**4nd Meeting of IBF Sunday 27/11/16**

**Analysis of Medical Record Data from Hospitals and Kupot Holim**

**Hebrew University, Mount Scopus, Jerusalem**
**Beit Meirsdorf, Room 405**

**9:00-9:30** **Registration and Coffee**

**9:30-9:40** **Welcome**

**9:40-10:30** **Natanel Goldschmidt**
**Ministry of Health**
**התוכנית הלאומית למדדי איכות בבתי חולים בישראל – דילמות סטטיסטיות ופתרונותיהן**

**10:30-11:20** **Laurence Freedman and Havi Murad**
**Gertner Institute of Epidemiology, Tel HaShomer**
**Exploring links between diabetes and cancer through analyzing a large Health Maintenance Organization database**

**11:20-11:50** **Coffee**

**11:50-12:40** **Dan Turner**
**Department of Pediatric Gastroenterology, Shaare Zedek Hospital**
**Utilizing Administrative Big Data for Population-Based Exploration of Inflammatory Bowel Diseases in Israel**

**12:40-14:00** **Lunch**

**14:00-14:50** **Michal Shauly-Aharonov**
**Department of Statistics, Hebrew University**
**Medical Applications of Change Detection Methods**

**14:50-15:40** **Avishai Mandelbaum**
**Faculty of Industrial Engineering and Management, Technion**
**Theompirical Research in Operations-Research and Stochastic-Processes: Theoretical + Empirical Journey through Service Systems (e.g. Hospitals, Phone-Centers, …)**

**15:40-16:10** **Discussant**
**Orly Manor**
**School of Public Health and Community Medicine, Hebrew University**

**16:10-16:30** **Floor discussion**

**Abstracts**

**Natanel Goldschmidt, Ministry of Health**

**התוכנית הלאומית למדדי איכות בבתי חולים בישראל – דילמות סטטיסטיות ופתרונותיהן**

התוכנית הלאומית למדדי איכות בבתי חולים בישראל, יצאה לדרך בשנת 2013. בשנה הראשונה נמדדו בתוכנית 5 מדדים בבתי החולים הכלליים. כעת, נמדדים בתוכנית כ-80 מדדים ב-6 תחומים שונים: בתי חולים כלליים, בתי חולים גריאטריים, בתי חולים פסיכיאטריים, תחנות טיפות חלב, חברות אמבולנס ושיקום חולי נפש.

המידע הרב שנאסף, מזמן התלבטויות סטטיסטיות רבות. הסוגיות שאנו עוסקים בהן סובבות נושאים רבים, וביניהם: דגימה, תקנון אוכלוסייה, זקיפת חסרים, רווח סמך, השוואות תקפות ופירסום תוצאות. ננסה להציג חלק מהדילמות הסטטיסטיות שנתקלנו בהן, ואת ההחלטות שהתקבלו.

**Laurence Freedman and Havi Murad**
**Gertner Institute of Epidemiology, Tel HaShomer**
**Exploring links between diabetes and cancer through analyzing a large Health Maintenance Organization database**

We describe a project to explore the link between diabetes and cancer through analyzing data from the Clalit Kupat Holim Database. We are addressing three broad questions in our investigation:

(a) are persons with diabetes at a higher (or lower) risk of developing cancer than persons without diabetes?

(b) among persons with diabetes are those with higher glucose levels or HbA1c levels at a higher (or lower) risk of developing cancer?

(c) among persons with diabetes do those taking certain glucose-lowering medications have a higher (or lower) risk of cancer?

In this lecture we will focus on questions (a) and (b). The Clalit database that we analyzed is a computerized file that includes information on over 2 million adult members of the Kupa over the age of 21y on January 1st 2002, who were the focus of our enquiry. We "followed" this historical cohort through the period 2002-12 using information in the database and information on cancer incidence and death from the Israel Cancer Registry that we linked to the Clalit file. The statistical analysis required us to take into account the special characteristics of the Clalit database, with regard to accuracy of information, completeness of information, and size of the file. Among the topics we will cover are: deriving from the information available a definition of diabetes, ethnicity, socio-economic status, body mass index and smoking; conducting computer-intensive survival analyses using the Cox model with time-dependent covariates, e.g. glucose or HbA1c levels; and handling missing values in such analyses. There was a large proportion of missing data in some time-dependent covariates (30%-50% in HbA1c; 20%-40% in glucose) at each time-point. We therefore developed a procedure for time-sequential multiple imputation at each time-point for the missing HbA1c and glucose values using the chained equations method, based on completed variables from previous time-points. Results from the analyses of questions (a) and (b) will

be presented, as well as results of simulations for examining the performance of the imputation method used when answering question (b).

**Dan Turner**
**Head, Juliet Keidan Institute of Pediatric Gastroenterology and Nutrition**
**Shaare Zedek Medical Center, Jerusalem**
**Utilizing Administrative Big Data for Population-Based Exploration of Inflammatory Bowel Diseases in Israel**

לפני 120 שנה לא היו בעולם מחלות מעי דלקתיות (קרוהן וקוליטיס כיבית) אך שכיחותן נמצאת בעליה מתמדת מאז שמחלת הקרוהן תוארה לראשונה בשנת 1932 על ידי ד״ר בריל קרוהן. מחלות מעי דלקתיות, כולל קרוהן וקוליטיס כיבית, הם מחלות כרוניות של דלקת במעי המתאפיינות בתקופות של רגיעה ותקופות של התלקחות אז יש שלשול, לעיתים דמי, כאב בטן, חולשה, תלונות שונות מחוץ למעי ובילדים גם פגיעה בגדילה.

עד לאחרונה לא היה לנו מידע על המספר המדויק של חולים בישראל. בשנת 2011 הוקמה רשת מחקר ארצית למחלות מעי דלקתיות בשם IIRN (Israeli IBD Research Nucleus). קבוצת חוקרי ה IIRN, ובמימון גרנט אקדמי מקרן הלמסלי, ריכזו נתוני חולים מבתי החולים הגדולים בישראל. בעזרת נתונים אלו פותחו אלגוריתמים ממוחשבים בארבעת קופות החולים למציאת חולי מחלות מעי דלקתיות בדיוק של כ 90%. בשיתוף עם האוניברסיטה העברית נפתח מרכז נתונים ארצי אליהם מועברים כל נתוני החולים מארבעת קופות החולים. בעקבות קול קורא לכל הגסטרואנטרולוגים בישראל נרשמו כבר 40 קבוצות מחקר סביב שאלות מחקר הקשורות למחלה וכעת עומלים על הקמת שרת שמאפשר גישה מרחוק לקבוצות המחקר לעבוד עם בסיס המידע שיכיל בסופו של התהליך נתוני 38 אלף חולים ועוד 130 אלף קונטרולות. פרוייקט זה ייאפשר לא רק הבנת האפידמיולוגיה של המחלה בישראל אלא גם ביצוע מחקרים תלויי אוכלוסיה בשאלות קליניות חשובות.

**Michal Shauly-Aharonov, Hebrew University**
**Medical Applications of Change Detection Methods**

Many physiological disturbances in processes within the human body are associated with diseases and severe clinical conditions. The physiological processes that are known to cause grave complications are usually strictly monitored. The problem is that a real disturbance may be masked by the variation of measurements, and only by accumulating information can it be detected early. Most of the existing methods of monitoring do not accumulate information; they are based on the last measurement or on recent past, and they usually examine the past "ad-hoc" and lean on the physician's intuition. We propose a detection method based on the Shiryaev-Roberts approach which considers the whole relevant history of measurements; with every new measurement, it gives an updated ("on-line") evaluation of the likelihood that a change has taken place somewhere along the time of monitoring. When this evaluation exceeds a certain upper control limit, a warning is given that the patient is at high risk. This alarm system is designed to have reasonable sensitivity and specificity, or alternatively to provide desired probability of detection and false-alarm, as well as sufficient time for intervention.

In particular, the talk will address two medical challenges of hospitals and kupot-holim:

(1) on-line detection of hazardous patterns in glucose level of pregnant women with type 1 diabetes.

(2) on-line detection of life-threatening changes in potassium level during hospitalization after a heart attack.

**Avishai Mandelbaum, Faculty of Industrial Engineering and Management, Technion**

**Theompirical Research in Operations-Research and Stochastic-Processes: Theoretical + Empirical Journey through Service Systems (e.g. Hospitals, Phone-Centers, …)**

I shall describe a personal research journey through service systems (e.g. telephone and chat centers, hospitals, banks,…). I typically view these systems through the lenses of a queueing scientist (e.g. "enjoying" congestion and flows); sometimes using operational characteristics (e.g. waiting, abandonment, priorities) as surrogates for financial, psychological and clinical performance; and always seeking phenomena that give rise to research challenges in Operations Research and Statistics.

The theory of queueing is ideally suitable for capturing the operational tradeoff that is at the core of any service: quality vs. efficiency. Three cases in point are the Erlang-A, -R and -S models: the first has become a common call center model, by accommodating the choice that customers enjoy, namely wait for service or abandon; the second arose from emergency departments, in which returns to service are prevalent; and the third captures operational symmetry between servers and customers. All three models, or their (asymptotic) fluid or diffusion counterparts, parsimoniously yet valuably portray complex realities. Here value is to be tested against real service systems, which is in contrast to prevalent practice in asymptotic-driven research. (In that practice, models are often remote from data, and the value of fluid/diffusion models is judged by its accuracy relative to alternative models.)

The ultimate goal of my research is an automatic creation, in real-time, of data-based models for service operations - analytical and simulation. The latter models will serve as a validation ground for the former, and both will be universally accessible for applications by researchers, students and practitioners. Prerequisites include, first and foremost, measurements of individual events (e.g. patient-physician transactions), which then support inference of model primitives, structure and protocols. The above goal has been pursued at the Technion IE&M, with data-support by its SEE Laboratory (SEE = Service Enterprise Engineering).