearly, there is a high intensity of contacts between age-peers and, to a somewhat lesser extent, between grandparents and grandchildren.

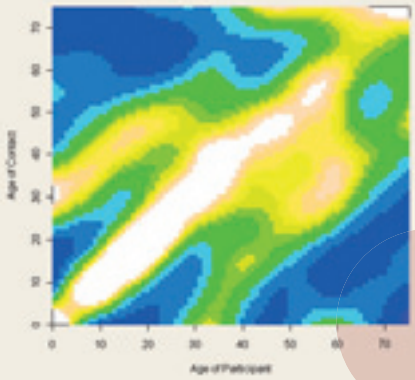


Figure 3: The intensity of the daily average number of contacts between people of various ages, based on a Belgian contact survey. Lighter colors correspond to higher intensity of contact.

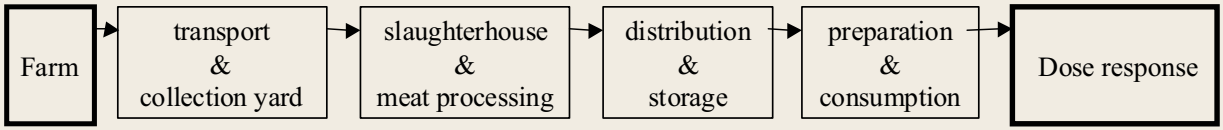
Thanks to the availability of such data and the application of advanced statistical methods, the effects of vaccination can be charted. It follows that the beneficial effects of childhood vaccination can be tremendously beneficial. At the same time, caution is needed since such effects do not hold for each and every infectious disease. Indeed, every infection has got its own specific characteristics, which have to be taken into account when modeling its spread. Think, for example, of specific risk groups such as intravenous drug users when modeling the spread of Hepatitis B.

►► Zoonoses

Some infectious diseases occur in both humans and animals. In some cases, transitions take place exclusively from humans onto animals, while for others it is exactly the other way around. A zoonosis is an infectious disease that allows transmission between humans and animals. BSE, the now well-known abbreviation for *bovine spongiform encephalopathy,* and better known still as "mad cow disease," is a paradigm in this respect. It is a cattle disease, caused by so-called prions. It is generally accepted that the disease can be transmitted to humans, thereby giving rise to a new version of Creutzfeldt-Jacob's Disease (NvCJD). Confrontation with this disease is highly unlikely and hence a large-scale epidemic is even less likely. However, there are other zoonoses for which higher levels of concern are warranted. One such case is salmonella. Salmonella is transmitted, first and foremost, via food (eggs, diary products, poultry). It is one of the most frequent food infections worldwide and global estimates of the number of cases range from 14 to 120 per 100,000 people.

Since 1997, the FAVV is responsible for monitoring the food chain for the presence of Salmonella. The World Health Organization issued guidelines and principles regarding standardized risk assessment for the contraction of food-related infections. Following these guidelines, such an assessment ought to be based on scientific standards. The Federal Government sponsored the contract research project METZOON ("Methodological Development for Quantitative Risk Assessment of Zoonoses in Belgium, applied to *Salmonella* in pigs") is conducted within the context of a consortium, encompassing the Center of Statistics of Hasselt University, the Veterinary Faculties of the Universities of Liège and Ghent, the Institute for Agriculture and Fishery, in Melle, the Center for Research in Veterinary Sciences and Agro-chemistry (CODA, Brussels), and the Scientific Institute of Public Health (Brussels). The project envisages the evaluation and improvement of existing methods for quantitative risk assessment of zoonoses in the food chain. Quality criteria play a pivotal role in this respect. It is one of the project's main goals to develop methodology to minimize uncertainty in risk assessment and to optimally provide information regarding the variability and hence uncertainty in the risk in the target population.

CenStat's role targets evaluation and optimization of the mathematical and statistical models, employed in the various stages of risk analysis. The research project encompasses all stages of food production, which implies that both the animal and human side receives attention, as sketched in Figure 4.



Figuur 4: Stages of the *Quantitative Microbial Risk Assessment* (QMRA) model for description of Salmonella Typhimurium contamination of pork meat in Belgium.

Within the METZOON project, the following sub-projects have been realized, among others: (1) the estimation of the dose-response relationship based on the most valid outbreak data; (2) estimation of the Salmonella serological prevalence, with correction for test-misclassification based on a so-called Bayesian analysis; (3) identification of companies at risk for Salmonella, using advanced regression techniques; (4) identification of risk factors associated with such companies; (5) based upon this type of factors, so-called "what if" evaluations of the risk analysis model are conducted; (6) the risk evaluation model has been optimized for the Belgian situation, using historical data and expert opinion, in conjunction with new data and further expert views collected within the framework of the METZOON project.

►► Modelling the Spatial Distribution of Infectious Diseases

Yearly, infectious diseases in animals cause tremendous economic losses. A paradigm in this respect is the sequence of outbreaks of foot and mouth diseases, avian influenza, and bluetongue in Belgian companies. During the outbreak, attempts are made to stop or at least slow the spread of such infectious diseases through such control measures as preventing the animals from ranging outdoors, the reduction of blockage of animal transportation between companies, the complete clearance of companies and the vaccination of animals.

Bluetongue is an infectious disease in sheep and cows. One of the symptoms to be found in infected animals is, perhaps not surprisingly, a swollen and blue tongue. The illness cannot be transferred to humans, but does cause important economic loss. Roughly 30 to 40 percent of the animals dies from the disease. The bluetongue virus is transmitted by certain types of mosquitoes, *Cullicoides,* the so-called vector. The mosquito transmits the disease by stinging an affected and thereafter an unaffected animal. Evidently, the infection's territory is tightly linked to that of the mosquito.
Until right before 2006, the disease was found exclusively in Southern Europe, in such countries as Spain, Greece, and Italy. However, owing to global warming, the mosquito's territory is expanding. In August 2006, the first outbreak of bluetongue in

the Netherlands became a fact. Thereafter, the virus spread extremely quickly, with outbreaks in Belgium and Germany within the time span of a week. In December 2006, the virus was spread throughout the country. Figure 5 displays the companies where bluetongue was present, until December 2006. During the outbreaks, control measures were frequently imposed to contain the outbreak, to the extent possible. One such measure is the establishment of a perimeter around infected companies; transportation of animals from within such a perimeter is prohibited. It was further advised to maintain animals inside their stables within such a perimeter, and to appropriately employ insecticides. Unfortunately, the set of measures taken was unable to prevent the infection from further spreading.

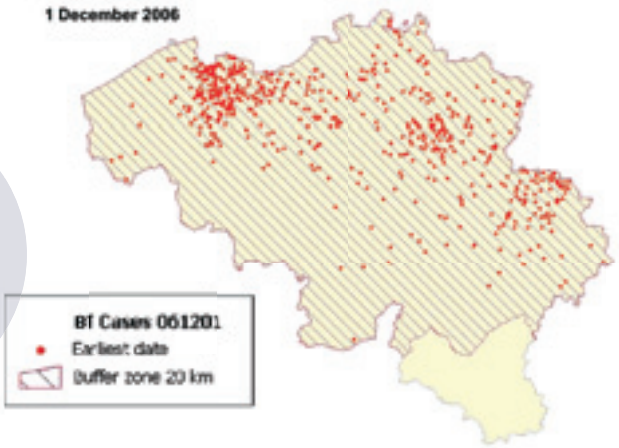
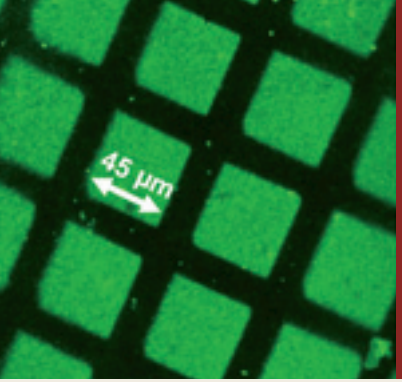


Figure 5: The spread of bluetonge in Belgium, until December 1, 2006.

One of the plausible hypotheses was that the zone established by the perimeter around an infected company is too small to prevent the diseases from spreading. An alternative hypothesis is that the mosquitoes can travel long distances via wind or transportation of animals or other freight. Finally, it is possible that the protective measures were taken too late and that the

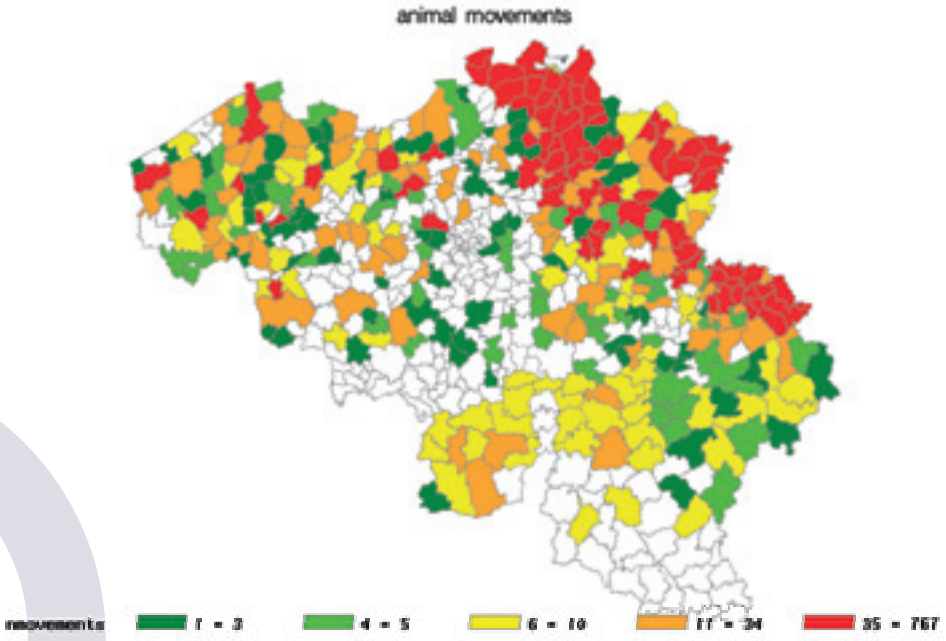


spread by that time already had become a fact. Such hypotheses need to be put to the test. The European Food Safety Agency, EFSA, set up a global epidemiological analysis, in which CenStat participated, in collaboration with the Center for Research in Agriculture and Agro-Chemistry, CODA.

Such analyses necessitate the use of so-called *spatio-temporal models*. This modeling strategy aims at mapping the geographical spread of a disease in a time-dependent fashion. Spatial correlation is taken into account, i.e., the likely possibility of nearby areas to exhibit larger correlation than areas further apart from each other is allowed for in the modeling exercise. Indeed, it is not unlikely for the spread of mosquitoes to be spatially heterogeneous, stemming from, for example, environmental factors such as altitude, temperature, and the ecosystem. To obtain an idea about the geographical spread of the disease, one can map the number of outbreaks, e.g., through the number of companies per town with presence of bluetongue. Of course, the overall number of companies needs to be factored into the equation as well. Otherwise, larger towns would by default be considered at higher risk than smaller town, which would lead to biased results. Furthermore, the spatial density of companies plays a role, too. Figure 6 displays the bluetongue risk during the 2006 outbreak, and results from such a spatio-temporal model.

Figure 6: Risk for a company to be infected during the 2006 outbreak. Red areas are more at risk, relative to the green areas.

The temperature, altitude, dominant wind direction, and animal transport are important factors that influence the further spread of the disease. The number of animals transported out of the bluetongue-infected area are depicted in Figure 7. CenStat researchers were able to show that the number of animals transported constituted an increased risk for the disease's spread.



Figuur 7. Number of animals transported out of the area of first infection, during the outbreak.

Until now, there is no clarity as to the control measures that ought to be taken to prevent the disease from spreading. CenStat will tackle this important problem in the future, starting from spatio-temporal models, designed to describe the disease's dynamics.

Key Competence 3: Bioinformatics

Introduction

A lot of current research in biology and medicine is aimed at understanding, what is the role of particular fragments of a living's organism genome from a point of view of different biological process taking place in a cell? In this context, the word "bioinformatics" is often used. What is bioinformatics? Different people define it differently. Very often, the "informatics" component is linked to computer science. In this interpretation, bioinformatics means something like "computer science techniques used to support gene research". However, this is a very limited view, as can be seen from the following definition formulated by The NIH Biomedical Information Science and Technology Initiative Consortium (http://www.bisti.nih.gov/CompuBioDef.pdf): *"Bioinformatics: Research, development, or application of computational tools and approaches for expanding the use of biological, medical, behavioral or health data, including those to acquire, store, organize, archive, analyze, or visualize such data."* In this view bioinformatics is really about extracting information. Thus, the "informatics" component should be rather linked to "information". Consequently, bioinformatics should not be seen as a discipline limited only to computer science techniques, but rather as an interdisciplinary science that includes methods allowing to obtain information about genes and their role. And when it comes to getting information from data, statistics enters the picture.

Center for Statistics and Bioinformatics

Thanks to its extensive expertise in statistics, Center for Statistics (CenStat) is in a good position to conduct research, provide high-quality scientific consulting services, and offer education in statistical methodology relevant for bioinformatics.

The group of CenStat researchers involved in the bioinformatics-related activities has been steadily growing over the past few years, thanks to the tUL and Hasselt University BOF funding, as well as to industry-sponsored grants. At this moment the group includes two permanent staff members (Tomasz Burzykowski and Ziv Shkedy), one post-doc (Ivy Jansen), and six doctoral researchers and assistants (Philippe Haldermans, Adetayo Kassim, Dan Lin, Dirk Valkenborg, Suzy Van Sanden, Qi Zhu). On occasional basis other CenStat members contribute to the group's activities as well.

Potential applications of statistical methodology in *genomics* (the study of genes and their function) and *proteomics* (the study of all the proteins expressed by a genome) are innumerable. To maintain high quality of research and services, even at large statistical centers some selection of application areas is inevitable. At Center for Statistics we concentrate on the tools

for the analysis of *gene*- and *protein expression*. In particular, we mainly focus on the analysis of microarray data and protein mass spectrometry data.

Microarrays

We mainly focus on the spotted cDNA arrays. In a standard cDNA microarray experiment (see Figure 1), mRNA from two samples (1) is reverse-transcribed to single-stranded DNA and differentially labelled with green and red fluorescent dyes (2). The obtained samples of single-stranded DNA, often referred to as targets, are mixed together and poured over a glass slide that contains spots with known single-stranded DNA fragments (probes). Hybridization takes place between the target and the complementary probe sequences (3). The intensity of the fluorescent signals at each spot is measured (4) and provides information about the expression level of the genes: the ratio of the (green and red) signal intensities approximates the ratio of the mRNA concentrations in the two samples.

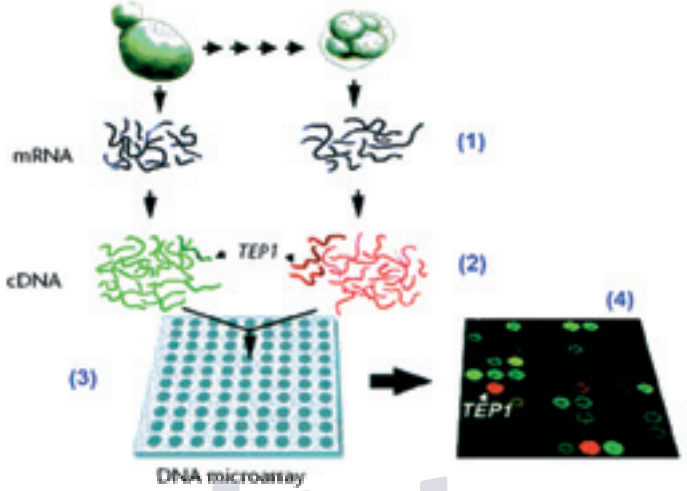


Figure 1: cDNA microarray technology

Pre-processing of microarray data

The ratio the (green and red) signal intensities is influenced by systematic effects from different sources that can introduce biases. These effects should be removed prior to undertaking an analysis of the obtained measurements.

The presence of such effects can be detected in ratio-by-intensity (RI) plots, sometimes also called MA-plots. The plots result from plotting the difference between the log intensity readings of the two dyes versus their mean log-value. Assuming that only a small number of genes are in fact differentially expressed, we should see on average a horizontal curve. If this not the case, some curvature might need to be removed.

At CenStat we have developed several methods for removing systematic effects from microarray data. Figure 2 illustrates the RI-plots for five cDNA microarrays before (left column) and after (right column) applying the transformation. The curvature of the plots seen in the left-hand side plots is clearly removed by the transformation. This indicates that the curvature might have been caused by a constant difference in the fluorescence measurements for the two dyes.

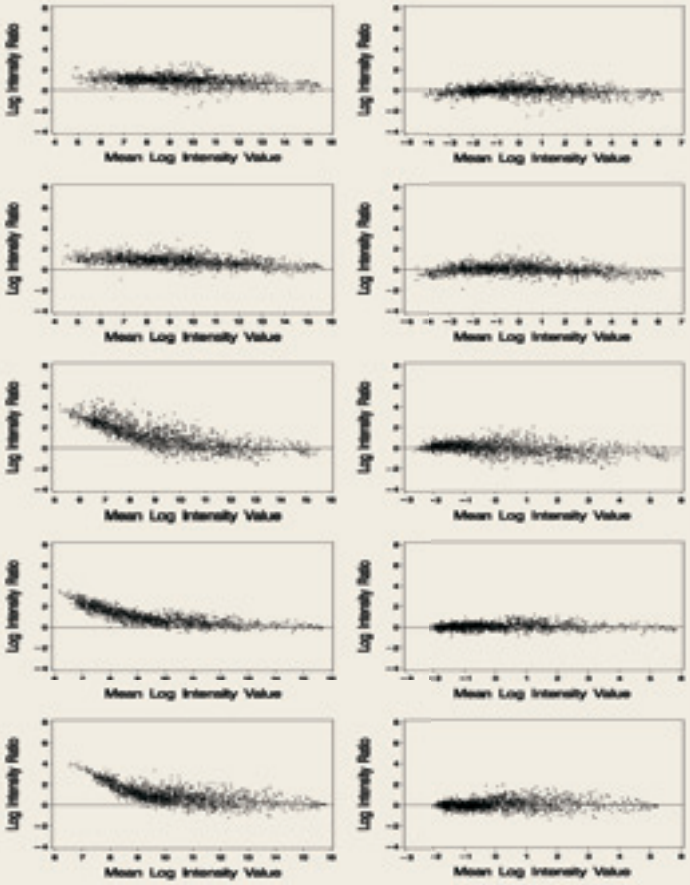


Figure 2: Normalized cDNA microarray data

An alternative, very flexible approach to the normalization of microarray data, has been proposed as well. It is based on the use of linear mixed effects models.

Analysis of microarray data

Properly pre-processed micrarray data can be used for various aims.

For instance, one can use microarrays for diagnostic purposes, i.e., to classify samples into several groups (e.g., normal/diseased). Many statistical classification methods are available, but their application to the microarray data encounters the problem of having many more variables (genes) than observations (arrays). Which method might be most suitable for such a situation? At CenStat, reseachers designed an extensive simulation study aimed at answering this question. From the results it follows that classical methods like diagonal linear discrimination analysis can perform equally well as more complex approaches like random forests or support vector machines.

Microarrays are more and more often used in drug development. In this context the investigation of a dose-response relationship is of primary interest, where the response is the gene expression of a subject at a certain dose level. In the microarray context, the application of classical dose-response analysis methods is not straightforward, as the analysis needs to be performed for (tens of) thousands of genes. In a series of papers, CenStat researchers have investigated many aspects of the analysis of microarray dose-response studies. In particular, these have proposed and compared several methods, that take into account the problem of multiple genes.

One of the potential uses of microarrays in drug development is to find genomic biomarkers, i.e., genes, whose expression level might serve as an indicator of, e.g., response to treatment. A group of CenStat researchers is investigating practical methods to find and evaluate candidate genomic biomarkers.

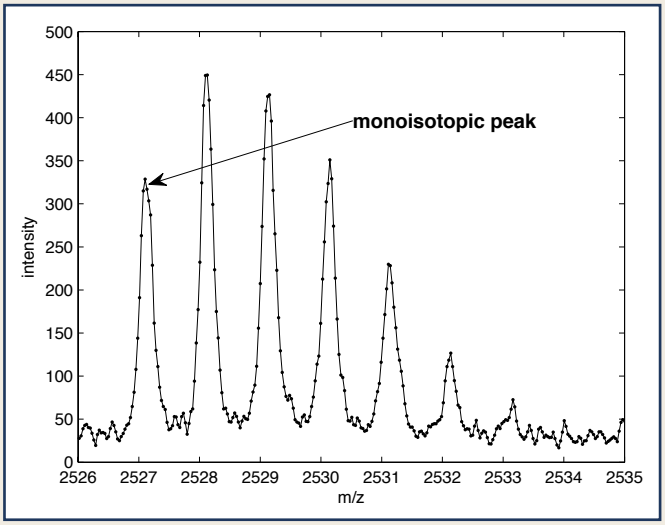
Last but not least, an important issue is that expression levels of genes are most likely dependent due to complex co-regulation processes. An analysis of the dependence might give some information about biological processes (pathways) that are operating in a particular biological condition. However, to this aim one should simultaneously analyze expression level of many (or all) genes. Research in this direction has been initiated by CenStat researchers, who, in collaboration with colleagues from the Theoretical Computer Science Group, has been investigating the use of statistical methods in an algorithm aimed at finding positive and negative co-regulated gene clusters in microarray data.

Protein mass spectrometry

Analysis of protein content of a sample is most commonly performed by mass spectrometry (MS). Protein mixtures are extracted from, e.g., cells, and digested with a specific protease, e.g., trypsin. In a commonly used approach, termed MALDI-MS, the resulting peptide mixture is co-crystallized in a surplus of matrix molecules on a solid support. After introduction into the mass spectrometer, peptides are evaporated and positively charged by Matrix-Assisted Laser Desorption-Ionisation (MALDI). After acceleration in an electric field, the mass/charge (m/z) ratio of the peptide ions is analysed by measurement of the Time-Of-Flight (TOF) in a field-free flight tube.

As a result of processing a peptide mixture by MS, a spectrum is obtained, i.e., a list of measurements of the intensity of particles of particular masses arriving at the detector. Figure 3 shows an example of a spectrum.

Figure 3: An example of (a part of) a mass spectrum



On the vertical axis the registered intensity of particles arriving at the detector for a particular mass (or mass to charge ratio, m/z, shown on the horizontal axis) is given. The peaks indicate an increased detection of particles that may be due to a presence of a peptide in a sample. The heights of the peaks are supposed to reflect the relative abundance of peptides with a particular mass in the mixture.

Intensity measurements contain systematic and random effects that may include, e.g., baseline effect, chemical noise (peaks from compounds such as matrix and solvent), and random (measurement) error. The systematic effects need to be removed. Most statistical methods proposed for processing mass spectra use data-driven smoothing techniques based on, e.g., wavelets. These methods suffer from several drawbacks. In particular, they are not capable of distinguishing true signal from that generated by, e.g., chemical noise.

At CenStat, an alternative approach, based on searching for particular features of the spectrum peaks, that would reflect biochemical properties of peptides, has been investigated. This approach allows to distinguish intensity measurements generated by peptides from those generated by other chemical particles or by random noise. CenStat researchers have implemented this method of peak finding and evaluated its suitability for mass spectra obtained by a combined fractional diagonal chromatography technique. Moreover, an improvement to the method has been developed that allows for a higher sensitivity of detection of peptides that, e.g., contain sulphur atoms. In principle, the method might be used to detect other types of modifications of a peptide like phosphorylation, or glycoprotein. Such modifications are of much interest to biologists. This is a topic of current research.

Another interesting research line focuses on so-called doubly-labelled mass spectra. The experimental technique resembles that of the cDNA microarrays. In particular, peptides are extracted from two samples, labelled with oxygen atoms with atomic mass 16 and 18, respectively, and then simultaneously analyzed by a mass spectrometer. As a result, in the single spectrum the spectra from the two samples should (ideally) be shifted by four m/z units. An analysis of the ratio between the intensities observed for the two spectra should give an idea about the relative abundance of peptides in the two samples. A method for analyzing such data is currently investigated.

Collaboration

The aforementioned research activities were initiated or undertaken as a result of a broadening collaboration between CenStat and other scientific and industry partners in the domain of genomics and proteomics. Within the Hasselt University our collaborators are colleagues at the Centrum of Environmental Sciences, at the Biomedical Research Institute, and at the Theoretical Computer Science Group. Within tUL, a collaboration with the Department of Health Risk and Toxicology of the Maastricht University is worth mentioning.

An important source of data and research topics are CenStat's scientific consultancy services. In this respect, the collaboration with, e.g., Johnson & Johnson Research & Development at Beerse, and with Pronota in Ghent, is very instrumental and rewarding.

CenStat's activity in the domain of bioinformatics is not limited to the scientific research and consultancy, though. An important component is education. In cooperation with the Theoretical Computer Science Group, CenStat contributed to the establishment of the tUL's Bachelor in Biomedical Sciences. As of the next year, within the newly organized two-year Master of Statistics program, a possibility to obtain a Master of Bioinformatics will be offered. We aim at providing the students with the necessary foundations of molecular biology and specialized knowledge and applied skills in database management, programming, statistical techniques, and knowledge discovery and integration, as required in the rapidly evolving research fields in genomics and proteomics.

Key Competence 4: The Watch Dog of the Clinical Trial

It is no exception for 10 to 15 years to elapse before a potentially effective molecule makes it to the pharmacy's shelves, if ever. Indeed, the development of a medicinal product is a complicated and hence slow process. Upon terminating the *in vitro* phase, experiments in living organisms begin. In the initial *in vivo* experiments, animal models are employed and then, ultimately, the leap to the human model is made. One starts in very small groups of volunteers and then gradually builds up towards large, so-called conclusive or *pivotal* clinical trials. Based on these, the decision is taken whether or not the project can and will be entered into the market. However, is it even allowed to experiment in humans? Is safety considered and, if so, by whom?

As a general rule, medical experiments in humans are prohibited, as stipulated by the universally adopted Helsinki Convention, in response to the horrendous experiments in the Nazi concentration camps. Thence, a clinical trial is, strictly speaking, always an exception to the convention. No wonder that it is bound by strict and severe rules, regulations, and restrictions.

Before initiating a clinical trial, a protocol is written, detailing every aspect of the study: Who receives the standard treatment and who will be allocated to the experimental group? How will the products be administered and by whom? What is being measured and at which measurement occasions? Which clinics and what specialized personnel is allowed to partake? How will study staff be trained? Which statistical analyses will be conducted and by whom? Merely by itself, the protocol already is a rather voluminous document, but it is then supplemented with a so-called *case report form* (CRF) for each and every participating volunteer. Virtually without exception, such a CRF is substantial in size. This allows for easy and detailed control, afterwards, not only by the sponsoring company, but also by the regulatory authority. In the United States, the latter is the *Food and Drug Administration* (FDA). In our part of the world, it is fair to say that some work in terms of refining and harmonizing the regulatory framework needs to be done. Indeed, presently, every EU member state has got its own administration, topped of with a umbrella organization, EMEA, the European Medicine's Agency. Furthermore, Japan has a well-developed regulatory framework and accompanying administration. Quite a few other countries will strongly rely on FDA's, EMEA's, and Japan's opinion.

The regulatory authority has two very important roles, at the beginning and at the end of a clinical trial, respectively. Before the study is initiated, the authority scrutinizes the protocol, invites the sponsoring company to hearings and, finally, perhaps upon amending the protocol, issues permission to start the study. At the end of the trial, the regulator receives all data, including statistical analyses conducted, but also, for example, the *case report forms*. In bygone days, this easily meant

truckloads full of documents. Currently, a few DVD boxes suffice. The regulatory administration studies the evidence provided, will oftentimes re-analyze the data, and/or supplement these done by the sponsor's statisticians with their own. Based here upon, the company will be heard. Upon completion of this process, a 'go/no-go' decision will be taken. Occasionally, the signal will neither be green nor red, but yellow: the decision is postponed; the biopharmaceutical version of a second chance exam, as it were.

We referred to *beginning* and *end* of a study. Needless to say that checks and balances are in place *during* a study's conduct, as well. The company is interested in knowing whether their product works, and does so better than the competitors' products. Next to this, the most important control question is whether the product is *safe*: is the amount and severity of side effects under control, in particular these that can be ascribed to the medication administered. Generally and not surprisingly, this is a very technical-biomedical matter, at the same time calling for data analysis. These control endeavors are assigned to what used to be called a DSMB or *Data Safety and Monitoring Board*. Presently, one rather refers to the *Independent Data Monitoring Committee* (IDMC). Still, the emphasis is placed on the product's safety but at the same time a much broader view is taken: the entire data flow is investigated. The word *"independent,"* in particular, is crucial. An IDMC consists of experts with biopharmaceutical and statistical competences, with profound knowledge in the therapeutic and data-analytic areas of relevance (one does not select ophthalmologists in a gastro-intestinal-cancer-related IDMC), but without direct or indirect links with the company or the study. Indeed, someone directly involved in the treatment of study subjects cannot be expected to be fully impartial, in spite of good intentions. The same is true for share holders of a company sponsoring a study.

Various CenStat members are experienced with IDMCs. Indeed, precisely to avoid unwarranted ties, one will often recruit medical and statistical personnel from academe, rather than from the industrial world. In the past, an IDMC virtually invariably implied an intercontinental trip of a few days, within which one would lock oneself up during 5-10 hours in a board room, tucked away in a large airport in Washington, DC or New York City. For large clinical trials, these could easily spill over into multi-day meetings. Thanks to the state of contemporary communication technology, such meetings frequently revert to tele- and video-conferencing facilities.

Key Competence 5: Repeated and Incomplete Data

An hypertensive patient requires a physicians' check-up at regular intervals. At every such visit, the doctor will record the patient's blood pressure and duly write it down, enter the information in the patient's electronic dossier, or perhaps even construct a patient-specific curve. Additionally, the doctor can decide to prescribe blood pressure lowering agents. Such agents do not lightly enter the pharmaceutical market, but rather undergo years of scrutiny, including clinical trials in which a group of experimental volunteers are followed during several months or even years. Let us assume that we dispose of 200 study subjects. As a general rule, one would assign half of these to standard therapy and the other half to the experimental treatment under investigation. At the end of the study, there are 200 blood pressure profiles, nicely divided over two groups. Figure 1 provides an examples of such a small *longitudinal* study.

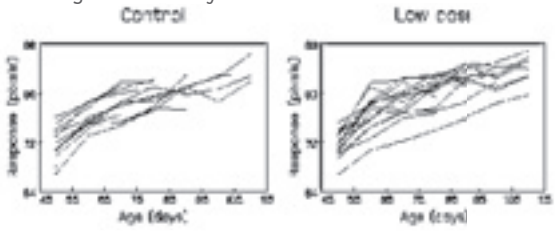


Figure 1: Example of data from a longitudinal study.

Both the doctor, with the single patient's profile for his or her patient, and the pharmaceutical company, are thus confronted with *repeated measures* or *longitudinal data*. The term refers to a specific class of data, necessitating specific modes of analysis. That is not the end of the story, though. An additional problem is that not all patients remain on study as long as scheduled by design. Some move, switch to alternative medication, on doctor's advice or by their own choice. Some simply may no longer want to face the burden placed upon them by the requirements from the study protocol and leave it. All of this is in line with human nature and, importantly, explicitly allowed for by the Convention of Helsinki, designed and globally adopted to protect patients' rights based on the terrifying events at Nazi concentration camps. Incompletely collected studies require special attention, and special methodology, from the data analyst. Both repeatedly and incompletely collected data, and the analysis thereof, are at the heart of the Center for Statistics in Diepenbeek and the Biostatistical Centre of Katholieke Universiteit Leuven. The latter center is headed by Geert Verbeke, together with Marc Aerts leading the new institute IB². Members of both centers have extensively published on repeated and incomplete data. Geert Verbeke is also a visiting faculty member of Hasselt University. It is no wonder, then, that these research themes will play a key role in the newly founded institute.

Education

As indicated earlier, over the past 25 years, UHasselt and CenStat established a Master in Statistics, with trajectories in biostatistics, biostatistics/ICP, applied statistics, and bioinformatics. We already pointed out the international dimension of advisory board, staff, and students.

The program is broad in concept and stably based on the cross-fertilization between research, education, and scientific consulting, a typical mix for CenStat. This directly translates in to a high level of satisfaction by both alumni as well as key players in the field. The students, indeed, invariably find they have to work extremely hard!

At this time, about two hundred students follow a master education in applied statistics, biostatistics, or bioinformatics at Hasselt University, whereas, 30 years ago not only Flanders and Belgium but in fact continental Europe in its entirety were blind spots on the map, as far as these disciplines went.

►► A Master in (Bio)statistics at Hasselt University

The senior colleagues, Herman Callaert, Noël Veraverbeke, and Paul Janssen maintained extensive contacts, as early as the 1970s and 1980s, with the Anglo-Saxon statistical community. Herman and Paul sojourned at such locations as Stanford (California), Rochester (New York State), Baltimore (Maryland), and Seattle (in the state of Washington). In the middle of the 1980s, an honorary doctorate was bestowed upon Professor Sir David Cox (Nuffield College, Oxford). Contacts were established with Harvard University in Boston, Massachusetts. Harvard boasts, apart from a statistics department, also a department of biostatistics with no less than 200 staff and faculty! What was the European situation at that time… Work was to be done.

At the start of the program, in 1988, with support from among others David Cox from Oxford and Stephen Lagakos, Professor of Biostatistics in Harvard, there was little or no biostatistical experience readily available at Hasselt University. To staff the at the time one-year program, faculty was recruited from other Flemish and French speaking Belgian universities, supplemented with faculty from abroad. Owing to the Anglo-Saxon origin of the discipline, it is hardly surprising that, over the years, a considerable contingent was drawn from the United Kingdom and the United States. CenStat's growth not withstanding, with increased numbers of faculty in biostatistics and bioinformatics present on site and hence more courses taught by such local faculty, a core of visiting faculty members has been maintained. This is generally regarded as an important added value.

Dutch was, as candidate working language, not an option, in view of the international nature of guest professors and student population and, more recently, also the international composition background of locally based faculty. At the same time, key literature and even technical jargon is principally English. Hence, the choice for Disraeli's language was self-evident.

►► Students from the South

At the start of the 1990s, the administration then abbreviated by ABOS, now University Development Cooperation of the Flemish Inter-University Council (VLIR-UOS), recognized the program as a so-called International Course Program (VLIR-ICP). This program, chaired by Paul Janssen, formally started in 1993. The first series of alumni essentially took the biostatistics curriculum, supplemented with a summer term, a preparatory set of modules immediately preceding the start of the program. This made the program very heavy, because the summer term was followed, without interruption, by the core three trimesters, to then top it of, again without interruption, with a summer internship.

►► A Heavy and Tough Program

Needless to say that all of this made the program rather heavy. Five terms of hard work, without lifting one's pen, is heavy, even if a program is expected to be hard work. In 1999, the program was re-designed to achieve a better equilibrium, giving birth to the *Master of Science in Applied Statistics*. For the VLIR-ICP students, five contiguous terms were replaced by two academic year consisting of three terms, with a well-deserved summer break in between. Students started in *Applied Statistics* and carried on in *Biostatistics*.

In the meantime, further developments have taken place. First, the masters in statistics have been subsumed into the Bologna BaMa (bachelor-master) structure. There now is a two-year master in *Biostatistics*, essentially the previously described combination of the "old" *Applied Statistics* and *Biostatistics*. Next to this, there now is a two-year master in *Applied Statistics*, too, an extension of the former one-year *Applied Statistics*, with emphasis less on depth rather than breadth. The VLIR-ICP students essentially follow the two-year version in *Biostatistics,* with specific attention for selected courses of direct and special relevance for their particular background. To complete the picture, there is a two-year version of the *Master in Bioinformatics as well*, in collaboration wit the tUL School of Life Sciences.

►► Four Well-populated Programs

The programs require day-to-day labor. There are four program directors: Marc Aerts for *Applied Statistics*, Paul Janssen for the VLIR-ICP program, Tomasz Burzykowski for *Bioinformatics*, and Geert Molenberghs for *Biostatistics*. Running the programs would be next to impossible without the devoted support by the *Program Managers*: Martine Machiels and Viviane Mebis. Martine had been the administrative engine, responsible, for over two decades, for the program. Recently, Viviane joined her efforts. Administration is to be interpreted in the widest possible sense. Evidently, there are core administrative duties to be taken up, such as enrolment, exam organization, composition of time tables, certificates, etc. In addition to these more straightforward tasks, the *Program Managers* ensure the student from an equatorial country are nevertheless warmly clad during the Belgian winter, they offer advice for medical and familial incidents, support occasionally a pregnant student... In sum, they are the heart beat and the soul of the program.

►► Are There Oter Programs in Biostatistics?

During many years, UHasselt occupied a unique position in continental Europe as far as biostatistics goes, next to, of course, educational programs in statistics. In particular mathematical, theoretical statistics had had a long tradition on the continent. Also other applied disciplines were well represented at the time, such as financial mathematics, insurance mathematics and statistics, econometrics, etc. In the meantime, various programs have been established. In Flanders, there are the different but related programs in Leuven and Ghent. The educational authorities in the French speaking part of the country have declared biostatistics as a priority discipline. UHasselt and KULeuven signed a memorandum of collaboration, ensuring decade-long competition is replaced by fruitful collaboration, leading to exchange of students and staff. Without exaggeration, it is fair to say that Flanders now is a lighting spot on the biostatistics world map.

Scientific output

Doctoral Theses

Lin Dan
Classification, Dose-response Modelling and the Evaluation of Biomarkers in a Micro-array Setting, 28 maart 2008 (promotor: Prof. Dr. Tomasz Burzykowski; copromotor Prof. Dr. Ziv Shkedy)

Wouters Kristien
Classification methods multi-class multivariate longitudinal data, 9 september 2008 (promotor: Prof. Dr. Geert Molenberghs, copromotor: Prof. Dr. Jose Cortinas Abrahantes)

Namata Harriet
Flexible statistical models for microbial risk assessment and infectious diseases, 15 september 2008 (promotor: Prof. Dr. Marc Aerts, copromotor: Prof. Dr. Christel Faes)

Vangeneugden Tonny
Applying Psychometric Validation Methodology to Longitudinal Clinical Trial Data, 19 mei 2008 (promotor: Prof. Dr. Geert Molenberghs)

Van Sanden Suzy
Statistical Methods for Microarray-based Analysis of Gene-expression, Classification and Biomarker Validation, 11 september 2008, (promotor: Prof. Dr. Tomasz Burzykowski; copromotor Prof. Dr. Ziv Shkedy)

Maringwa John
Flexible modelling techniques and use of historical controls in animal studies, 26 september 2008 (promotor: Prof. Dr. Helena Geys, copromotor: Prof. Fr. Christel Faes)

Massonnet Goele
Contributions to frailty and copula modelling with applications to clinical trials and dairy cows data, 8 september 2005 (promotor: Prof. Dr. Paul Janssen, copromotor: Prof. Dr. Tomasz Burzykowski)

Valkenborg Dirk
Statistical Methods for the Analysis of High-resolution Mass Spectrometry Data, 12 september 2008, (promotor: Prof. Dr. Tomasz Burzykowski)

Laenen Annouschka
Psychometric Validation of Continuous Rating Scales from Complex Data, 12 december 2008 (promotor: Prof. Dr. Geert Molenberghs)

Scientific Publications

Published Articles in Web of Science (A1)

- Abatih, E.; Van Oyen, H.; Bossuyt, N. & Bruckers, L. (2008) Variance Estimation Methods For Health Expectancy By Relative Socio Economic Status. European Journal Of Epidemiology, 23(4). p. 243 249
- Alonso Abad, A. & Molenberghs, G. (2008) Evaluating Time To Cancer Recurrence As A Surrogate Marker For Survival From An Information Theory Perspective. Statistical Methods In Medical Research, 17(5). p. 497 504
- Alonso Abad, A.; Litiere, S. & Molenberghs, G. (2008) A Family Of Tests To Detect Misspecifications In The Random Effects Structure Of Generalized Linear Mixed Models. Computational Statistics & Data Analysis, 52(9). p. 4474 4486
- Alonso, A. and Molenberghs, G. (2008). Surrogate markers: Hopes and Perils. Pharmacoeconomics and Outcomes Research, 8, 255-259.
- Beunckens, C., Sotto, C. and Molenberghs, G. (2008). A simulation study comparing weighted estimating equations with multiple imputation based estimating equations for longitudinal binary data. Computational Statistics and Data Analysis. 52, 1533-1548.
- Beunckens, C.; Molenberghs, G.; Verbeke, G. & Mallinckrodt, C. (2008) A Latent Class Mixture Model For Incomplete Longitudinal Gaussian Data. Biometrics, 64(1). p. 96 105
- Bilau, M., De Henauw, S., Schroijen, C., Bruckers, L., Den Hond, E., Koppen, G., Matthys, C., Van de Mieroop, E., Keune, H., Baeyens, W., Nelen; V., Van Larebeke, N., Willems, J.L. and Schoeters, G. (2009) The relation between the estimated dietary intake of PCDD/Fs) and levels in blood in a Flemish population (50-65 years). Environment International, 35, 9-13.
- Bilau, M.; Matthys, Ch.; Baeyens, W.; Bruckers, L.; De Backer, G.; Den Hond, E.; Keune, H.; Koppen, G.; Nelen, V.; Schoeters, G.; Van Larebeke, N.; Willems, J.L. & De Henauw, S. (2008) Dietary Exposure To Dioxin Like Compounds In Three Age Groups: Results From The Flemish Environment And Health Study. Chemosphere, 70(4). p. 584 592
- Bogaerts, K. & Lesaffre, E. (2008) Modeling The Association Of Bivariate Interval Censored Data Using The Copula Approach. Statistics In Medicine, 27(30). p. 6379 6392
- Bollaerts, K.; Aerts, M.; Faes, Ch.; Grijspeerdt, K.; Dewulf, J. & Mintiens, K. (2008) Human Salmonellosis: Estimation Of Dose Illness From Outbreak Data. Risk Analysis, 28(2). p. 427 440
- Bollaerts, K.; Aerts, M.; Ribbens, S.; Van Der Stede, Y.; Boone, I. & Mintiens, K. (2008) Identification Of Salmonella High Risk Pig Herds In Belgium By Using Semiparametric Quantile Regression. Journal Of The Royal Statistical Society Series A Statistics In Society, 171. p. 449 464
- Braekers, R. & Veraverbeke, N. (2008) A Conditional Koziol Green Model Under Dependent Censoring. Statistics & Probability Letters, 78(7). p. 927 937
- Braekers, R. and Van Keilegom, I. (2008). Flexible modeling base on copulas in nonparametric median regression. Journal of Multivariate Analysis (Doi: 10.1016/j.jmva.2008.11.2009).
- Burzykowski, T.; Buyse, M.; Piccart Gebhart, M.J.; Sledge, G.; Carmichael, J.; Lueck, H.J.; Mackey, J. R.; Nabholtz, J.M.; Paridaens, R.; Biganzoli, L.; Jassem, J.; Bontenbal, M.; Bonneterre, J.; Chan, S.; Basaran, G.A. & Therasse, P. (2008) Evaluation Of Tumor Response, Disease Control, Progression Free Survival, And Time To Progression As Potential Surrogate End Points In Metastatic Breast Cancer. Journal Of Clinical Oncology, 26(12). p. 1987 1992
- Burzykowski, T.; Buyse, M.; Yothers, G.; Sakamoto, J. & Sargent, D. (2008) Exploring And Validating Surrogate Endpoints In Colorectal Cancer. Lifetime Data Analysis, 14(1). p. 54 64 [
- Buyse, M. (2008) Reformulating The Hazard Ratio To Enhance Communication With Clinical Investigators. Clinical Trials, 5(6). p. 641 642
- Buyse, M.; Burzykowski, T.; Michiels, S. & Carroll, K. (2008) Individual And Trial Level Surrogacy In Colorectal Cancer. Statistical Methods In Medical Research, 17(5). p. 467 475
- Claeskens, G.; Nguti, Rosemary & Janssen, P. (2008) One Sided Tests In Shared Frailty Models. Test, 17(1). p. 69 82
- Cortinas Abrahantes, J.; Shkedy, Z. & Molenberghs, G. (2008) Alternative Methods To Evaluate Trial Level Surrogacy. Clinical Trials, 5(3). p. 194 208
- De Coster, S., Koppen, G., Bracke, M., Schroijen, C., Den Hond, E., Nelen, V., de Mierrop, E.V., Bruckers, L., Bilau, M., B aeyens, W., Schoeters, G. and van Larebeke, N. (2008). Pollutant effects of genotoxic parameters and tumor-associated protein levels in adults: a cross sectional study. Environmental Health, 7.
- De Coster, S.; Koppen, G.; Bracke, M.; Schroijen, C.; Den Hond, E.; Nelen, V.; De Mieroop, E.; Bruckers, L.; Bilau, M.; Baeyens, W.; Schoeters, G. & Van Larebeke, N. (2008) Pollutant Effects On Genotoxic Parameters And Tumor Associated Protein Levels In Adults: A Cross Sectional Study. Environmental Health, 7

- Delhalle, L.; De Sadeleer, L.; Bollaerts, K.; Farnir, F.; Saegerman, C.; Karsak, N.; Dewulf, J.; De Zutter, L. & Daube, G. (2008) Risk Factors For Salmonella And Hygiene Indicators In The 10 Largest Belgian Pig Slaughterhouses. Journal Of Food Protection, 71(7). p. 1320 1329
- El Ghouch, A. & Van Keilegom, I. (2008) Non Parametric Regression With Dependent Censored Data. Scandinavian Journal Of Statistics, 35(2). p. 228 247
- Faes, C., Aerts, M., Molenberghs, G., Geys, H., Teuns, G. and Bijnens, L. (2008). A high-dimensional joint model for longitudinal endpoints of different nature. Statistics in Medicine, 27, 4408-4427.
- Faes, Ch.; Aerts, M.; Molenberghs, G.; Geys, H.; Teuns, G. & Bijnens, L. (2008) A High Dimensional Joint Model For Longitudinal Outcomes Of Different Nature. Statistics In Medicine, 27(22). p. 4408 4427
- Faes, Ch.; Geys, H.; Molenberghs, G.; Aerts, M.; Cadarso Suarez, C.; Acuna, C. & Cano, M. (2008) A Flexible Method To Measure Synchrony In Neuronal Firing. Journal Of The American Statistical Association, 103(481). p. 149 161
- Fairclough, D.L.; Thijs, H.; Huang, I Chan; Finnern, H.W. & Wu, A.W. (2008) Handling Missing Quality Of Life Data In Hiv Clinical Trials: What Is Practical?. Quality Of Life Research, 17(1). p. 61 73
- Gaddah, A. and Braekers, R. (2008). Weak convergence for the conditional distribution function in a Koziol-Green model under dependent censoring. Journal of Statistical Planning and Inference (Doi: 10.1016/j.jspi.2008.06.001).
- Genc, Y.; Gulsahi, K.; Gulsahi, A.; Yavuz, Y.; Cetinyurek, A.; Ungor, M. & Col, M. (2008) Assessment Of Possible Risk Indicators For Apical Periodontitis In Root Filled Teeth In An Adult Turkish Population. Oral Surgery Oral Medicine Oral Pathology Oral Radiology And Endodontology, 106(4). p. E72 E77
- Ghidey, W.; Lesaffre, E. & Verbeke, G. (2008) A Comparison Of Methods For Estimating The Random Effects Distribution Of A Linear Mixedmodel. Statistical Methods In Medical Research
- Goethals, K., Janssen, P. and Duchateau, L. (2008). Frailty models and copulas: Similarities and differences; Journal of Applied Statistics 35, 1071-1079.
- Goethals, K; Janssen, P. & Duchateau, L. (2008) Frailty Models And Copulas: Similarities And Differences. Journal Of Applied Statistics, 35(9). p. 1071 1079
- Gyselaers, W.; Verswijvel, G.; Molenberghs, G. & Ombelet, W. (2008) Interlobar Venous Flow Is Different Between Left And Right Kidney In Uncomplicated Third Trimester Pregnancy. Gynecologic And Obstetric Investigation, 65(1). p. 6 11
- Hens, N., Aerts, M., Shkedy, Z., Kimani, P.K., Kojouhariva, M., Van Damme, P. and Beutels, P. (2008). Modeling age-time dependent incidence rates of hepatitis B and estimating the prevalence and force of infaction using generalized additive models. Epidemiology and Infection, 136, 341-351.
- Hens, N.; Aerts, M.; Shkedy, Z.; Kimani, P.K.; Kojouhorova, M.; Van Damme, P. & Beutels, P. (2008) Estimating The Impact Of Vaccination Using Age Time Dependent Incidence Rates Of Hepatitis B. Epidemiology And Infection, 136. p. 341 351
- Hens, N.; Aerts, M.; Shkedy, Z.; Theeten, H.; Van Damme, P. & Beutels, P. (2008) Modelling Multisera Data: The Estimation Of New Joint And Conditional Epidemiological Parameters. Statistics In Medicine, 27(14). p. 2651 2664
- Hughes, S.; Keene, O.; Howitt, N.; Roes, K.; Ashby, D.; Beltangady, M.; Bird, S.M.; Burton, C.; Buyse, M.; Day, S.; Van Ewijk, P.; Fletcher, Ch.; Grieve, A.P.; Guinot, Ch.; Hand, D.J.; Hemmings, R.; Holt, D.; Hothorn, L.; Jagers, P.; Keiding, N.; Korhonen, P.; Lesaffre, E.; Lewis, J.A.; Molenberghs, G.; Neumann, N.; Olsen, K.J.; Van Osta, G.; Pocock, S.; Rockhold, F.; Rossi, A.; Senn, S.; Stijnen, T.; Vaeth, M.; Wiklund, S.J. & Wolfram, J. (2008) European Regulatory Agencies Should Employ Full Time Statisticians. British Medical Journal, 336(7638). p. 250 250
- Ip, E. and Molenberghs, G. (2008). Empirical Bayes. In: International Encyclopedia in Education.
- Jacobs, T.; De Ridder, F.; Rusch, S.; Van Peer, A.; Molenberghs, G. & Bijnens, L. (2008) Including Information On The Therapeutic Window In Bioequivalence Acceptance. Pharmaceutical Research, 25(11). p. 2628 2638
- Jacobs, T.; Rossenu, S.; Dunne, A.; Molenberghs, G.; Straetemans, R. & Bijnens, L. (2008) Combined Models For Data From In Vitro In Vivo Correlation Experiments. Journal Of Biopharmaceutical Statistics, 18(6). p. 1197 1211
- Janssen, I. & Molenberghs, G. (2008) A Flexible Marginal Modelling Strategy For Non Monotone Missing Data. Journal Of The Royal Statistical Society Series A Statistics In Society, 171(2). p. 347 373
- Kadankov, V. & Kadankova, T. (2008) A Two Sided Exit Problem For A Difference Of A Compound Poisson Process And A Compound Renewal Process With A Discrete Phase Space. Stochastic Models, 24(1). p. 152 172
- Kadankov, V. & Kadankova, T. (2008) Exit Problems For The Difference Of A Compound Poisson Process And A Compound Renewal Process. Queueing Systems, 59(3 4). p. 271 296
- Kellen, E.; Zeegers, M.P..; Bruckers, L. & Buntinx, F. (2008) The Investigation Of A Geographical Cluster Of Bladder Cancer. Acta Clinica Belgica, 63(5). p. 313 320
- Keune, H., Loots, . I., Bruckers, L., Bilau, M. Koppen, G., van Larebeke, N., Schoeters, G. and Nelen, V. (2008). Monitoring environment, health and perception; an experimental survey on health and environment in Flanders, Belgium. International Journal of global environmental issues, 8, 90-111.

- Lin, D., Shkedy, Z., Burzykowski, T., Ion, R., Göhlmann, H.W., Bondt, A.D., Perer, T., Geerts, T., Van den Wyngaert, I. and Bijnens, L. (2008). An investigation on performance of Significance Analysis of Microarray (SAM) for the comparisions of several treatments with one control in the presence of small-variance genes. Biometrical Journal, 50, 801-823.
- Litiere, S.; Alonso Abad, A. & Molenberghs, G. (2008) The Impact Of A Misspecified Random Effects Distribution On The Estimation And The Performance Of Inferential Procedures In Generalized Linear Mixed Models. Statistics In Medicine, 27(16). p. 3125 3144
- Maringwa, J.; Geys, H.; Shkedy, Z.; Faes, Ch.; Molenberghs, G.; Aerts, M.; Van Ammel, K.; Teisman, A. & Bijnens, L. (2008) Analysis Of Cross Over Designs With Serial Correlation Within Periods Using Semi Parametric Mixed Models. Statistics In Medicine, 27(28). p. 6009 6033
- Maringwa, J.T.; Geys, H.; Shkedy, Z.; Faes, Ch.; Molenberghs, G.; Aerts, M.; Van Ammel, K.; Teisman, A. & Bijnens, L. (2008) Application Of Semiparametric Mixed Models And Simultaneous Confidence Bands In A Cardiovascular Safety Experiment With Longitudinal Data. Journal Of Biopharmaceutical Statistics, 18(6). p. 1043 1062
- Massonnet, G., Janssen, P. and Burzykowski, T. (2008). Fitting conditional survival models to meta-analytic data by using a transformation toward mixed-effects models. Biometrics, 64, 834-842.
- Massonnet, G.; Janssen, P. & Burzykowski, T. (2008) Fitting Conditional Survival Models To Meta Analytic Data By Using A Transformation Toward Mixed Effects Models. Biometrics, 64(3). p. 834 842
- Meroc, E.; Faes, Ch.; Herr, C.; Staubach, C.; Verheyden, B.; Vanbinst, T.; Vandenbussche, F.; Hooybergs, J.; Aerts, M.; De Clercq, K. & Mintiens, K. (2008) Establishing The Spread Of Bluetongue Virus At The End Of The 2006 Epidemic In Belgium. Veterinary Microbiology, 131(1 2). p. 133 144
- Mintiens, K.; Meroc, E.; Faes, Ch.; Cortinas Abrahantes, J.; Hendrickx, G.; Staubach, C.; Gerbier, G.; Elbers, A.R.W.; Aerts, M. & De Clercq, K. (2008) Impact Of Human Interventions On The Spread Of Bluetongue Virus Serotype 8 During The 2006 Epidemic In North Western Europe. Preventive Veterinary Medicine, 87(1 2). p. 145 161
- Molas, M. & Lesaffre, E. (2008) A Comparison Of Three Random Effects Approaches To Analyze Repeated Bounded Outcome Scores With An Application In A Stroke Revalidation Study. Statistics In Medicine, 27(30). p. 6612 6633
- Molenberghs, G. and Orman, C. (2008). Surrogate endpoints: application in pediatric clinical trials. In: Concepts and Applications of Pediatric Drug Development Mulberg, A. (ed.)
- Molenberghs, G.; Beunckens, C.; Sotto, C. & Kenward, M. G. (2008) Every Missingness Not At Random Model Has A Missingness At Random Counterpart With Equal Fit. Journal Of The Royal Statistical Society Series B Statistical Methodology, 70. p. 371 388
- Molenberghs, G.; Burzykowski, T.; Alonso Abad, A.; Assam Nkouibert, P.; Tilahun Eshete, A. & Buyse, M. (2008) The Meta Analytic Framework For The Evaluation Of Surrogate Endpoints In Clinical Trials. Journal Of Statistical Planning And Inference, 138(2). p. 432 449
- Mossong, J.; Hens, N.; Friederichs, V.; Davidkin, I.; Broman, M.; Litwinska, A.; Siennicka, J.; Trzcinska, A.; Van Damme, P.; Beutels, P.; Vyse, A.; Shkedy, Z.; Aerts, M.; Massari, M. & Gabutti, G. (2008) Parvovirus B19 Infection In Five European Countries: Seroepidemiology, Force Of Infection And Maternal Risk Of Infection. Epidemiology And Infection, 136(8). p. 1059 1068
- Mossong, J.; Hens, N.; Jit, M.; Beutels, P.; Auranen, K.; Mikolajczyk, R.; Massari, M.; Salmaso, S.; Tomba, G.; Wallinga, J.; Heijne, J.; Sadkowska Todys, M.; Rosinska, M. & Edmunds, W. (2008) Social Contacts And Mixing Patterns Relevant To The Spread Of Infectious Diseases. Plos Medicine, 5(3). p. 381 391
- Namata, H.; Aerts, M.; Faes, Ch. & Teunis, P. (2008) Model Averaging In Microbial Risk Assessment Using Fractional Polynomials. Risk Analysis, 28(4). p. 891 905
- Namata, H.; Meroc, E; Aerts, M.; Faes, Ch.; Cortinas Abrahantes, J.; Imberechts, H. & Mintiens, K. (2008) Salmonella In Belgian Laying Hens: An Identification Of Risk Factors. Preventive Veterinary Medicine, 83(3 4). p. 323 336
- Nia, P.S.; Colpaert, C.; Vermeulen, P.; Weyler, J.; Pezzella, F.; Van Schil, P. & Van Marck, E. (2008) Different Growth Patterns Of Non Small Cell Lung Cancer Represent Distinct Biologic Subtypes. Annals Of Thoracic Surgery, 85(2). p. 395 405
- Piccart Gebhart, M.J.; Burzykowski, T.; Buyse, M.; Sledge, G.; Carmichael, J.; Luck, H.J.; Mackey, J.R.; Nabholtz, J.M.; Paridaens, R.; Biganzoli, L.; Jassem, J.; Bontenbal, M.; Bonneterre, J.; Chan, S.; Basaran, G.A. & Therasse, P. (2008) Taxanes Alone Or In Combination With Anthracyclines As First Line Therapy Of Patients With Metastatic Breast Cancer. Journal Of Clinical Oncology, 26(12). p. 1980 1986
- Piedbois, P. & Buyse, M. (2008) Endpoints And Surrogate Endpoints In Colorectal Cancer: A Review Of Recent Developments. Current Opinion In Oncology, 20(4). p. 466 471
- Rizopoulos, D.; Verbeke, G. & Molenberghs, G. (2008) Shared Parameter Models Under Random Effects Misspecification. Biometrika, 95(1). p. 63 74

- Schroijen, C.; Baeyens, W.; Schoeters, G.; Den Hond, E.; Koppen, G.; Bruckers, L.; Nelen, V.; De Mieroop, E. Van; Bilau, M.; Covaci, A.; Keune, H.; Loots, I.; KleinJ.s, J.; Dhooge, W. & Van Larebeke, N. (2008) Internal Exposure To Pollutants Measured In Blood And Urine Of Flemish Adolescents In Function Of Area Of Residence. Chemosphere, 71(7). p. 1317 1325
- Smeets, K.; Ruytinx, J.; Semane, B.; Van Belleghem, F.; Remans, T.; Van Sanden, S.; Vangronsveld, J. & Cuypers, A. (2008) Cadmium Induced Transcriptional And Enzymatic Alterations Related To Oxidative Stress. Environmental And Experimental Botany, 63(1 3). p. 1 8
- Smeets, K.; Ruytinx, J.; Van Belleghem, F.; Semane, Brahim; Lin, D.; Vangronsveld, J. & Cuypers, A. (2008) Critical Evaluation And Statistical Validation Of A Hydroponic Culture System For Arabidopsis Thaliana. Plant Physiology And Biochemistry, 46(2). p. 212 218
- Speybroeck, N., Marcotty, T., Aerts, M., Dolan, T., Williams, B., Lauer, J., Molenberghs, G., Burzykowski, T., Mulumba, M. and Berkvens, D. (2008). Titrating Theileria parva: single stocks against combination of stocks. International Journal for Parasitology, 118, 522-530.
- Sterna, J. & Burzykowski, T. (2008) Assessment Of The Usefulness Of The Fenestration Method In Cases Of Disc Extrusion In The Cervical And Thoraco Lumbar Spine In Chondrodystrophic Dogs. Polish Journal Of Veterinary Sciences, 11(1). p. 55 62 [http://hdl.handle.net/1942/9586]
- Tibaldi, F.; Renard, D. & Molenberghs, G. (2008) Accounting For Variability In Individual Hierarchical Clinical Trial Data. Pharmaceutical Statistics, 7(4). p. 285 293
- Tilahun Eshete, A.; Assam Nkouibert, P.; Alonso Abad, A. & Molenberghs, G. (2008) Information Theory Based Surrogate Marker Evaluation From Several Randomized Clinical Trials With Binary Endpoints, Using Sas. Journal Of Biopharmaceutical Statistics, 18(2). p. 326 341
- Tilahun, A., Maringwa, J.T., Geys, H., Alonso, A., Raeymaekers, L., Molenberghs, G., Van Den Kieboom, G., Drinkenburg, P. and Bijnens, L. (2009). Investigating association between behavior, corticosterone, heart rate, and blood pressure in rats using surrogate marker evaluation methodology. Journal of Biopharmaceutical Statistics, 19, 1-17.
- Trinh, X.B.; Tjalma, W.A.A.; Makar, A.P.; Buytaert, G.; Weyler, J. & Van Dam, P. A. (2008) Use Of The Levonorgestrel Releasing Intrauterine System In Breast Cancer Patients. Fertility And Sterility, 90(1). p. 17 22
- Valkenborg, D.; Janssen, I. & Burzykowski, T. (2008) A Model Based Method For The Prediction Of The Isotopic Distribution Of Peptides. Journal Of The American Society For Mass Spectrometry, 19(5). p. 703 712
- Valkenborg, D.; Van Sanden, S.; Lin, D.; Kasim, A.; Janssen, I.; Shkedy, Z.; Burzykowski, T.; Haldermans, Ph. & Zhu, Q. (2008) A Cross Validation Study To Select A Classification Procedure For Clinical Diagnosis Based On Proteomic Mass Spectrometry. Statistical Applications In Genetics And Molecular Biology, 7(2)
- Van Barel, G.; De Ceuninck, W. & Witvrouw, A. (2008) The Influence Of Geometrical Imperfections In Micromachined Cantilevers On The Extracted Young's Modulus Using A Simple Model. Journal Of Micromechanics And Microengineering, 18(11). p. 115027 ...
- Van Effelterre, T., Shkedy, Z., Aerts, M., Molenberghs, G., Van Damme, P. and Beutels, P. (2009). Contact patterns and their implied basic reproductive numbers: an illustration for varicella-zoster virus. Epidemiology and Infection, 127, 48-57.
- Van Sanden, S.; Lin, D. & Burzykowski, T. (2008) Performance Of Gene Selection And Classification Methods In A Microarray Setting: A Simulation Study. Communications In Statistics Simulation And Computation, 37(2). p. 409 424
- Vangeneugden, T., Molenberghs, G., Laenen, A., Alonso, A. and Geys, H. (2008). Generalizability in non-Gaussian longitudinal clinical trial data based on generalized linear mixed models. Journal of Biopharmaceutical Statistics, 18, 1-22.
- Veraverbeke, N. (2008). Conditional residual life under random censorship. In: Advances in Statistics. Felicitation volume in honour of B.K.Kale. B.C. Armold, U. Gather, S.M. Bendre (eds). MacMillan India, New Dehli, pp. 174-185.
- Verbeke, G.; Molenberghs, G. & Beunckens, C. (2008) Formal And Informal Model Selection With Incomplete Data. Statistical Science, 23(2). p. 201 218
- Verhulst, S.L.; Vael, C.; Beunckens, C.; Nelen, V.; Goossens, H. & Desager, K. (2008) A Longitudinal Analysis On The Association Between Antibiotic Use, Intestinal Microflora, And Wheezing During The First Year Of Life. Journal Of Asthma, 45(9). p. 828 832
- Vloeberghs, E; Van Dam, D; Franck, F; Staufenbiel, M; Serroyen, J.; Molenberghs, G. & De Deyn, Pp (2008) Altered Ingestive Behavior, Weight Changes, And Intact Olfactory Sense In An App Overexpression Model. Behavioral Neuroscience, 122(3). p. 491 497
- Vyt, P.; Maes, D.; Quinten, C.; Rijsselaere, T.; Deley, W.; Aerts, M.; De Kruif, A. & Van Soom, A. (2008) Detailed Motility Evaluation Of Boar Semen And Its Predictive Value For Reproductive Performance In Sows. Vlaams Diergeneeskundig Tijdschrift, 77(5). p. 291 298
- Wang, L. & Veraverbeke, N. (2008) Mse Superiority Of Bayes And Empirical Bayes Estimators In Two Generalized Seemingly Unrelated Regressions. Statistics & Probability Letters, 78(2). p. 109 117

- Wouters, K.; Cortinas Abrahantes, J.; Molenberghs, G.; Ahnaou, A.; Drinkenburg, W.H.I.M. & Bijnens, L. (2008) A Comparison Of Doubly Hierarchical Discriminant Analyses For Multiple Class Longitudinal Data From Eeg Experiments. Journal Of Biopharmaceutical Statistics, 18(6). p. 1120 1135

Published Articles in Web of Science by Visiting Faculty of CenStat

- Baigent, C., Harrell, F., Buyse, M., Emberson, J. and Altman, D. (2008). Ensuring trial validity by data quality assurance and diversification of monitoring methods. Clinical Trials 5, 49-55.
- Beijerink, N.J., Bhatti, S.F.M., Okkens, A.C., Dieleman, S.J., Duchateau, L. and Kooistra, H.S. (2008). Pulsatile plasma profiles of FSH and LH before and during medroxyprogesterone acetate treatment in the bitch. Theriogenology 70, 179-185.
- Catry, D., Duchateau, L. Van De Ven, J. Laevens, H., Opsomer, G., Hasebrouck, F. and De Kruit, A. (2008). Efficacy of metaphylactic florfenicol therapy during natural outbreaks of bovine respiratory disease. Journal of Veterinary Pharmacological Therapy, 31, 479-487.
- Crop, F., Van Rompaye, B., Paelinck, L., Vakaet, L., Thierens, H. and De Wagter, C. (2008). On the calibration process of film dosimetry: OLS inverse regression versus WLS inverse prediction. Physics in Medicine and Biology, 53, 3971-3984.
- Desmedt, C., Haibe-Kains, B., Wirapati, P., Buyse, M., Bontempi, G., Delorenzi, M. Piccart M. and Sotiriou, C. (2008). Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. Clinical Cancer Research, 14, 5158-5165.
- Einmahl, J. and Van Keilegom, I. (2008). Specification tests in nonparametric regression. Journal of Econometrics, 143, 88-102.
- Fieuws, S., Verbeke, G., Maes, B. and Vanrenterghem Y. (2008). Predicting renal graft failure using multivariate longitudinal profiles. Biostatistics, 9, 419-431.
- Komárek, A. and Lesaffre, E. (2008). Generalized linear mixed model with a penalized Gaussian mixture as a random-effects distribution. Computation Statistics and Data Analysis 52, 3441-3458.
- Linton, O., Sperlich, S. and Van Keilegom, I. (2008). Estimation of a semiparametric transformation model. Annals of Statistics, 36, 686-718.
- Mwalili, S., Lesaffre, E. and Declerck, S. (2008). The zero-inflated negative binomial regression model dith correction for misclassification: an example in caries research. Statistical methods in Medical Research, 17, 123-139.
- Rizopoulos, D. Verbeke, G., Lesaffre, E. and Vanrenterghem, Y. (2008). A two-part joint model for the analysis of survival and longitudinal binary data with excess zeros. Biometrics, 64, 611-619.
- Rizopoulos, D., Verbeke, G. and Lesaffre, E. (2008). Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. Journal of the Royal Statistical Association, Series B, in press.
- Sardari Nia, P., Colpaert, C., Vermeulen, P., Weyler, J., Pezzella, F., Van Schil, P. and Van Marck, E. (2008). Different growth patterns of non-small cell lung cancer represent distinct biologic subtypes. Annals of Thorac Surgery, 85(2), 395-405.
- Tas, O., De Rooster, H., Baert, E., Doom, M.H. and Duchateau, L. (2008). The accuracy of the lactate pro hand-held analyser to determine blood lactate in healthy dogs. Journal of Small Animals Practice, 49, 504-508.
- Van Dessel, E., Hubens, G., Ruppert, M., Balliu, L., Weyler, J. and Vaneerdeweg W. (2008). Roux-en-Y gastric bypass as a re-do procedure for failed restrictive gastric surgery. Surg Endosc. 22, 1014-1018.

Articles in Journals With International Reading Committee (A2)

- Alonso Abad, A. & Molenberghs, G. (2008) Surrogate End Points: Hopes And Perils. Expert Review Of Pharmacoeconomics & Outcomes Research, 8(3). p. 255 259 A2
- Keune, H.; Loots, I.; Bruckers, L.; Bilau, M.; Koppen, G.; Van Larebeke, N.; Schoeters, G. & Nelen, V. (2008) Monitoring Environment, Health And Perception: An Experimental Survey On Health And Environment In Flanders, Belgium. International Journal Of Global Environmental Issues, 8(1/2). p. 90 111 A2

Published Meeting Abstracts, Letters, etc. (A5)

- Burzykowski, T. (2008) Surrogate Endpoints: Wishful Thinking Or Reality?. Statistical Methods In Medical Research, 17(5). p. 463 466 A5 [http://hdl.handle.net/1942/8742]
- Buyse, M. E.; Squifflet, P.; Rowe, J.M.; Whisnant, J.K.; Bhagwat, D.; Allard, S. E.; Hellstrand, K. & Brune, M.L. (2008) Post Consolidation Immunotherapy With Histamine Dihydrochloride And Interleukin 2 In Aml: Assessment Of Consistency And Robustness Of Treatment Benefit For Patients In First Complete Remission.. Blood, 112(11). p. 680 680 A5 [http://hdl.handle.net/1942/9649]

- Jacobs, N.; Claes, N.; Thijs, H.; Dendale, P. & De Bourdeaudhuij, I. (2008) A Six Month Evaluation Of A Tailored Behaviour Change Programme For A Highly Educated Study Sample. Psychology & Health, 23(Suppl. 1). p. 151 152 A5
- Robaeys, G.; Nevens, F.; Van Eyken, P.; Starkel, P.; Colle, I.; Bruckers, L.; Van Ranst, M. & Buntinx, F. (2008) Intravenous Substance Use And The Outcome Of Liver Transplantation For Chronic Hepatitis C. Journal Of Hepatology, 48(Supp.2). p. 230 ... A5
- Standaert, A. R.; Van Holderbeke, M.; Cornelis, C.; Torfs, R.; Nelen, V; Berghmans, P.; Bruckers, L.; Van Gestel, G.; Claeys, N.; Van Campenhout, K.; Wildemeersch, D. & Verlaek, M. (2008) Modeling Human Exposure To Cadmium And Arsenic In The Northern Campine Region. Epidemiology, 19(6). p. S203 S203 A5

Books (author and co-author) (B1)

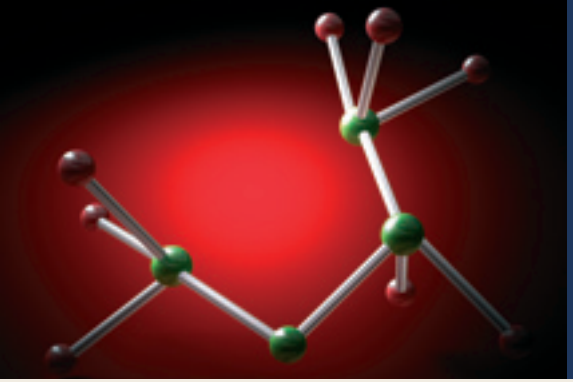
- Duchateau, L. & Janssen, P. (2008) The Frailty Model. (Springer New York) [http://hdl.handle.net/1942/9641]
- Fitzmaurice, G.; Davidian, M.; Molenberghs, G. & Verbeke, G. (2008) Longitudinal Data Analysis. (Chapman & Hall/CRC)

Book Chapter (B2)

- Buntinx, F.; Aertgeerts, B.; Aerts, M.; Bruyninckx, R.; Knottnerus, J.A.; Van Den Bruel, A. & Van Den Ende, J. (2008) Multivariable Analysis In Diagnostic Accuracy Studies. What Are The Possibilities? Knottnerus, J.A. & Buntinx, F. (Ed.) The Evidence Base Of Clinical Diagnosis, P. 146 167.
- Faes, Ch.; Geys, H. & Catalano, P. (2008) Joint Models Of Discrete And Continuous Longitudinal Data. Fitzmaurice, G. & Davidian, M. & Molenberghs, G. & Verbeke, G. (Ed.) Handbooks Of Modern Statistical Methods: Longitudinal Data Analysis, P. 327 348.
- Hens, N. (2008) Exploiting Data For Risk Evaluation. Scientific Exploitation Of Databases Within The Framework Of Food Safety Risk Assessment.
- Moons, Elke; Wets, G. & Aerts, M. (2008) Investigating The Performance If Rule Based Models With Increasing Complexity On The Prediction Of Trip Generation Distribution. Rossi, C. (Ed.) Frontiers In Brain, Vision And Ai, P. 167 182.
- Veraverbeke, N. (2008) Conditional Residual Lifetime Under Random Censorship. Arnold, B.C. & Gather, U. & Bendre, S.M. (Ed.) Advances In Statistics. Felicitation Volume In Honour Of B.K. Kale, P. 174 185.

Articles in Proceedings of Scientific Conferences (registered at ISI) (C1)

- Aerts, M., Bollaerts, K., Ribbens, K., Van der Stede, Y., Boone, I. and Mintiens, K. (2008). Semi-parametric regression to identify farms with high Salmonella infection burden. In: Proceedings of the 23rd International Workshop on Statistical Modelling, P. Eilers (ed.). Utrecht, pp.89-92.
- Bollaerts, K., Aerts, M., Faes, C., Dewulf, J. and Mintiens, K. (2008). Dose-illness models for human salmonellosis based on outbreak data. In: Proceedings of 23rd International Workshop on Statistical Modelling. P. Eilers (ed.), Utrecht, the Netherlands, pp.1245-130.
- Bollaerts, K., Messens, W., Delhalle, L., Aerts, M., Dewulf, J., Debusser, E., Boone, I. Grijspeerdt, K. (2008). Development of a modular quantitative microbial risk assessment to evaluate zoonotic risks in Belgium: Salmonellosis through consumption of pork as an example. In Proceedings of 10th European Symposium on Statistical Methods for the Food Industry, Louvain-la-Neuve, Belgium.
- Buntinx, F., Aertgeerts, B., Aerts, M., Bruyninckx, R., Knottnerus, J.A., Van den Bruel, A., Van den Ende, J. (2008). Multivariable analysis in diagnostic accuracy studies. What are the possiblities? In: The evidence base of clinical diagnosis. J.A. Knottnerus and F. Buntinx (eds.) John Wiley and Sons (ISBN 1405157879 / 9781405157872), pp. 146-167.
- Burzykowski, T., Valkenborg, D. (2008). Processing of high-resolution MALDI-TOF mass spectrometry data by using an average isotopic distribution. Proceedings of the Fourteenth National Conference on Application of Mathematics in Biology and Medine. Warsaw, Poland. Bodnar, M., Forys, U., (eds), 33-38. ISBN 83-903893-4-7.
- Callaert, H., (2008). Understanding confidence intervals. Proceedings of the Fifth Conference of the European Society for Research in Mathematics Education. CD-ROM edited by D. Pantazi and G. Philippou. ISBN 978-9963-671-25-0.
- Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G., Kenward, M.G. Shared-parameter models and missingness at random. Proceedings of the 24th International Biometric Conference, Dublin, July 13-18, 2008. ISBN 978-0-9821919-0-3.
- Faes, C., Geys, H. and Catalano, P. (2008). Joint models of discrete and continuous longitudinal data. In: Handbooks of Modern Statistical Methods: Longitudinal Data Anlysis. Fitzmaurice, G., Davidian, M., Molenberghs, G. and Verbeke, G. (Eds.) London: Chapman Hall/CRC, pp.327-348.



- Goeyvaerts, N., Hens, N., Ogunjimi, B., Aerts, M., Van Damme, P. and Beutels, P. (2008). Estimating transmission parameters using social contact data and serological data. Proceedings of the 24th International Biometric Conference, Dublin, July 13-18, 2008.
- Hens, N. (2008). Exploiting data for risk evaluation. In: Scientific exploitation of databases within the framework of food safety risk assessment. Federal Food Safety Authority. Brussels, Belgium.
- Hens, N., Aerts, M., Melagaro, A., Gay, N., Edmunds, J. Elucidation of the factors producing any observed differences in the epidemiological patterns of airborne infections across Europe. Proceedings of the 24th International Biometric Conference, Dublin, July 13-18 2008. ISBN 978-0-9821919-0-3.
- Hens, N., Wienke, A., Aerts, M. and Molenberghs, G. (2008). On the correlated frailty model for bivariate current status data. Proceedings of the workshop on composite likelihood methods. Warwick, UK, 15-17 April 2008.
- Jacobs, T., Straetemans, R., Rossenu, S., Dunne, A., Molenberghs, G., Bijnens, L. (2008). A new convolution-based approach to develop a level a in-vitro-in-vivo correlation (IVIVC) for an extended-release oral dosage form. 2007 Proceedings of the American Statistical Association, Biopharmaceutical Section [CD-ROM], Alexandria, VA: American Statistical Association, p.141-144.
- Leanen, A., Alonso, A., Molenberghs, G. and Vangeneugden, T. (2008). Reliability of a longitudinal sequence of scale ratings. Proceedings of the 24th International Biometric Conference, Dublin, Ireland. ISBN 978-0-9821919-0-3.
- Lejeune, O., Hens, N., Bilcke, J. and Beutels, P. Proceedings of the International workshop facing the challenge of Infectious Diseases – Integrating mathematical modeling, computational thinking and ICT applications, ISI Foundations, Turin (Italy), October 13-18, 2008.
- Leuridan, E., Hens, N. Ieven, G., Aerts, M., Van Damme, P. (2008). Maternal antibodies against measles after natural infection or vaccination in parturient women. Proceedings of the 26th Annual meeting of the European Society for Paediatric Infectious Diseases ESPID, Graz, Austria, 14-16 May 2008.
- Lin, D., Shkedy, Z., Burykowski, T., Aets, M., Göhlmann, H.W.H., De Bondt, A., Perera, T., Geerts, T., Van den Wyngaert, I., Bijnens, L. (2008). Information theory based marker evaluation 2007 Proceedings of the American Statistical Association, Biopharmaceutical Section [CD-ROM], Alexandria, VA: American Statistical Association, p.472-479.
- Lin, D.; Shkedy, Z.; Burzykowski, T.; Ion, R.; Goehlmann, H.W.H.; De Bondt, A.; Perera, T.; Geerts, T.; Van Den Wyngaert, I. & Bijnens, L. (2008) An Investigation On Performance Of Significance Analysis Of Microarray (Sam) For The Comparisons Of Several Treatments With One Control In The Presence Of Small Variance Genes. 5th International Conference On Multiple Comparison Procedures. p. 801 823. C1 [http://hdl.handle.net/1942/9006]
- Massonnet, G., Janssen, P. and Duchateau, L. (2008). Using copulas to model four-dimensional udder enfection data. Programme and abstract book 29th Annual conference of the International Society for Clinical Biostatistcs, Copenhagen, Denmark.
- Messens, W., Bollaerts, K., Delhalle, L., Aerts, M., Van der Stede, Y., Quoilin, S., De Busser, E., Dewulf, J., Maes, D., Boone, I., Mintiens, K.,Grijspeerdt, K. (2008). Development of a quantitaive microbial risk assessment to evaluate zoonotic risks in belgium: human salmonellosis through household consumption of minced pork meat. In First Belgian Symposium on Salmonella Research and Control in Pigs, Ghent Belgium, Book of proceedings, 43-47. F. Boyens, F. Pasmans (eds) (ISBN 9789058641434).
- Molenberghs, G., Aerts, M., Beunckens, C., Creemers, A., Hens, N., Kenward, M.G., Sotto, C. and Verbeke, G. Missing not at random models and their missing at random counterparts, in various modelling frameworks in livro in ACTAS do XVI Congresso da SPE, UTAD – Vila REAL, 2008.
- Moons, E., Wets, G. and Aerts, M. (2008). Investigating the Performance if rule-based models with increasing complexity on the prediction of trip generation and distribution. In: Frontiers in Brain, Vision and AI. C. Rossi (ed), pp.167-182. ISBN 978-953-7619-04-6

Articles in Proceedings of Scientific Conferences

- Aerts, M.; Bollaerts, K.; Ribbens, K.; Van Der Stede, Y.; Boone, I. & Mintiens, K. (2008) Semi Parametric Regression To Identify Farms With High Salmonella Infection Burden. Eilers, P. (Ed.) Proceedings Of The 23rd International Workshop On Statistical Modelling. p. 89 92.
- Bollaerts, K.; Aerts, M.; Faes, Ch.; Dewulf, J. & Mintiens, K. (2008) Dose Illness Models For Human Salmonellosis Based On Outbreak Data. Eilers, P. (Ed.) Proceedings Of 23rd International Workshop On Statistical Modelling. p. 1245 1300.
- Bollaerts, K.; Messens, W.; Delhalle, L.; Aerts, M.; Dewulf, J.; Debusser, E.; Boone, I. & Grijspeerdt, K. (2008) Development Of A Modular Quantitative Microbial Risk Assessment To Evaluate Zoonotic Risks In Belgium: Salmonellosis Through Consumption Of Pork As An Example. Proceedings Of The 10th European Symposium On Statistical Methods For The Food Industry.
- Burzykowski, T. & Valkenborg, D. (2008) Processing Of High Resolution Maldi Tof

Mass Spectometry Data By Using An Average Isotopic Distribution. Bodnar, M. & Forys, U. (Ed.) Proceedings Of The Fourteenth National Conference On Application Of Mathematics In Biology And Medicine. p. 33 38.

- Callaert, H. (2008) Understanding Confidence Intervals. Pantazi, D. & Philippou, G. (Ed.) Proceedings Of The Fifth Conference Of The European Society For Research In Mathematics Education.
- Creemers, A.; Hens, N.; Aerts, M.; Molenberghs, G.; Verbeke, G. & Kenward, M.G. (2008) Shared Parameter Models And Missingness At Random. Proceedings Of The 24th International Biometric Conference.
- Goeyvaerts, N.; Hens, N.; Ogunjimi, B.; Aerts, M.; Van Damme, P. & Beutels, P. (2008) Estimating Transmission Parameters Using Social Contact Data And Serological Data. Proceedings Of The 24th International Biometric Conference.
- Hens, N.; Aerts, M.; Melagaro, A.; Gay, N. & Edmunds, J. (2008) Elucidation Of The Factors Producing Any Observed Differences In The Epidemiological Patterns Of Airborne Infections Across Europe. Proceedings Of The 24th International Biometric Conference.
- Hens, N.; Wienke, A.; Aerts, M. & Molenberghs, G. (2008) On The Correlated Frailty Model For Bivariate Current Status Data. Proceedings Of The Workshop On Composite Likelihood Methods.
- Jacobs, T.; Straetemans, R.; Rossenu, S.; Dunne, A.; Molenberghs, G. & Bijnens, L. (2008) A New Convolution Based Approach To Develop A Level In Vitro In Vivo Correlation (Ivivc) For An Extended Release Oral Dosage Form. 2007 Proceedings Of The American Statistical Association, Biopharmaceutical Section. p. 141 144.
- Laenen, A.; Alonso Abad, A.; Molenberghs, G. & Vangeneugden, T. (2008) Reliability Of A Longitudinal Sequence Of Scale Ratings. Proceedings Of The 24th International Biometric Conference.
- Lejeune, O.; Hens, N.; Blicke, J. & Beutels, P. (2008) Integrating Mathematical Modeling, Computational Thinking And Ict Applications. Proceedings Of The International Workshop Facing The Challenge Of Infectious Diseases.
- Leuridan, E.; Hens, N.; Ieven, G.; Aerts, M. & Van Damme, P. (2008) Maternal Antibodies Against Measles After Natural Infection Or Vaccination In Parturient Women. Proceedings Of The 26th Annual Meeting Of The European Society For Paediatric Infectious Diseases Espid.
- Lin, D.; Shkedy, Z.; Burzykowski, T.; Aerts, M.; Göhlmann, H.W.H.; De Bondt, A.; Perera, T.; Geerts, T.; Van Den Wyngaert, I. & Bijnens, L. (2008) Information Theory Based Marker Evaluation. 2007 Proceedings Of The American Statistical Association, Biopharmaceutical Section. p. 472 479
- Massonnet, G.; Janssen, P. & Duchateau, L. (2008) Using Copulas To Model Four Dimensional Udder Enfection Data. Programme And Abstract Book 29th Annual Conference Of The International Society For Clinical Biostatistics.
- Messens, W.; Bollaerts, K.; Delhalle, L.; Aerts, M.; Van Der Stede, Y.; Quoilin, S.; De Busser, E.; Dewulf, J.; Maes, D.; Boone, I.; Mintiens, K. & Grijspeerdt, K. (2008) Development Of A Quantitative Microbial Risk Assessment To Evaluate Zoonotic Risks In Belgium: Human Salmonellosis Through Household Consumption Of Minced Pork Meat. Boyens, F. & Pasmans, F. (Ed.) First Belgian Symposium On Salmonella Research And Control In Pigs, Book Of Proceedings. p. 43 47.
- Molenberghs, G.; Aerts, M.; Beunckens, C.; Creemers, A.; Hens, N.; Kenward, M.G.; Sotto, C. & Verbeke, G. (2008) Missing Not At Random Models And Their Missing At Random Counterparts, In Various Modelling Frameworks. Livro In Actas Do Xvi Congresso Da Spe.

Other

- Gyselaers, W.; Molenberghs, G.; Van Mieghem, Walter & Ombelet, W. (2008) Prospective Doppler Study Of Maternal Renal Interlobar Venous Flow During Normal Pregnancy. [http://hdl.handle.net/1942/9187]

- Gyselaers, W.; Molenberghs, G.; Van Mieghem, Walter & Ombelet, W. (2008) Gestation Dependant Increase Of Renal Interlobar Venous Impedance Index In Pre Eclampsia.

Partners

CenStat maintains close collaborative links with various industrial partners. Biostatisticians are employed by academe and the industry alike. What are similarities and differences between both environments? What are the advantages of collaboration?

►► The academic versus the industrial biostatistician

In industry, one maintains direct and close contact with one's data. On the other hand, methodology is primarily developed in an academic setting en then transferred to industry. Of course, academic statisticians entertain close links with their data capturing partners, too. This is especially true for applied groups, such as CenStat. Evidently, direct contact with data is less prevalent in theoretical groups. At the same time, there is at times a somewhat conservative attitude in industry relative to novel methodology. This is one of the areas where the university can usefully contribute. Furthermore, education plays an important role. A mathematician, biologist, chemist, doctor… also trained in statistics, all take their particular angle when looking at and thinking about data. In industry, one becomes more quickly deeply acquainted with one's data; this stems not in the least from extensive work in the same substantive area. Others work on a variety of projects, each one coming with its own background. Variation is an added value to some, but needs to be contrasted with depth, of course. At any rate, groups where high-level and interdisciplinary collaboration is the standard, are becoming more prevalent and, at the same time, larger.

►► Student exchange between university and industry

Accepting thesis students for master or doctoral level training, implies an initial investment for the industry. It is nevertheless worth every cent. The student enters fresh and with the most recent insights and views, wants to work, and contribute. Return on investment starts, typically, around 6 months to one year into the project; hence, not of day one, of course. At times, the collaboration during an internship is so satisfying that it leads to scientific publications. Many students need to acquire expertise and practice in in report writing, the redaction of publications, oral communication, etc. Briefly, they have to become familiar and practiced with scientific communication. The thesis supervisor must take up his/her responsibility, i.e., making an investment in this respect.

An intensive collaborative effort

CenStat is very active in interdisciplinary research and collaborative endeavors. For example, seven CenStat students received a Janssen Pharmaceutica doctoral grant. Hence, this a very specific and satisfying form of collaboration. Like the majority of collaborations and business relations within industry, it originates from trust and from the conviction that the two sides have mutual contributions to make. What starts with mutual trust, based on successful collaboration, ultimately leads to freeing budgets and hence grants. From the start of the preclinical research group within Janssen, the collaboration has led to successful innovation. As such, the culture within the company changes. The preclinical department within Janssen assumes a leadership role, rather than that of a service unit. This is the difference between the "secretaries that can count" and "partners and leaders in science."

Many success stories are of a bottom-up nature, originating from the collaboration between two people or small groups of people who collaborate intensely, and from where expansion then gradually follows. Scientists and statisticians are bound to learn from one another, in order to stay on top of success. Collaboration begins in the planning phase and the design of the study. Initiating collaboration when the study is already ongoing and data need to be analyzed leads to "too little too late." Problems, and solutions, start with education. Classically, statistics was considered the typical difficult, failure-bound course. Teachers discussed less relevant, even boring, topics, and did so in an overly technical fashion. Currently, teaching is often done by people who get involved in practice; they understand the (ir)relevance of matters, and what an optimal level is at which courses need to be offered.

The department of (bio)statistics, within many biopharmaceutical companies, is populated with graduates from Hasselt University, i.e., with colleagues holding a master and/or doctoral level degree from Diepenbeek. Indeed, the PhD holding fraction is considerably larger than what it used to be. This has been the trend in the United States too, and for a long time. At the same time, it is fair to say that the Flemish and Belgian doctors in statistics are very well prepared and definitely can stand the test of international comparison.

Activities Organized in 2008

Marc Aerts

- Workshop on Survey Methodology by Geert Molenberghs, Maputo, 15-17 december 2008. Co-organizer, as part of the IUC Partner Programme with the University Eduardo Mondlane, Mozambique.

Paul Janssen

- IAP Workshop on Missing Information in Survival Data beyond Right Censoring, Gent, 18-19 september 2008. Member of the scientific committee and local co-organizer.

Geert Molenberghs

- XXX Conference of the International Society for Clinical Biostatistics, ISCB30, Prague, Czech Republic, 23-27 augustus 2009. Chair of the Scientific Program Committee
- International Seminar on Nonparametric Inference, ISNI2008, Vigo, Spain, Member of the scientific committee.

Ziv Shkedy

- NSC2008: Statistical Methods for Pharmaceutical Research and Early Development, Leuven, 23-25 september 2008. Member of the organizing committee.

Noël Veraverbeke

- AF Math Conference: Actuarial and Financial Mathematics Conference (international conference), Brussel 5-6 februari 2009. Member of the scientific committee.
- IAP Workshop on Missing Information in Survival Data beyond Right Censoring, Gent, 18-19 september 2008. Member of the scientific committee.
- 58th ISI Conference, Dublin, 21-27 augustus 2011. Chairman of the General Topics Committee
- International Seminar on Nonparametric Inference ISNI2008, Vigo, Spain. Member of the scientific committee

Prizes, Honors, and Awards

Kaatje Bollaerts

- Prize of the best International Workshop on Statistical Modelling 2008 Student presentation

Dan Lin

- Quetelet prize for the best master thesis biostatistics, UH, 2007

Geert Molenberghs

- 2009 Excellence in Continuing Education Award of the American Statistical Association for Course on Discrete Longitudinal Data, presented by Verbeke, G. and Molenberghs, G. at the Joint Statistical Meetings, Denver, Colorado (04/08/2008). Awarded at the Joint Statistical meetings in Washington, District of Columbia, XX/08/2009.
- 2008-2009: Belgian Francqui Chair, Universiteit Antwerpen.

Notes