

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

University of Minnesota

514 Vincent Hall

206 Church Street S.E.

Minneapolis, Minnesota 55455

FAX (612) 626-7370

telephone (612) 624-6066

e-mail: ima-staff@ima.umn.edu

IMA Schedules via finger: finger_seminar@ima.umn.edu

Newsletters, Updates and preprints available via

anonymous ftp: [ftp.ima.umn.edu](ftp://ftp.ima.umn.edu), gopher: [gopher.ima.umn.edu](gopher://gopher.ima.umn.edu), www: <http://www.ima.umn.edu>

IMA NEWSLETTER # 252

July 1-31, 1997

1996-97 Program

MATHEMATICS IN HIGH-PERFORMANCE COMPUTING

1997 Summer Program

STATISTICS IN THE HEALTH SCIENCES

See the Winter 1997 IMA Update for a full description of the Summer 1997 program.

News and Notes

Special IMA Workshop:

Template-Driven Automatic Differentiation for Large-Scale Scientific and Engineering Applications

June 29-July 3, 1997

Organizers: Thomas Coleman (Cornell University), Fadil Santosa
(University of Minnesota) and William Symes (Rice University)

Sponsored jointly with the Cornell Theory Center
and CRPC, Rice University

IMA Summer Program:

STATISTICS IN THE HEALTH SCIENCES

July 7-August 22, 1997

Organizers: Donald A. Berry, Seymour Geisser, Patricia Grambsch,
Joel Greenhouse, Elizabeth Halloran, Nicholas Lange, Barry
Margolin, Sandy Weisberg, Scott Zeger (Chair), and Marvin Zelen.

PARTICIPATING INSTITUTIONS: Centre National de la Recherche Scientifique, Consiglio Nazionale delle Ricerche, Georgia Institute of Technology, Indiana University, Iowa State University, Kent State University, Michigan State University, Northern Illinois University, Ohio State University, Pennsylvania State University, Purdue University, Seoul National University (RIM - GARC), Texas A&M University, University of Chicago, University of Cincinnati, University of Houston, University of Illinois (Chicago), University of Illinois (Urbana), University of Iowa, University of Kentucky, University of Manitoba, University of Maryland, University of Michigan, University of Minnesota, University of Notre Dame, University of Pittsburgh, University of Southern California, University of Wisconsin, Wayne State University.

PARTICIPATING CORPORATIONS: Bellcore, Eastman Kodak, EPRI, Ford, Fujitsu, General Motors, Honeywell, IBM, Lockheed Martin, Motorola, Siemens, 3M.

Improved IMA Home Page

The IMA has substantially improved its home page on the World-Wide Web, accessible through netscape or other web-reading applications at

<http://www.ima.umn.edu>.

The page is continually under construction. We invite comments or suggestions, which may be addressed to

webmaster@ima.umn.edu.

In particular, we appreciate any information about World-Wide Web links appropriate to current and upcoming IMA programs.

Schedule for July 1–31, 1997

Special IMA Workshop:

Template-Driven Automatic Differentiation for Large-Scale Scientific and Engineering Applications

June 29–July 3, 1997

Organizers: Thomas Coleman (Cornell University), Fadil Santosa (University of Minnesota) and William Symes (Rice University)

Sponsored jointly with the Cornell Theory Center and CRPC, Rice University

There has been much progress in development of general-purpose Automatic Differentiation (AD) packages in the past few years. The developers have put a great deal of emphasis into differentiation of functions defined by large and complex legacy code. However, practical application of AD to large-scale simulation problems often neither requires, nor particularly benefits from, this degree of generality. Indeed, the generality built into the design of AD codes means that, by default, it does not exploit certain structure in the function evaluation, and thus puts a limit on the size of problems that can be tackled.

Modern simulators are often well organized and modular in style, and express cleanly the structure of the underlying algorithm at various levels of abstraction. Such code organization opens the possibility of a “product structure” of AD: differentiation applied separately at two or more levels of abstraction may be a far simpler task than monolithic differentiation of an entire code, yet yield equivalent or even superior results. For example, time-stepping schemes decompose into time loops and stencil calls; several researchers have produced very efficient derivative code by formal differentiation (usually by hand) of the time loop followed by AD applied to the stencil subprograms. Both steps are simpler than AD of the complete code, especially for production of the adjoint linearized operator (“reverse mode”). However, the decomposition permits fine control and many optimizations of the derivative code which would be difficult to achieve in a monolithic AD application.

This workshop will bring together researchers from a variety of fields who have an interest in practical applications of AD to contemporary scientific and engineering problems of very large scale, *e.g.* inverse problems and optimal design, to explore the benefits of algorithmic decomposition as described above. It will also involve AD package authors, who will offer their insights and expertise in design and application of AD and perhaps take away some revised goals for AD research. A number of AD users and inventors will present lectures on fundamentals and issues. Participants will be invited to bring code of current interest and will apply various AD tools during the workshop, reporting results in seminar as the week progresses.

Sunday, June 29

Talks today are in Lecture Hall EE/CS 3-180

| | | |
|---------------|---|---|
| 8:15 am | Registration and Coffee | Reception Room EE/CS 3-176 |
| 8:45 am | Welcome and Orientation | R. Gulliver, F. Santosa |
| 9:00-9:45 am | Jean Utke Argonne National Lab | ADOL-C Usage, Features and New Developments |
| 9:50-10:35 am | Lucas Roh Argonne National Labs | Differentiating C Code Using ADIC |
| 10:35 am | Coffee Break | Reception Room EE/CS 3-176 |
| 11:15-12:00 | Alan Carle Rice University | ADIFOR, ADIFOR-MP and ADJIFOR |
| 2:00-2:45 pm | John Guckenheimer Cornell University | Using AD to compute periodic orbits |
| 2:50-3:35 pm | Matthias Heinkenschloss Rice University | Derivatives in Optimal Control |
| 4:00 pm | IMA Tea (and more!) | Vincent Hall 502 (The IMA Lounge) |

A variety of appetizers and beverages will be served.

Monday, June 30

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------------|--|--|
| 8:30 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:00-10:00 am | Fadil Santosa University of Minnesota | Model problems from application and team selection |
| 10:00 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 10:30-11:15 | Christele Faure INRIA Sophia Antipolis | Main Features of Odyssee |
| 11:20-12:05 | Ralf Giering MPI für Meteorologie, Hamburg | TAMC: Source-to-source translation in forward and reverse mode |
| 2:00-2:45 pm | Thomas Coleman Cornell University | Automatic Differentiation and Structure |
| 2:50-3:35 pm | William Symes Rice University | A template approach to AD for time-marching schemes |

6:30 pm **Workshop Picnic** River Flats, behind Coffman Union

Participants will be asked to contribute \$8.00

Tuesday, July 1

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------------|---|--|
| 8:30 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:00-9:45 am | Uwe Naumann University of Dresden | Efficient Calculation of Jacobians by Optimal Node Elimination in Computational Graphs |
| 9:50-10:35 am | Koichi Kubota Chuo University | The second-order derivatives and rounding error estimates of AD |
| 10:35 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:15-12:00 | Victor Goldman University of Twente | Specification-driven adjoint model generation: an application to bounded domain problems |
| 1:30-2:15 pm | David Dobson Texas A&M University | Photonic bandgap structures |
| 4:30-5:15 pm | Progress report from teams | |

Wednesday, July 2

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------------|--|--|
| 8:30 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:00-9:45 am | Arun Verma Cornell University | ADMIT – AD and MATLAB Interface Toolbox |
| 9:50-10:35 am | Isabelle Charpentier Lab. de Model. & Calcul, Grenoble | Computation of the adjoint code of the atmospheric model MESO_NH by Odyssee |
| 10:35 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 1:30-2:15 pm | Olivier Talagrand École Normale Supérieure | Automatic Differentiation in Numerical Modeling of the Atmospheric and Oceanic Circulation. Present State and Perspectives |
| 4:30-5:15 pm | Progress report from teams | |

Thursday, July 3

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------------|--|---|
| 8:30 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:45–10:35 am | Linda Petzold University of Minnesota | Automatic Differentiation in Sensitivity Analysis and Optimal Control of Large-Scale Dynamical Systems |
| 10:35 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00–12:30 | Workshop wrap-up discussion | |

Friday, July 4

Independence Day, a University of Minnesota holiday. IMA offices will be closed.

IMA Summer Program:
STATISTICS IN THE HEALTH SCIENCES

July 7–August 22, 1997

Organizers: Donald A. Berry, Seymour Geisser, Patricia Grambsch, Joel Greenhouse, Elizabeth Halloran, Nicholas Lange, Barry Margolin, Sandy Weisberg, Scott Zeger (Chair), and Marvin Zelen.

Many important contributions to health care have been made by statistical scientists. Examples include development of randomized-control-trial and case-control methods of investigations. Nearly every aspect of health research has statistical components. These components are sometimes well developed and sometimes not. Frequently, researchers from different disciplines develop their own approach to a statistical problem, but with little interaction among the various disciplines. The overall aim of the workshop is to bring together statisticians with other substantive scientific workers, who are working in health areas with theoretical and methodological statisticians, to discuss and explore current statistical methods in the health sciences and to develop new methods where needed. In view of the rapidly changing health-care environment, the time is ripe for such an interchange.

There are five topics, one for each week of the program, with the exception of the Clinical Trials topic, which will take place over two weeks. The topics, dates and organizers are:

Week 1: Genetics. July 7–11 (Elizabeth Halloran, Seymour Geisser)

Week 2: Imaging. July 14–18 (Scott Zeger, Joel Greenhouse, Nick Lange)

Week 3: Diagnosis and Prediction. July 21–25 (Patricia Grambsch, Seymour Geisser)

Weeks 4 & 5: Design & Analysis of Clinical Trials. July 28–Aug. 8 (Donald Berry, Marvin Zelen)

Week 6: Statistics and Epidemiology: Environment and Health. August 18–22 (Joel Greenhouse, Elizabeth Halloran, Marvin Zelen, Barry Margolin)

| |
|------------------------------------|
| Week 1: Genetics. July 7–11 |
|------------------------------------|

Organizers: Elizabeth Halloran, Seymour Geisser

Statistical methods for population genetics are finding increasing areas of application, including DNA profiles, evolutionary phylogenies, hereditary disease patterns based on mitochondrial DNA, and even the origins and spread of language.

DNA profiles of several genetic loci are an important tool for identification to measure alleles in hypervariable loci, whether based on RFLP or PCR techniques. The use of these techniques, however, to ascertain similarity of two profiles and to estimate the relative frequency of a profile in a specified population and their underlying statistics have been the subject of much heated debate.

The building of evolutionary trees to establish phylogenies, or the relations between organisms, families, or species, is another important area of population genetics. In the medical sciences this is particularly important in the area of microbiology and parasitology. Examples include malaria, HIV, and tuberculosis. It is also used in studies of descent based on mitochondrial DNA. Mitochondrial DNA is inherited from the mother, so requires different statistical approaches than the usual linkage analysis based on nuclear DNA. Several difficult statistical issues arise in phylogeny building. One is the alignment of sequences. Second is the estimation of the distances based on some assumed evolutionary model. Third is the building of a tree structure based on a clustering algorithm. Fourth, is the problem of inference about the complex tree structure that was obtained. Methods of inference include likelihood methods and bootstrapping, but neither is well-understood in this application.

The statistical methods of population genetics and for the analysis of disease patterns based on mitochondrial DNA have lagged behind those for linkage analysis based on nuclear DNA. This workshop will be a unique opportunity for top-level statisticians to meet with quantitative geneticists to advance the statistical underpinnings of these emerging problems.

Monday, July 7

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------|---|---|
| 8:45 am | Registration and Coffee | IMA Lounge Vincent Hall 502 |
| 9:15 am | Welcome and Orientation | A. Friedman, R. Gulliver, E. Halloran |
| 9:30 am | Bruce Weir North Carolina State Univ. | Effects of population structure on paternity calculations |

Abstract: When an alleged father is not excluded on genetic grounds from being the father of a child, it is customary to present a “paternity index.” This is the ratio of the probability of the child’s genotype if the alleged father is the father, divided by the probability of the child’s genotype if some other man is the father. Although attention has been paid to the situation where the alleged father is related to the father, little attention has been given to the case when these two men, although different, have some evolutionary relationship by virtue of belonging to the same subpopulation. Both family and evolutionary relationships can be accommodated by the same set of population structure parameters. The net result in either case is to lessen the weight of genetic evidence in paternity disputes.

| | | |
|----------|---|--|
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | Seymour Geisser University of Minnesota | A Critique of Forensic Laboratories’ Use of Statistics for RFLP-DNA Matching and Profiling |

Abstract: We present the statistical methods used by the two principal forensic laboratories in the United States and critically discuss their claims and attendant flaws. In particular, we demonstrate that the assumptions that their relative frequency estimates of a DNA profile in a population are based on are flawed. Also, the claims that these estimates are conservative are shown to be false.

| | | |
|---------|---|---------------------------------------|
| 2:00 pm | Laurence Mueller Univ. of California-Irvine | The DNA Typing Controversy and NRC II |
|---------|---|---------------------------------------|

Abstract: The examination of genetic patterns found in the DNA of biological specimens at crime scenes has become an important and widespread tool for forensic analysis. An important component of this technology is the assessment of how likely it is for the genetic patterns in two different samples to match when in fact they are not from the same person. The techniques for estimating these statistics currently employed by essentially all forensic laboratories in the United States is to assume statistical independence among multiple gene loci and to ignore the possibility that laboratory errors may give rise to erroneous matches. A recent report by the National Research Council (NRC) has made many recommendations which generally support the current practices in forensic laboratories and contradict recommendations of an earlier report by the NRC on this same topic. I present an assessment of the recommendations of the new NRC report. In particular I address what is known about false positives in forensic laboratories, the extent to which forensic laboratories have followed recommendations in actual case work and what I consider rational solutions to these problems.

| | | |
|---------|----------------------------|-----------------------------------|
| 3:00 pm | Discussion Session | Future Directions in DNA Typing |
| 4:00 pm | IMA Tea (and more!) | Vincent Hall 502 (The IMA Lounge) |

A variety of appetizers and beverages will be served.

Tuesday, July 8

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------|---|--|
| 9:15 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:30 am | Susan Holmes Cornell University | Evaluating Phylogenies built from Molecular Data |

Abstract: This is a survey of methods for evaluating phylogenetic trees built from molecular data.

From the nonparametric bootstrap to model-based simulation methods many users evaluate their trees by assigning a confidence *index* to the clades of their trees.

I will give some examples carried out using currently available programs such as PHYLIP and SEQ-GEN.

I then propose some other indicators of quality of a phylogeny inspired from the statistical notions of STRESS, power and robustness.

| | | |
|----------|---|--|
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | Elizabeth Halloran Emory University | Bootstrap Confidence Levels for Phylogenetic Trees |

Abstract: In 1985, Felsenstein suggested use of the bootstrap to answer the question of how much confidence we should have in estimated phylogenetic trees. Since then, the method has been widely criticized as systematically biased. In this talk, we consider the use of the bootstrap for inference about phylogenetic trees. We show that Felsenstein's method is not biased, but that it can be corrected to agree better with standard ideas of confidence levels. The correction can be made using a second level of bootstrap. We demonstrate the method on set of aligned DNA sequences from 11 malaria parasites.

This is joint work with Brad Efron and Susan Holmes.

| | | |
|---------|---|------------------------|
| 2:00 pm | Bradley Efron Stanford University | The Problem of Regions |
|---------|---|------------------------|

Abstract: The figure below shows an example of the problem of regions. A bivariate normal vector having an unknown expectation vector and identity covariance matrix has been observed to fall into the region " R_{quad} ",

rather than the regions “ R_{con} ” or “ R_{lin} ”. How confident should we be that the expectation vector itself lies in R_{quad} ? A bootstrap methodology is proposed for answering this question, in a way that combines frequentist and Bayesian measures of confidence. The figure relates to the choice of the degree of a polynomial regression. Other examples will include assigning confidence to the branches of an observed phylogenetic tree, and saying how confident we can be that a density function is bimodal given that its kernel estimate appears that way.

This is joint work with Robert Tibshirani.

| | | |
|---------|---|--|
| 3:00 pm | Terry Speed Univ. of California, Berkeley | Mapping disease genes using identity by descent data |
|---------|---|--|

Abstract: Many people are currently searching for genes which contribute to the susceptibility of individuals to complex diseases such as schizophrenia, diabetes, coronary heart disease and asthma. Researchers typically have no solid knowledge concerning the nature or mode of action of such genes, or their frequencies in study populations. This has some relevance to the statistical methods used in searching for such genes. For example, traditional linkage analyses require a fully specified probability model, including gene frequencies and penetrances. In recent years there has been an increase in the use of Identity-By-Descent data, and statistical methods using only affected family members in this context, perhaps arising from a perception that these methods are to some extent more model-robust. Very recently a debate has arisen concerning the extent to which this last suggestion is true, and on the merits of affected only methods using IBD data, rather than traditional linkage analyses. This is the background to the talk.

Wednesday, July 9

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------|--|--|
| 9:15 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:30 am | Herman Chernoff Harvard University | The Bootstrap in Phylogenetic Analysis With Low Coverage |

Abstract: Felsenstein introduced the use of the bootstrap in phylogenetic analysis. Criticisms of the efficacy of this approach were addressed by Efron, Halloran and Holmes who pointed out some common misunderstandings of what the bootstrap is supposed to accomplish. They also pointed out that in this high dimensional problem,

the bootstrap provides an asymptotic approximation to a P-value for the test of the hypothesis that an observed clade is not a proper clade of the true phylogeny. Their illustration uses the Escalante-Ayala data on 11 species of plasmodium (malaria) with 1620 loci on part of the genome. Their analysis is confined to 223 nonmonotypic loci, i.e. loci where the 11-dimensional vector had more than one distinct component. Of these 223 loci, there were 119 singletons or vectors which appeared only once. This implies that the observed vectors represent a coverage of about 45% distribution of nonmonotypic vectors. It suggests that in this highly discrete and discontinuous problem of phylogenetic analysis, the sample size is not sufficiently large for the asymptotic first approximation to be meaningful. A simulation on a synchronous model, resembling that of the plasmodium data, but exaggerating the low coverage to 30% shows that the bootstrap can be very overoptimistic in assessing the reliability of observed clades.

This is joint work with Susan Holmes.

| | | |
|----------|---|--|
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | Wen-Hsiung Li University of Texas | Coalescent theory in population genetics |

Abstract: The coalescent approach to population genetic problems has proved to be very powerful and has become a standard approach in treating evolutionary history of populations from DNA sequence data. This approach involves many challenging mathematical, statistical and computational skills. Recent studies from our group and others will be reviewed in my talk. In particular, I will explain how this approach can be used to infer the age of the most recent common ancestor of the DNA sequences in a sample with special application to human evolution, how it can be applied to detect the presence of natural selection among the DNA sequences in a sample, and how it can be applied to estimate the long-term effective population size and mutation rate. Future directions of research will also be mentioned.

| | | |
|---------|---------------------------|--|
| 2:00 pm | Discussion Session | Inference in Evolutionary Theory and Population Genetics: How to deal with the statistically difficult problems? |
|---------|---------------------------|--|

Thursday, July 10

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------|--|------------------------------|
| 9:15 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:30 am | Warren J. Ewens University of Pennsylvania | Statistics in human genetics |

Abstract: Genetics and statistics are twin subjects, each of which has significantly affected the development of the other. Here we discuss various applications of statistical methods to problems in human genetics. One main application concerns linkage analysis, specifically how to use statistical methods to test whether a given marker locus is linked (and in practice closely linked) to a disease locus. Another application describes some statistical theory necessary to estimate genetic parameters when all the standard assumptions of statistics (independently and identically distributed random variables, in particular) do not hold.

| | | |
|----------|---|--|
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | Eleanor Feingold Emory University | Mapping disease genes by searching for shared genomic segments among distant relatives |

Abstract: Traditional linkage analysis and allele-sharing methods locate disease genes using family data. Linkage disequilibrium mapping uses unrelated individuals. Each of these methods has various strengths and weaknesses. I will discuss a new method that could combine some of the strengths of both: searching for shared genomic

tests are formed by comparing the observed score with the expected score under the null hypothesis of only nuclear autosomal mutation involvement. We study the optimality of score functions and weighting schemes under different models.

| | | |
|-------------|---|---|
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | Charles Geyer University of Minnesota | Markov Chain Monte Carlo and Pedigree Analysis |
| 12:00–12:30 | Discussion Session | Statistics in Disease Genetics; Statistics in the Genome Project |

| |
|------------------------------------|
| Week 2: Imaging. July 14–18 |
|------------------------------------|

Organizers: Scott Zeger, Joel Greenhouse, Nick Lange

Medical sciences, in particular radiology, have experienced an explosion of new technologies for imaging the structure and function of cells, tissues and organs. Magnetic resonance imaging (MRI), fractional MRI and PET scanning are examples of techniques that can provide practically real-time three-dimensional images. But the explosion of data has created important questions of how best to design studies to analyze the information they contain.

Methods for constructing and enhancing images, for identifying and classifying particular anatomical features such as lesions and tumors and for mapping the functional response to a stimulus require further development. Interactions among radiologists, biomedical engineers and statisticians will advance this emerging field.

Monday, July 14

Talks today are in Lecture Hall EE/CS 3-180

| | | |
|----------|---|--|
| 8:45 am | Registration and Coffee | Reception Room EE/CS 3-176 |
| 9:15 am | Welcome and Orientation | R. Gulliver, J. Greenhouse |
| 9:30 am | Nicholas Lange Harvard University | Overview |
| 10:30 am | Coffee Break | Reception Room EE/CS 3-176 |
| 11:00 am | Kamil Ugurbil University of Minnesota | Magnetic Resonance Imaging of Human Brain Function at High Magnetic Fields |
| 2:00 pm | Keith Thulborn Univ. of Pittsburgh Medical Center | Functional MRI: Technology and Clinical Applications |

Abstract: The application of fMRI for measuring sensory, motor and cognitive brain function has considerable importance not only for neuroscience in humans but has the potential to aid in medical management of many forms of neuropathology. Several issues must be addressed in the medical setting that are not present in a research environment. As the results of a given patient are to be applied to that patient's care, reliability and error assessment are of crucial importance and must be available as a part of the functional data. Not only must the studies of patients use selective and brief paradigms but must often be performed on a clinical

service by the MR technologist and nurse. To be clinically useful, the results need to be available immediately for interpretation. These challenges can be met by appropriate fMRI technology. This equipment must be maintained with appropriate quality assurance. Head motion is an important source of technical failure and methods to prevent its occurrence in acquisition or to provide correction in post processing are essential. The use of blood oxygenation level dependent (BOLD) contrast for fMRI of patients differs from normal cooperative volunteers in that not only are the patients ill and more anxious but also are often on medications and have neuropathology that may disrupt the neuronal hemodynamic response on which such fMRI is dependent. These issues require that the underlying physiology and metabolism of the brain be characterized for each patient and incorporated into the interpretation. The clinical uses of fMRI in patients will be discussed in terms of importance to medical management and on the ways that the above challenges can be met through a team approach that crosses disciplines including imaging, cognitive science, computer science, electrical engineering and, most importantly, statistics.

4:00 pm **IMA Tea (and more!)** Vincent Hall 502 (The IMA Lounge)

A variety of appetizers and beverages will be served.

Tuesday, July 15

Talks today are in Lecture Hall EE/CS 3-180

9:15 am **Coffee** Reception Room EE/CS 3-176

9:30 am **John A. Detre** Clinical applications of functional neuroimaging
Hospital of the University of Penna.

Abstract: Functional imaging refers to methodology used to image tissue function rather than anatomical structure. A variety of tissue functions can be probed, ranging from receptor binding to blood flow. In basic and clinical neuroscience, functional neuroimaging methods promise not only to provide new understanding of normal brain function, but also to improve the diagnosis and management of patients with diseases of the central nervous system. In some cases, neuroimaging techniques can directly and quantitatively measure a physiological parameter of interest, making interpretation straightforward. However, in many instances, a surrogate marker for the parameter of interest is measured, for example blood flow as a surrogate marker for task-specific functional brain activation. In such cases, alterations in the physiology of the surrogate function can modify the results obtained and their interpretation. Changes in coupling between a neurological function and its surrogate marker may vary across subjects, across brain regions, or with time. Even when surrogate systems retain normal function, behavioral performance must be considered in interpreting the results of patient studies. This talk will focus on the use of magnetic resonance neuroimaging methods for assessing brain function, and their potential applications to neurological disorders. Potential approaches to avoiding pitfalls of data interpretation or even capitalizing on potential pitfalls will be discussed.

10:30 am **Coffee Break** Reception Room EE/CS 3-176

11:00 am **Robert M. Weisskoff** Modeling Physics and Physiology for Imaging Func-
Harvard Medical School tion with Magnetic Resonance

Abstract: Unlike most medical imaging modalities, Magnetic Resonance Imaging (MRI) is based on effects that cross multiple biological levels: contrast depends on interactions between the local chemistry, water mobility, microscopic magnetic environment at the subcellular, cellular or vascular level, cellular integrity, etc. These interactions potentially allow for imaging functional changes in the same reference frame as the anatomic information. However, to tap this potential, we need methodologies that robustly incorporate the best models of the underlying physiology/physics interactions in order to extract these physiological data. This presentation will focus on two examples: pharmacokinetic modeling to help image breast cancer and noise and signal modeling in functional MRI to help quantify neuronal activation.

| | | |
|---------|--|--|
| 2:00 pm | Doug Noll Univ. of Pittsburgh Medical Center | Movement Correction in Functional Magnetic Resonance Imaging |
|---------|--|--|

Abstract: Functional Magnetic Resonance Imaging (fMRI) is a technique that may help us achieve a greater understanding of how the human brain functions. In fMRI, an MRI scanner is used to detect regional changes in cerebral metabolism or in blood flow, volume or oxygenation in response to brain activity. Changes in these physiological parameters lead to very small, yet detectable changes in MR images. Because the changes related to brain activity are small, fMRI is highly susceptible to sources of artifactual signal change, for example, physiological fluctuations, hardware instability and head movement. Here, we examine the challenge of correcting for head movement. Most movement correction algorithms have two basic steps - estimation of movement and correction of movement (interpolation). We have studied the error in both steps. We have developed a Fourier domain approach to reduce interpolation error for in-plane movements. For through-plane movements, we have developed a theoretical framework, based on sampling theory, to understand the sources of error and have devised acquisition and processing methods for reducing errors. Our approaches differ from commonly used approaches in that we have explicitly considered the details of image acquisition in development of methods for the correction of movement in fMRI.

| | | |
|---------|----------------|----------------------------------|
| 3:00 pm | Roundup | Thulborn, Detre, Weisskoff, Noll |
|---------|----------------|----------------------------------|

Wednesday, July 16

Talks today are in Lecture Hall EE/CS 3-180

| | | |
|---------|---------------|----------------------------|
| 9:15 am | Coffee | Reception Room EE/CS 3-176 |
|---------|---------------|----------------------------|

| | | |
|---------|---|--|
| 9:30 am | Christopher Genovese Carnegie-Mellon University | Towards New Methods of Inference for Functional Magnetic Resonance Imaging |
|---------|---|--|

Abstract: Functional Magnetic Resonance Imaging (fMRI) is an exciting and rapidly developing tool that enables cognitive psychologists and neuroscientists to study the human brain *in action*. During an fMRI experiment, subjects perform a set of cognitive tasks while images of their brain are acquired. Psychologists use these data to build and test models of human cognitive processing.

Most current statistical methods for the analysis of fMRI data are based on the classification of “active” locations in the brain. While identifying the location of neural activity is an important first step, a more sophisticated approach is needed to address many of the scientific questions to which fMRI data is applied. Moreover, many of the classification methods—based on hypothesis tests—are founded on simplistic statistical models, and few of these can evolve as new information about the underlying processes comes to light.

I will describe a family of statistical models for fMRI data that can account for some of the complexity in fMRI data. I will also present a framework for inference that enables investigators to address a much broader range of scientific questions, to test the predictions of cognitive theories, and to integrate fMRI results with other sources of information. This work offers several lessons—both computational and conceptual—for dealing with data realized from general spatio-temporal processes. Using examples from fMRI experiments, I illustrate the approach and demonstrate its advantages.

| | | |
|----------|---------------------|----------------------------|
| 10:30 am | Coffee Break | Reception Room EE/CS 3-176 |
|----------|---------------------|----------------------------|

| | | |
|----------|--|--|
| 11:00 am | William F. Eddy Carnegie-Mellon University | A Method for Quantitative Assessment of fMRI Data Processing |
|----------|--|--|

Abstract: As noted by Weisskoff (1996) in reference to fMRI experiments, “MR scanners can have quite variable stability even when the ‘normal’ stability measures are well within the instrumental norms.” Consequently, measurement of instrument variability and methods which correct data for the systematic components of that variability are of considerable importance. Equally important in fMRI are methods which estimate and correct

sometimes subsequently prove to be related to systematic anatomic differences between the groups rather than to true functional differences. Further refinement of statistical methods for analysis of PET blood flow studies will be important in maximizing the quality of the conclusions drawn using this technique.

| | | |
|---------|---|--|
| 2:00 pm | Edward H. Herskovits Johns Hopkins School of Medicine | Considerations in the Examination of a Large-Cohort Brain-Image Database |
|---------|---|--|

Abstract: Large-cohort neuroscience database projects, such as the Cardiovascular Health Study and the Baltimore Longitudinal Study of Aging, have as a critical component neuroradiology variables, such as infarct-like lesions on magnetic-resonance examination of the brain. Statistical analysis of the images and functional variables is central to this lesion-deficit research. I will describe a scalable relational-database architecture for this brain-image database, will present preliminary results of analysis of these data, and will propose future work for analyzing the spatial and signal components, in conjunction with clinical variables.

| | | |
|---------|----------------|----------------------------|
| 3:00 pm | Roundup | Worsley, Woods, Herskovits |
|---------|----------------|----------------------------|

| | |
|---------|--|
| 6:00 pm | Workshop Buffet Second floor, Radisson Hotel |
|---------|--|

Friday, July 18

Talks today are in Lecture Hall EE/CS 3-180

| | | |
|---------|---------------|----------------------------|
| 9:15 am | Coffee | Reception Room EE/CS 3-176 |
|---------|---------------|----------------------------|

| | | |
|---------|--|---|
| 9:30 am | Robert W. Cox Medical College of Wisconsin | Mathematical and Statistical Issues in the Imaging of Brain Function with MRI |
|---------|--|---|

Abstract: The detection, mapping, and quantification of neural activation with fMRI relies on extracting information from noisy measurements. In any such problem, it is necessary to understand both the signal that is to be detected, and the noise that interferes with the detection process. I will discuss some characteristics of the fMRI signal and noise, and will explain why both parts of the problem are active areas of research.

When one knows the form of the signal, then detection of activation is essentially a pattern matching problem: which parts of the brain show the expected pattern, and which do not? The most widely used method for this is a least squares fit technique. When there is no expected form to the received signal, then detection of activation involves looking for unknown temporal patterns. There are two methods widely used for this: principal components analysis, and fuzzy clustering.

Characterizing the response of the MR signal to changes in brain tissue state is one of the most active areas of research. Understanding, detecting, and correcting for the many artifacts that can occur in MRI time series of the human brain is also important. These include scanner effects, subject motion, and spatial correlation of the “noise” from physiological fluctuations in the subject.

If each voxel is considered separately, so that any possible spatial activation pattern is allowed, then a large (i , 10000) number of activation decisions must be made – this is the “multiple comparison” problem. To ensure that few of these decisions produce false activations, the threshold for detection must be made very stringent, making the detection of weak activations unlikely. One way to overcome this is to restrict the allowable spatial activation patterns; for example, to detect only in clusters above a certain volume.

| | | |
|----------|---------------------|----------------------------|
| 10:30 am | Coffee Break | Reception Room EE/CS 3-176 |
|----------|---------------------|----------------------------|

| | | |
|----------|--|--|
| 11:00 am | Steve Small University of Maryland | fMRI of Stroke Recovery: Longitudinal Assessment and Treatment Effects |
|----------|--|--|

Abstract: Stroke is a leading cause of morbidity and mortality, with many patients developing severe disability. We have initiated a series of studies to uncover the neurobiological course of stroke recovery with and without

interventions. Neural reorganization is assessed with task dependent BOLD functional magnetic resonance imaging (fMRI).

In this talk, I will focus on two investigations that present particular data analytic problems. First is a longitudinal imaging study of motor stroke recovery, in which patients with strokes are evaluated with fMRI and behavioral measures monthly during the period of maximal recovery. Second is a treatment study of language recovery in which reading therapy administered to a stroke patient was followed with fMRI and language assessment.

Three aspects of these studies present unique problems: (1) Imaging patients is more difficult than imaging normal subjects, since they have more voluntary and involuntary movements, do not tolerate well the scanner, and have abnormal tissue and vessels in the brain; (2) Longitudinal imaging requires anatomical and functional reliability, as well as methods to compare formally the images from different sessions; and (3) Treatment studies are complicated by the issues of reliability, spontaneous recovery, and relating changes in the brain to changes in behavior.

12:00–12:30 **Roundup** Cox, Small, Greenhouse, Zeger

| |
|---|
| Week 3: Diagnosis and Prediction. July 21–25 |
|---|

Organizers: Patricia Grambsch, Seymour Geisser

The overall goal is to understand how patient information should be used to diagnose disease, both for the individual patient and at the level of population studies. Classical statistical techniques, including discriminant analysis, regression, and classification have recently been joined by newer technologies, including probabilistic expert systems, neural networks and fuzzy logic. At the population level, public screening programs aim at early detection of disease in the hope that early diagnosis will lead to more cures or longer survival. Implementation costs for those programs amount of billions of dollars. Issues such as the initial age to begin participating in a screening program, the number of and timing of diagnostic exams, optimal methods for incorporating accruing patient information over a sequential series of exams and differential screening strategies for populations at varying levels of risk are open questions. There are expanding efforts to plan programs to detect some chronic diseases early in the expectation that earlier diagnosis combined with therapy will result in greater cure rates and/or increase survival. This is especially true in cancer, diabetes and heart disease. The range of problems ranges from the planning and analysis of early-detection clinical trials to the planning of public health programs. There are serious controversies in many of the applications to specific diseases. For example in breast cancer, randomized trials have shown that mammography is effective in reducing mortality only for women over 50 years of age. The American Cancer Society recommends that women over 50 should receive an annual mammogram, yet in the Scandinavian countries the recommendation is every two years and in the United Kingdom every three years. Nearly all of these controversies may be clarified by development of new statistical theories. Probabilistic models, incorporating trial results, can lead to the solutions of many of the problems for specific chronic diseases. The design of mass screening trials and programs can benefit from decision-theoretic analysis methods aiming at minimizing costs and optimizing health outcomes. The statistical issues raised here impact most areas of statistics and health.

The statistical issues include classification/discriminate analysis, regression, classification and newer technologies, including expert systems, notably probabilistic systems and neural networks. At the population level, screening studies are important. The statistical issues include length-based sampling, number and scheduling of screening studies, and differential screening strategies in populations at varying levels at risk. The statistical issues raised here impact most areas of statistics and health.

Monday, July 21

Talks today are in Seminar Room Vincent Hall 570

8:45 am **Registration and Coffee** IMA Lounge Vincent Hall 502

| | | |
|----------|---|---|
| 9:15 am | Welcome and Orientation | A. Friedman, R. Gulliver, S. Geisser |
| 9:30 am | Scott Zeger John Hopkins University | To be announced |
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | Sharon-lise T. Normand Harvard University | Delivering Medical Care: Statistical Methods for Assessing Quality of Care Using Observational Data |

Abstract: Two decades of health services research have produced numerous findings related to variations in the use of many medical and surgical procedures. For example, studies based on elderly Medicare patients who have had a heart attack indicate utilization of cardiac therapies decrease with age and are less likely to be used in women, in blacks, and in rural hospitals.

If the underlying risks across populations are *comparable*, then observed variations in processes of care and patient outcomes suggest that there may be inappropriate delivery of medical therapies and procedures. In part for these reasons, there is a national effort to develop and disseminate information on indicators of quality of care. In this talk, I discuss analytic methods for assessing quality of care, or *profiling*, using observational data with an application to heart attack patients: hierarchical probit models for developing a medical guideline regarding the appropriateness of coronary angiography, an invasive diagnostic technique for assessing the extent of heart disease, and Bayesian methods for inferring adherence to & validation of the guideline.

This is joint work with Mary Beth Landrum.

| | | |
|---------|--|--|
| 2:00 pm | Chap T. Le University of Minnesota | Estimating and Using the Receiver Operating Characteristic (ROC) Curve |
|---------|--|--|

Abstract: The speaker will propose a new estimator for the receiver operating characteristic (ROC) curve. This estimator converges to the ROC function uniformly on the interval $[0, 1]$ and has advantages as compared to the conventional empirical estimate, especially in the presence of tied observations. Two applications will be discussed: (1) a measure of goodness-of-fit for logistic regression models, and (2) evaluation of confounding effects in ROC studies.

| | | |
|---------|----------------------------|-----------------------------------|
| 4:00 pm | IMA Tea (and more!) | Vincent Hall 502 (The IMA Lounge) |
|---------|----------------------------|-----------------------------------|

A variety of appetizers and beverages will be served.

Tuesday, July 22

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------|--|--|
| 9:15 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:30 am | Leonardo Epstein Battelle Memorial Institute | Selection of thresholds for a continuous response assay to maximize the agreement with and ordinal standard: a Bayesian approach |

Abstract: Intense research is underway to develop in vitro methods or tests to provide alternatives to in vivo safety and toxicological studies. The purpose of many of these in vivo studies is to classify substances into ordered categories according to their potential hazard to humans. Often, regulatory agencies establish these categorical classifications to identify precautions when humans handle hazardous substances or are at risk of exposure to them.

If the purpose of an in vitro method is to replace such an in vivo study, then the transit of the in vitro method from development to routine use requires demonstration that the in vitro classification agrees, to a satisfactory degree, with the in vivo classification. The *in vivo* classification, therefore, plays the role of a standard.

The response of in vitro methods is usually continuous, for instance the fraction of cells in an engineered tissue

that survive exposure to a substance. However, often the purpose of in vitro methods is to predict an in vivo hazard classification of substances. In such cases, the test developer usually partitions the range of the response into intervals that correspond to the categories of the classification.

To exploit the full potential of an in vitro assay to classify substances correctly, we put forth criteria and methods to select a partition of the assay's response range to maximize the agreement between the in vitro results and the standard classification. These selection methods use the predictive densities of in vitro responses from substances whose standard classification is known. From these training data one computes the posterior predictive probability of correct classification, or some other criterion, as a function of the thresholds that define the partition. An optimal partition is one that maximizes the criterion. To our knowledge, in vitro test manufacturers use informal heuristic procedures to select the partition.

| | | |
|----------|--|--|
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | George Papandonatos SUNY-Buffalo | Bayesian Prediction for Weibull regression models with Gamma Frailty |

Abstract: In Bayesian hypothesis testing involving censored lifetime data, it is often of interest to predict whether continuation of the experiment will significantly alter the inferences drawn at an interim stage. A possible approach to this prediction problem has been described for Weibull-distributed data in Papandonatos & Geisser (1997b). However, the proportional hazards assumption may fail to hold in practice, due to a selection effect arising from the omission of important confounders. To accommodate this unmeasured heterogeneity in the target population, we introduce a Gamma frailty term and show that it leads to a marginal model that falls within the accelerated failure time framework examined in Papandonatos & Geisser (1997a). We illustrate the proposed methodology using the Malignant Melanoma data of Andersen et al. (1993).

| | | |
|---------|---|---|
| 2:00 pm | Wesley Johnson University of California-Davis | Predictive Inference for Binomial, Survival, and Screening Models |
|---------|---|---|

Abstract: We indicate the relative ease of performing Bayesian predictive inferences in a variety of settings. Modern computational tools, in addition to some new methods of prior elicitation, facilitate the methodology. We illustrate with several examples.

Wednesday, July 23

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------|---|---|
| 9:15 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:30 am | Seymour Geisser University of Minnesota | Comparing Two Tests used for Diagnostic or Screening Purposes |

Abstract: We present a Bayesian approach to the comparison of two binary diagnostic tests in a decision framework. It is shown, under a sensible loss function, that both a larger sensitivity and specificity are sufficient for one test to be superior to another. However, even if both the predictive value positive and negative of one test are larger than another test, it needs not to be superior, with respect to that loss function. An example of this approach is given that compares two Elisa screening tests for antibodies to the AIDS virus. We also show how a binary test that depends on a random variable may be optimally dichotomized with respect to the loss function.

| | | |
|----------|--|--|
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | Ming-Dauh Wang University of Minnesota | Optimal Administration of Multiple Screening Tests |

Abstract: We consider the case where mass screening for a disease or a characteristic is required. We then

develop the methodology of how to optimally administer one or more diagnostic tests, either sequentially or simultaneously in this mass screening situation. The use of different diagnostic tests as well as the repeated use of the same test are considered.

This is joint work with Seymour Geisser.

| | | |
|---------|--|--|
| 2:00 pm | Joseph L. Gastwirth George Washington University | The Utility of the Hui-Walter Paradigm for the Evaluation of Diagnostic Tests in the Analysis of Social Science Data |
|---------|--|--|

Abstract: After reviewing the basic framework for the evaluation of the accuracy rates of diagnostic tests some potential uses in law and social science will be described. The sensitivity of the original Hui-Walter model to violations of various underlying assumptions will be summarized. Modifications, relying on alternative sets of assumptions that seem suitable for the new area of application, will be described and illustrated. Insights into the effect misclassification errors have on the interpretation of unemployment data will be given. If time permits a potential use of a one population model, instead of the two subpopulations required by the Hui-Walter method, in the medical screening test will be illustrated.

This is joint work with Michael D. Sinclair.

Thursday, July 24

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------|--|---|
| 9:15 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:30 am | Marvin Zelen Harvard Sch. of Public Health | Issues and Problems in the Early Detection of Disease |

Abstract: This talk presents a general overview of problems relating to detecting disease early through special screening programs. Among the issues which will be discussed are: optimal planning of trials to evaluate benefit, assessment of scientific evidence and the planning of public health programs. Modeling the natural history of disease is an important factor in assessing benefit due to lead time bias and length biased sampling. Relationships between meta-analysis and modeling will be discussed. The talk will conclude with a discussion of open problems and recent work on some of these problems.

| | | |
|----------|---|---|
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | Sandra Lee Dana-Farber Cancer Inst. & Harvard | Planning Public Health Programs for the Early Detection of Disease: Applications to Breast Cancer |

Abstract: Public Health programs to diagnose chronic diseases are becoming an important part of the medical care system in developed countries. Such programs have been common in the past for tuberculosis, but are now beginning to be recommended for high blood pressure, diabetes, breast, colon, ovary and prostate cancers. These programs involve special diagnostic examinations which are ordinarily given periodically. However there are no guidelines to assist public health planners in planning such programs. The main issues are: (i) at what age should individuals begin to participate in such programs; (ii) how should one choose the spacings between examinations as a function of age; and (iii) how does one choose examination schedules for individuals who are at elevated risk because of family history and/or genetic or other markers. This investigation will outline a general stochastic model for the early detection process. Illustrations will be made to scheduling examinations to detect female breast cancer using mammography.

| | | |
|---------|---|--|
| 2:00 pm | Karen Kafadar University of Colorado-Denver | Estimating Lead Time and Screening Benefit in Randomized Cancer Screening Trials |
|---------|---|--|

Abstract: Screening tests are used frequently for control of diseases such as cancer. The increased survival time of screen-detected cases over those that are detected clinically may be due in part to “lead time,” or the length

of time by which the disease is diagnosed earlier by screening in the presence or absence of any real extension in survival time. A realistic evaluation of screening needs to assess the true benefit of screening; i.e. the length of time by which survival has been extended, beyond merely the time of the advanced diagnosis.

The comparison of survival measured from time of entry between cases in a screening arm and in a control arm in randomized studies avoids the lead time bias. If the effects of average lead time and average benefit on survival are additive, these effects can be estimated by recognizing that (a) the difference in survival curves since time of diagnosis confounds benefit time and lead time, but (b) the difference in survival curves since time of start of study involves benefit only. The method is evaluated on simulated data for its accuracy and also on data from randomized studies and may have important consequences for (1) planning required follow-up of future randomized studies, and (2) providing a method for estimating the benefit in nonrandomized studies where an external estimate of the lead time can be borrowed.

This research represents joint work with Philip C. Prorok

IMA Industrial Postdoc Seminar

The seminar will meet from 1:00 – 4:00 pm today in Murphy Hall 130. The format of the seminar is:

1. Presentation of projects and problems from industry (Honeywell, Lockheed Martin and Kodak) on which the industrial postdocs are working.
2. Informal suggestions and discussion among the participants.

The seminar is directed by Avner Friedman and Walter Littman. Visitors who plan to attend are requested to inform Dr. Friedman.

5:00 pm **Workshop Buffet**
 Courtyard and Vincent Hall 120

Friday, July 25

Talks today are in Seminar Room Vincent Hall 570

| | | |
|---------|---|---|
| 9:15 am | Coffee | IMA Lounge Vincent Hall 502 |
| 9:30 am | Giovanni Parmigiani Duke University | Timing Medical Examinations via Intensity Functions |

Abstract: This talk presents a decision theoretic method for timing medical examinations. The specific model is motivated by screening—that is examining asymptomatic individuals for hidden disease or risk-increasing conditions. Ideas and strategies may be applied more broadly to problems in which a stochastic process is monitored over time with a costly and possibly faulty data collection procedure. From a technical viewpoint, the approach of this article is based on modeling the decision space as a space of functions, termed screening intensity functions. Results include explicit rules for deciding whether or not an individual of given age and risk factors should be screened for a disease, and for deciding when an individual examined today should be examined again. For example, in a special case, the optimal frequency of examinations is proportional to the square root of the incidence of the disease. Results are illustrated using data on breast cancer screening.

| | | |
|----------|---|---|
| 10:30 am | Coffee Break | IMA Lounge Vincent Hall 502 |
| 11:00 am | Claudia Henschke Cornell University Medical College | Screening for Lung Cancer: A new Paradigm |

Abstract: Lung cancer is the malignancy accounting for most deaths in the United States. In 1991, it accounted for 143,000 deaths. In smokers the lifetime risk can exceed 10%. Currently, overall five-year survival rate as

of time of diagnosis is only slightly in excess of 10%; however, with (diagnosis and) treatment in Stage I, the five-year survival rate is as high as 70%. In the 1980's, two large randomized screening studies addressed the value of quarterly sputum cytology and concluded that this did not change the lung cancer mortality. Although they did not address the value of CXR screening, another large randomized study, the Mayo study, contrasted CXR screening every 4 months, using stereo PA views and coupled with sputum cytology, against routine care and also showed no differences in the mortality rates.

Based largely on the Mayo study, prominent national organizations (i. e., National Cancer Institute, American Cancer Society, American Medical Association) recommended against screening for lung cancer even in high-risk individuals in periodic health examinations. Nevertheless, many physicians continue to do so. These screening practices by clinicians are aimed not merely at reduction in mortality but cost containment as well: treatment of incurable lung cancer averages more than \$35,000 per patient. These screening practices may be justifiable despite the recommendations. Added justification for subsequent CXR screening for lung cancer derives from the fact that newer CXR techniques are superior to those employed in the Mayo study. PA and lateral views have been adopted throughout the world in place of two stereo views used in the Mayo study. In reference to current state-of-the-art CXR techniques, the frequency of detectable small solitary pulmonary nodules (SSPNs) and of their malignancy have been studied, but not the attendant gain in survival. Nodules less than 1 cm in diameter are, however, essentially undetectable even by today's highest-quality CXR.

CT imaging as a novel and highly promising approach to the detection and characterization of SSPN raises three types of questions:

1. How often does it lead to detection of SSPN ?
2. How often is such an SSPN malignant ?
3. If malignant, how often does resection etc. result in cure ?

The first question pertains to decisions about screening use of CT (clinical or mass screening), the second to decisions about CT screening for SSPN and decisions about further diagnostics (biopsy, etc.) given an SSPN detected by CT screening or incidentally, and the third to decisions about CT screening for SSPN, diagnostics given an SSPN, and intervention given a malignant SSPN.

Recognition that the pursuit of CT-detectable SSPN (screening) are pertinent only to persons with relatively high risk for lung cancer allowed us to develop a cost-effective paradigm for the evaluation of CT screening. This model allows us to determine, first, the probability of SSPN detection as a function of risk-relevant person characteristics among high-risk people, separately for baseline and repeat screenings. If these probabilities are P_0 and P_1 , respectively, then the probability of detecting an SSPN in the course of x years of annual screening is $1 - (1 - P_0)(1 - P_1)^x$. Furthermore, it allows us to determine the probability that the nodule is malignant and can be cured.

| | | |
|---------|--|---|
| 2:00 pm | Steven Skates Harvard University | Early detection of ovarian cancer with longitudinal marker levels |
|---------|--|---|

Abstract: Ovarian cancer has a high cure rate if detected in early stage, and a low cure rate if detected in late stage, which is commonly the case. Hence an early detection program is an appealing approach to reducing mortality from this disease. However, the low incidence of the disease presents a difficult hurdle for early detection programs, requiring a very high specificity and simultaneous high sensitivity to be effective. We demonstrate how the extra information in the longitudinal value of a marker can substantially improve the specificity and sensitivity simultaneously, and indicate methods whereby the longitudinal approach could be used in practice.

| | | |
|---------|--|---------|
| 3:00 pm | Marvin Zelen Harvard Sch. of Public Health | Summary |
|---------|--|---------|

Monday, July 28

Weeks 4 & 5: Design & Analysis of Clinical Trials. (July 28–August 8)
will be described in detail in IMA Newsletter 253.

Tuesday, July 29

Wednesday, July 30

Thursday, July 31

| |
|---------------------------------|
| CURRENT IMA PARTICIPANTS |
|---------------------------------|

POSTDOCTORAL MEMBERS FOR 1996-97 PROGRAM YEAR

| NAME | PREVIOUS INSTITUTION |
|--------------------|-----------------------------------|
| GOBBERT, MATTHIAS | Arizona State University |
| LOTOTSKY, SERGEY | University of Southern California |
| MALIASSOV, SERGUEI | Texas A&M University |
| NGUYEN, BRIAN | University of Michigan |
| NIE, QING | Ohio State University |
| SARKAR, SANHITA | University of Minnesota |
| SUCHOMEL, BRIAN | University of Wyoming |
| YANG, DAOQI | Wayne State University |

POSTDOCTORAL MEMBERSHIPS IN INDUSTRIAL MATHEMATICS FOR 1996-97

| NAME | PREVIOUS INSTITUTION | INDUSTRIAL AFFILIATION |
|--------------------|--------------------------|------------------------|
| CHAWLA, SANJAY | University of Tennessee | Honeywell |
| KOURITZIN, MICHAEL | Carleton University | Lockheed Martin |
| LOPEZ, GILBERTO | Northwestern University | Eastman Kodak |
| WANG, LEI | University of Washington | Honeywell |

VISITORS IN RESIDENCE (as of 6/9)

| | | |
|-----------------------|---------------------------------------|----------------|
| BARTKO, JOHN | National Institute of Mental Health | JUL 19 - 25 |
| BERRY, DON | Duke University | JUL 28 - AUG 8 |
| BROWN, STEPHEN | University of Hertfordshire | JUN 27 - JUL 3 |
| BRYAN, NICK | Johns Hopkins Hospital | JUL 13 - 18 |
| BURNS, JOHN | Virginia Poly. and State Univ. | JUN 27 - JUL 3 |
| CARLE, ALAN | Rice University | JUN 28 - JUL 3 |
| CHALONER, KATHRYN | University of Minnesota | JUL 28 - AUG 8 |
| CHANG, WON-JAE | Seoul National University | JUL 1 - 31 |
| CHARPENTIER, ISABELLE | Lab. de Model. et Calcul, Grenoble | JUN 27 - JUL 3 |
| CHEN, CHIN-TU | University of Chicago | JUL 13 - 18 |
| CHERNOFF, HERMAN | Harvard University | JUL 6 - 11 |
| CHRISTENSEN, RONALD | University of New Mexico | JUL 20 - 25 |
| COCKBURN, BERNARDO | University of Minnesota | SEP 1 - AUG 31 |
| COLEMAN, THOMAS | Cornell University | JUN 29 - JUL 2 |
| COX, ROBERT W. | Medical College of Wisconsin | JUL 13 - 18 |
| DAVIDOV, ORI | Hutchinson Cancer Rsch. Ctr., Seattle | JUL 28 - AUG 8 |
| DETRE, JOHN A. | Hospital of the Univ. of Pennsylvania | JUL 13 - 18 |
| DIXON, DENNIS | National Institutes of Health | JUL 28 - AUG 8 |
| DOBSON, DAVID | Texas A&M University | JUN 28 - JUL 3 |
| DORR, DAVID | Washington Univ. -St. Louis | JUL 6 - AUG 22 |
| EDDY, WILLIAM | Carnegie-Mellon University | JUL 13 - 18 |
| EFRON, BRAD | Stanford University | JUL 6 - 11 |
| EPSTEIN, LEONARDO | Battelle Memorial Institute | JUL 20 - 25 |
| EWENS, WARREN | University of Pennsylvania | JUL 5 - 11 |

| | | |
|--------------------------|--|----------------|
| FAURE, CRISTELE | INRIA | JUN 28 - JUL 3 |
| FEINGOLD, ELEANOR | Emory University | JUL 5 - 11 |
| FRIEDMAN, AVNER | Institute for Mathematics | SEP 1 - AUG 31 |
| GASTWIRTH, JOSEPH L. | George Washington University | JUL 20 - 25 |
| GEISSER, SEYMOUR | University of Minnesota | JUL 20 - 25 |
| GEISSER, SEYMOUR | University of Minnesota | JUL 7 - 11 |
| GENOVESE, CHRISTOPHER | Carnegie-Mellon University | JUL 13 - 18 |
| GEORGE, STEVE | Duke University | JUL 28 - AUG 8 |
| GEYER, CHARLES | University of Minnesota | JUL 7 - 11 |
| GIERING, RALF | Max-Planck-Inst. für Meteorologie, Hamburg | JUN 27 - JUL 3 |
| GOLDMAN, VICTOR | University of Twente | JUN 27 - JUL 3 |
| GRAMBSCH, PATRICIA | University of Minnesota | JUL 20 - 25 |
| GREENHOUSE, JOEL | Carnegie Mellon University | JUL 13 - 18 |
| GU, CHI | Washington University | JUL 5 - AUG 9 |
| GUCKENHEIMER, JOHN | Cornell University | JUN 27 - JUL 3 |
| GULLIVER, ROBERT | Institute for Mathematics | SEP 1 - AUG 31 |
| HALLORAN, ELIZABETH | Emory University | JUL 6 - 18 |
| HEINKENSCHLOSS, MATTHIAS | Rice University | JUN 27 - JUL 3 |
| HEJHAL, DENNIS | University of Minnesota | SEP 1 - AUG 31 |
| HENSCHKE, CLAUDIA | Cornell Univ. Medical College | JUL 20 - 25 |
| HERSKOVITZ, EDWARD | Johns Hopkins School of Medicine | JUL 13 - 18 |
| HOLMES, SUSAN | Cornell University-Biometry | JUL 6 - 11 |
| HORN, PAUL S. | University of Cincinnati | JUL 27 - AUG 8 |
| HU, PING | Dana-Farber Cancer Institute | JUL 28 - AUG 8 |
| IBRAHIM, JOSEPH | Harvard University | JUL 20 - 25 |
| JOHNSON, WESLEY | University of California-Davis | JUL 20 - 25 |
| KAFADAR, KAREN | University of Colorado-Denver | JUL 20 - 25 |
| KUBOTA, KOICHI | Chuo University | JUN 27 - JUL 3 |
| LANGE, NICK | McLean Hospital, Harvard | JUL 13 - 18 |
| LE, CHAP | University of Minnesota | JUL 20 - 25 |
| LEE, SANDRA | Harvard University | JUL 20 - 25 |
| LEWIS, ROGER J. | Univ. of California-Los Angeles | JUL 27 - AUG 8 |
| LI, WEN-HSIUNG | University of Texas | JUL 6 - 11 |
| LITTMAN, WALTER | University of Minnesota | SEP 1 - AUG 31 |
| LOUIS, THOMAS | University of Minnesota | JUL 28 - AUG 8 |
| LOWENGRUB, JOHN | University of Minnesota | SEP 1 - AUG 31 |
| LUCAS, ROH | Argonne National Labs | JUN 27 - JUL 3 |
| LUSKIN, MITCHELL | University of Minnesota | SEP 1 - AUG 31 |
| MANKE, JOE | Boeing Information/Support Serv. | JUN 28 - JUL 3 |
| MARGOLIN, BARRY | Univ. of North Carolina-Chapel Hill | JUL 27 - AUG 8 |
| MIGLIORETTI, DIANA | Johns Hopkins University | JUL 13 - 18 |
| MUELLER, LAURENCE | University of California-Irvine | JUL 6 - 11 |
| NAUMANN, UWE | University of Dresden | JUN 27 - JUL 3 |
| NEWTON, MICHAEL | University of Wisconsin-Madison | JUL 6 - 11 |
| NOLL, DOUG | Univ. of Pittsburgh Medical Center | JUL 13 - 18 |
| NORMAND, SHARON-LISE | Harvard University | JUL 20 - 25 |
| PAPANDONATOS, GEORGE | SUNY-Buffalo | JUL 20 - 25 |
| PARMIGIANI, GIOVANNI | Duke University | JUL 20 - 25 |
| PETTIGREW, KAREN | National Institutes of Health | JUL 20 - 25 |
| PETZOLD, LINDA | University of Minnesota | JUN 27 - JUL 3 |
| PRENTICE, ROSS | University of Washington | JUL 28 - AUG 8 |
| RACINE-POONE, AMY | Ciba-Geigy, Basel | JUL 28 - AUG 8 |
| RACKNER, BARRY | Minnesota Supercomputer Center | SEP 1 - AUG 31 |
| RAZ, JONATHAN | University of Michigan | JUL 13 - 18 |
| ROH, LUCAS | Argonne National Labs | JUN 27 - JUL 1 |
| SCHORK, NICOLAS | MetroHealth Med. Ctr., Cleveland | JUL 6 - 11 |

| | | |
|--------------------------------|------------------------------------|----------------|
| SCHUCANY, WILLIAM R. | Southern Methodist University | JUL 13 - 18 |
| SEILLIER-MOISEWITCH, FRANCOISE | University of North Carolina | JUL 6 - 11 |
| SELL, GEORGE | University of Minnesota | SEP 1 - AUG 31 |
| SHIH, JOANNA | Nat. Heart, Lung and Blood Inst. | JUL 20 - 25 |
| SHIH, JOE | Merck & Company, Inc. | JUL 20 - 25 |
| SIEGMUND, DAVID | Stanford University | JUL 6 - 11 |
| SIMES, JOHN | University of Sydney | JUL 28 - AUG 8 |
| SIMON, RICHARD | National Institutes of Health | JUL 28 - AUG 8 |
| SIMONSEN, KATY | North Carolina State Univ. | JUL 5 - 11 |
| SKATES, STEVEN | Harvard University | JUL 20 - 25 |
| SMALL, STEVEN | University of Maryland | JUL 13 - 18 |
| SPEED, TERENCE | Univ. of California-Berkeley | JUL 6 - 9 |
| STANGL, DALENE | Duke University | JUL 28 - AUG 8 |
| STOFFER, DAVID | University of Pittsburgh | JUL 13 - 17 |
| STROTHER, STEPHEN | University of Minnesota | JUL 13 - 18 |
| SUN, FENGZHU | Emory University | JUL 6 - 11 |
| ŠVERÁK, VLADIMIR | University of Minnesota | SEP 1 - AUG 31 |
| SYMES, WILLIAM | Rice University | JUN 27 - JUL 3 |
| TALAGRAND, OLIVIER | École Normale Supérieure | JUN 27 - JUL 3 |
| THULBORN, KEITH | Univ. of Pittsburgh Medical Center | JUL 13 - 18 |
| UGURBIL, KAMIL | University of Minnesota | JUL 14 - 18 |
| UTKE, JEAN | Argonne National Laboratory | JUN 28 -29 |
| UTTS, JESSICA | University of California, Davis | JUL 20 - 25 |
| VERMA, ARUN | Cornell University | JUN 27 - JUL 3 |
| WANG, MING-DAUH | University of Minnesota | JUL 21 - 25 |
| WEI, L. J. | Harvard University | JUL 28 - AUG 8 |
| WEIR, BRUCE | North Carolina State University | JUL 6 - 11 |
| WEISBERG, SANDY | University of Minnesota | JUL 7 - AUG 29 |
| WEISS, ROB | Univ. of California-Los Angeles | JUL 20 - 25 |
| WEISSKOFF, ROBERT M | MGH-NMR Center | JUL 13 - 18 |
| WIJSMAN, ELLEN | University of Washington | JUL 6 - 11 |
| WITTES, JANET | Statistics Collaborative, Inc. | JUL 20 - 25 |
| WOODS, ROGER | Univ. of California-Los Angeles | JUL 13 - 18 |
| WORSLEY, KEITH | McGill University | JUL 13 - 18 |
| YIN, GEORGE | Wayne State University | JUL 18 - 23 |
| YLVISAKER, DON | Univ. of California-Los Angeles | JUL 6 - 11 |
| YOUNG, STANLEY | Glaxo, Inc. | JUL 28 - AUG 8 |
| ZEGER, SCOTT | Johns Hopkins University | JUL 13 - 25 |
| ZELLEN, MARVIN | Harvard Sch. of Public Health | JUL 20 - AUG 8 |
| ZELTERMAN, DANIEL | Yale University | JUL 20 - 25 |
| ZHANG, CHAOMING | Rice University | JUN 27 - JUL 3 |