Statistical Practice in Cancer Conference Program March 3, 2017 SRB, Ted & Marty Couch Auditorium (unless otherwise specified)	
8:00am – 8:30am	**Registration & Refreshments** 1st Floor Atrium
SEGMENT 1: CLINICAL TRIALS **Facilitator: Brooke Fridley, Ph.D.**	
8:30am – 8:35am	**Welcome and Introduction of Keynote Speaker** Michael Schell, Moffitt Cancer Center
8:35am – 9:30am	**Keynote Speaker**: ***Clinical trials in the twenty-first century: Emerging issues and controversies*** Susan Ellenberg, University of Pennsylvania
9:30am – 10:30am	**Workshop: *Likelihood Methods in Clinical Trials*** Jeffrey Blume, Vanderbilt University
10:30am – 10:45am	**BREAK**
10:45am – 11:45am	**Clinical Trials Panel Discussion** Susan Ellenberg, University of Pennsylvania J. Jack Lee, MD Anderson Cancer Center Dan Sullivan, MD, Moffitt Cancer Center
11:45am – 12:15pm	**Poster Session** (1st Floor Atrium) 1. *An evaluation of bias in 2 stage residual inclusion survival models*, Fei Wan, University of Arkansas 2. *A Novel Statistical Tool for Differential Expression Analysis of NanoString nCounter Data*, Chi Wang, University of Kentucky 3. *Simon 2-stage design and the likelihood method*, Dan Ayers, Vanderbilt University **Software in Action** 4. *Software by Emory*, Yuan Liu, Emory University 5. *Continual reassessment method*, Nolan Wages, University of Virginia 6. The Role of R Packages in Reproducible Research, Shawn Garbett, Vanderbilt University 7. *Launching 80+ statistical software tools from Houston*, Clift Norris, MD Anderson
12:15pm – 1:00pm	**LUNCH (provided)** 1st Floor Atrium

SEGMENT 2: STATISTICAL SOFTWARE **Facilitator: Dung-Tsa Chen, Ph.D**	
1:00pm – 1:15pm	**Session 1**: ***Taking the heavy load of routine statistical analyses and generating a professional looking report: Introduction of a series of high-performance SAS macros in BBISR at Winship Cancer Institute***
1:15pm – 1:30pm	**Session 2**: ***A web application for conducting the continual reassessment method*** Nolan Wages, University of Virginia
1:30pm – 1:45pm	**Session 3**: ***The Role of R Packages in Reproducible Research*** Shawn Garbett, Vanderbilt University
1:45pm – 2:00pm	**Session 4**: ***Design Strategies for the Development of Robust and User-friendly Software*** Clift Norris, MD Anderson
2:00pm – 2:30pm	**Statistical Software Panel Discussion** Yuan Liu, Emory University Nolan Wages, University of Virginia Shawn Garbett, Vanderbilt University Clift Norris, MD Anderson
2:30pm – 2:45pm	**BREAK**
SEGMENT 3: STATISTICAL MISCELLANEA **Facilitator: Y. Ann Chen, Ph.D.**	
2:45pm – 3:15pm	**Session 5: Multi-state models, time-dependent covariates, and competing risks for the analysis of BMT studies** Gary Rosner, Johns Hopkins
3:15pm – 3:45pm	**Session 6**: ***Identifying Statistical Innovations Ready for Practice from the Literature*** Michael J. Schell, Moffitt Cancer Center Ji-Hyun Lee, University of New Mexico
3:45pm – 4:00pm	**BREAK**
4:00pm – 4:30pm	**Session 7**: ***Application and Development of Statistical Methods for Cancer Registry Data to Facilitate Population-Based Cancer Research*** Bin Huang, University of Kentucky
4:30pm – 5:00pm	**Session 8**: ***Staff Statistician Career Pathway Discussion*** Moderators: Ji-Hyun Lee, University of New Mexico Richie Reich, Moffitt Cancer Center

Statistical Practice in Cancer Conference Presenters

Conference Host & Organizer  **Michael J. Schell, PhD** Senior Member, Department of Biostatistics & Bioinformatics Scientific Director, Biostatistics Shared Resource H. Lee Moffitt Cancer Center & Research Institute	Dr. Schell is a Senior Member in the Biostatistics and Bioinformatics Department and Scientific Director of the Biostatistics Shared Resource at the Moffitt Cancer Center. He is also the Director of the Biostatistics Core for the Skin SPORE grant. He has over 30 years of experience in statistical analysis related to cancer research, with primary areas of expertise in clinical trials, flexible regression methods, and next-generation sequencing analysis. He authored "Identifying Key Statistical Papers From 1985-2002 Using Citation Data for Applied Biostatisticians", The American Statistician, 2010.
Keynote Speaker  **Susan S. Ellenberg, Ph.D.** Professor of Biostatistics Biostatistics & Epidemiology University of Pennsylvania	Dr. Ellenberg's research interests have focused on issues in the design and analysis of clinical trials, and assessment of medical product safety. Particular areas of interest include efficient trial designs, interim monitoring and the operation of data monitoring committees, evaluation of surrogate endpoints, ethical issues in clinical research, and special issues in trials of cancer and AIDS therapies, and of vaccines. She is an Associate Editor of Clinical Trials and of the Journal of the National Cancer Institute. Dr. Ellenberg is a Fellow of the American Statistical Association, the American Association for the Advancement of Science, the Society for Clinical Trials, and an elected member of the International Statistical Institute. She has served as President of the Society for Clinical Trials and the Eastern North American Region of the International Biometric Society, and has chaired the Statistics Section of the AAAS and the Board of Trustees for the National Institute of Statistical Sciences. Her book on clinical trials data monitoring committees, co-authored with Drs. Thomas Fleming (University of Washington) and David DeMets (University of Wisconsin), was named WileyEurope Statistics Book of the Year for 2002.

Statistical Practice in Cancer Conference Presenters Continues

Jeffrey D. Blume, PhD
Associate Professor of Biostatistics
Associate Professor of Biomedical Informatics
Director of Graduate Studies
Vanderbilt University

I am an Associate Professor in the Department of Biostatistics at Vanderbilt University School of Medicine. I am the founding Director of Graduate Studies in Biostatistics at Vanderbilt. I received my Ph.D. in Biostatistics from the John Hopkins School of Public Health, where I was the recipient of a National Eye Institute Traineeship in Clinical Trials. I am the leading expert in likelihood methods for measuring statistical inference and I publish on the foundations of statistical inference. Recent work has led to the development of a "second-generation p-value" that has significantly improved statistical and scientific properties. Throughout my career I have published broadly on the foundations of statistical inference, methodology for analyzing and interpreting receiver operating characteristic curves, clinical trials design and analysis, and neuroimaging. My collaborative experience is quite diverse, with a continuing emphasis on diagnostic and cancer trials, large-scale trials in emergency medicine, translational biomedicine, and radiologic and fMRI studies.

Yuan Liu, PhD, MS
Assistant Research Professor, Dept. of Biostatistics And Bioinformatics
Rollins School of Public Health Emory University
Assistant Research Professor,
Bioinformatics Shared Resource, Winship Cancer Institute of Emory University

Yuan Liu, PhD, is an Assistant Professor in the Department of Biostatistics and Bioinformatics at Rollin School of Public Health. She joined the Biostatistics and Bioinformatics Shared Resource at Winship Cancer Institute in January 2010. Dr. Liu has extensive collaborative experience with Winship investigators in the areas of breast cancer, lung cancer, pancreatic cancer, prostate cancer, head and neck cancer, hematology and radiation oncology. Her expertise includes study design, data management & analysis for retrospective studies, prognostic biomarker validation, survival/recurrence data analysis and propensity score approach.

Statistical Practice in Cancer Conference Presenters

Nolan Wages, Ph.D.
Assistant Professor,
Division of Translational Research & Applied Statistics
Department of Public Health Services,
University of Virginia

Dr. Wages is an Assistant Professor in the Division of Translational Research and Applied Statistics in the Department of Public Health Sciences. As a member of this division, the majority of his research effort is related to cancer research and its applications. Dr. Wages is an active member of the UVA Cancer Center Biostatistics Shared Resource, for which he works with cancer center members on the design and analysis of clinical trials. With funding from a NCI/NIH K25 Award, his current methodological research involves the design and analysis of early-phase clinical trials, with a particular focus on studies of combined immunotherapies.

Shawn Garbett
Sr. App Developer
Department of Biostatistics
Vanderbilt Unviersity

Modeling leads to key useful insights. Shawn studied engineering at Tulane University and Statistics at Penn State. Shawn came to Vanderbilt after contracting for the military on state models of computation. Before that he modeled for TVA on burn optimization and reservoir scheduling. Passionate about applied mathematics, and a desire for progressing science VUMC was a natural fit. In his free time, Shawn loves to play classical piano and hiking.

Statistical Practice in Cancer Conference Presenters

Clift Norris, Ph.D.
Mgr, Systems Analyst Services
Biostatistics Dept
MD Anderson Cancer Center

Clift Norris has been involved in technical software development for more than 25 years in the fields of engineering and medicine. He has been drawn to this field due to the opportunity to create tools that enable professionals to solve problems that would otherwise be difficult or impossible. With degrees in Mechanical Engineering and Computer Science he currently works with a team developing software for cancer research at MD Anderson Cancer Center in Houston, Texas.

Gary L. Rosner, ScD
Director of the Division of Biostatistics and Bioinformatics, Department of Oncology
E. K. Marshall Jr. Professor of Oncology
Johns Hopkins University School of Medicine
Professor of Biostatistics
Johns Hopkins Bloomberg School of Public Health

Dr. Gary Rosner, Professor of Oncology, is Director of the Division of Biostatistics and Bioinformatics, in the Department of Oncology, since joining the faculty at Johns Hopkins University in February 2010. Dr. Rosner leads the Biostatistics Shared Resource in the Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins. He received his Sc.D. in Biostatistics from the Harvard School of Public Health in 1985. He was a member of the faculties of Yale University, Duke University, and The University of Texas M. D. Anderson Cancer Center. He has been actively involved in collaborative research, teaching, and mentoring. Dr. Rosner has extensive experience in clinical research. In addition to collaborating with clinical researchers at his home institutions, he was a faculty statistician in the Statistical Office of the Cancer and Leukemia Group B (CALGB) for 12 years. Dr. Rosner's biostatistical research concerns population pharmacokinetics, pharmacodynamic modeling, pharmacogenetics, Bayesian methods, the analysis of repeated measurement and longitudinal data, clinical trial design, and survival analysis.

Statistical Practice in Cancer Conference Presenters

Ji-Hyun Lee, DrPH
Director, Biostatistics Shared Resource
University of New Mexico Comprehensive Cancer Center;
Professor of Biostatistics
University of New Mexico School of Medicine

Ji-Hyun Lee, DrPH, is a Professor in the Department of Internal Medicine at the University of New Mexico School of Medicine. She is Director of the Biostatistics Shared Resource and a full member of the Cancer Control Research Group at the UNM Comprehensive Cancer Center. Ji-Hyun is also 2017 President of Caucus for Women in Statistics (CWStat.org). She holds a doctorate of Public Health in Biostatistics from the University of North Carolina at Chapel Hill. Dr. Lee has extensive experience in overseeing biostatisticians and collaborating on numerous scientific studies. Her research interests include clinical trials; group randomized trials based on communities; methods for the analysis of observational data and repeated measurements; statistical methods in epidemiology; and best statistical practices.

Bin Huang, DrPH
Associate Professor,
Department of Biostatistics, College of Public Health
Markey Cancer Center, College of Medicine
University of Kentucky

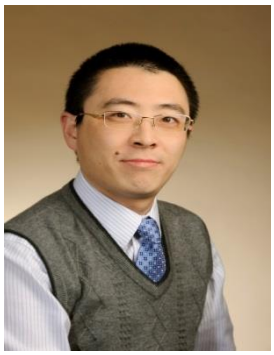
Dr. Bin Huang is an Associate Professor in the Department of Biostatistics and the Markey Cancer Center at the University of Kentucky. Dr. Huang's research focuses on incorporating novel statistical and epidemiological methodologies and study designs for population-based cancer research, particularly on utilization of cancer registry data. He has expertise in data linkage with administrative claims data, relative survival and life table and multiple imputations in missing data. He has extensive collaboration experience with clinicians, basic scientists, epidemiologist, behavior scientists and has been PI for several federally funded projects utilizing cancer registry data for cancer outcome research.

Statistical Practice in Cancer Conference Presenters

J. Jack Lee, Ph.D.
Associate Vice Provost of Quantitative Research,
Professor of Biostatistics
University of Texas MD Anderson Cancer Center.

Dr. Lee's areas of research interest include design and analysis of clinical trials, survival analysis, longitudinal data analysis, statistical computation/graphics, statistical methods for determining drug interaction in combination studies, and cancer chemoprevention. Dr. Lee has been working on the development and application of innovative Bayesian methods for cancer clinical trials. He also actively participates in many multidisciplinary translational research in head/neck and lung cancer teams funded by National Institute of Health (NIH) and the Department of Defense (DoD). He has particular interests in incorporating multiple biomarkers and adaptive designs to develop more efficient and ethical clinical trials.

Dr. Lee is a Statistical Editor for the *Journal of the National Cancer Institute* and *Cancer Prevention Research*. He is a Fellow of the American Statistical Association.



Chi Wang, Ph.D.
Associate Professor, Department of Biostatistics, College of Public Health and the Biostatistics and Bioinformatics Shared Resource Facility, Markey Cancer Center at the University of Kentucky

Dr. Chi Wang is an Associate Professor in the Department of Biostatistics, College of Public Health and the Biostatistics and Bioinformatics Shared Resource Facility, Markey Cancer Center at the University of Kentucky. Prior to joining the University of Kentucky, he was an Assistant Professor in the Department of Statistics at the University of California, Riverside. He received the B.S. and M.S. degrees in statistics from Peking University, and the Ph.D. degree in biostatistics from the Johns Hopkins University. His research focuses on developing statistical and bioinformatics methods for high-throughput genomic, transcriptomic, and metabolomic data. He is also interested in applying statistical methods to cancer-related research projects.

Statistical Practice in Cancer Conference Presenters

Fei Wan, Ph.D.
Professor, Department of Biostatistics
University of Arkansas for Medical Sciences

Fei Wan is assistant professor of the Department of Biostatistics at the University of Arkansas for Medical Sciences (UAMS) since 1996. Before joining UAMS, he was a faculty member at Group Health Research Institute in Seattle between 2015 and 2016. Dr. Wan received his Ph.D. in Biostatistics from the University of Pennsylvania in 2015. His research focuses on causal inference and survival analysis.

Dan Ayers
Senior Associate, Department of Biostatistics
Vanderbilt-Ingram Cancer Center

Dan Ayers is a Senior Associate in Biostatistics at the Vanderbilt-Ingram Cancer Center and a member of the Center for Quantitative Sciences at Vanderbilt. He was previously at three other NCI-designated cancer centers: St. Jude Children's Research Hospital, The University of Arkansas Cancer Center, and the MD Anderson Cancer Center. Dan is an expert in clinical trials analyses.

Statistical Practice in Cancer Conference Presenters

Richie Reich, Ph.D.
Biostatastics Core Faculity Manager
Moffitt Cancer Center

I have been the Biostatistics Core Facility Manager at Moffitt Cancer Center since May, 2016. He received his PhD in Experimental Psychopathology at USF in 2002 and was a faculty member in the College of Arts and Sciences at USF Sarasota-Manatee from 2005-2016. While a faculty member, I began working as a biostatistician for Moffitt Cancer Center in 2008 and Sarasota Memorial Hospital in 2014. This biostatistics experience, coupled with my leadership experience as a faculty member, led to my current position. I previously have published on numerous psychology and cancer-related studies using multiple methods including growth curve models, structural equation modeling, multi-dimensional scaling, cluster analysis, and meta-analysis.

Dan Sullivan, Ph.D.
ACD Clinical Science
Moffitt Cancer Center

Dr. Sullivan has over 20 years of experience in blood and marrow transplantation and the treatment of patients with multiple myeloma, and sees patients with multiple myeloma in the out-patient clinics at Moffitt. He has chaired and served on several NCI and ACS grant review panels and committees, and currently serves on the ACS Council for Extramural Grants, NCI NExT Special Emphasis Panel, and the NCI Investigational Drug Steering Committee. He has significant experience in the writing and conduct of early-phase translational clinical trials and his clinical interests are in novel approaches to treat multiple myeloma. Dr. Sullivan's basic research activities are in defining the role of nuclear-cytoplasmic trafficking of proteins and how this determines multiple myeloma cell drug sensitivity (currently supported by R01 CA194051-01). He also has significant experience in mentoring junior faculty—through oversight of the Moffitt Physician-Scientist Mentorship Program, as PI of the NCI N01 Phase 2 contract for the Southeast Phase 2 Consortium for the past 10 years, as the Moffitt Site PI of the Princess Margaret Phase I Consortium UM1 grant, and as MPI and Career Development and Mentoring Director of the NCI U54 PHSU-MCC Partnership. Finally, Dr. Sullivan is also the Chief Medical Officer of M2Gen, and in this capacity is deeply involved in precision medicine trials for patients with cancer.

Workshop, Jeffrey D. Blume, Ph.D., Vanderbilt University

Likelihood Methods in Clinical Trials

In a classic 1966 paper on sequential trials and the likelihood principle, Jerome Cornfield explained why sequential analyses ought to be rooted in the likelihood principle. However, fifty years later, the vast majority of sequential analyses employed in practice violate this principle. In this talk, I'll explain why the direct interpretation of the likelihood function, as metric of statistical evidence, adheres to both the likelihood principle and classical frequentist error conservation principles. Because of this, likelihood methods are naturally suited for the analyses of clinical trials. Key is the ability of the likelihood framework to distinguish between three evidential quantities: (1) a measure of the strength of evidence, (2) the probability that a particular study design will generate misleading or weak evidence, and (3) the probability that observed evidence is misleading. I will illustrate the likelihood approach, and the role of these evidential quantities, in the context of a few simple examples and re-analyses of well-known clinical trials.

Poster Session, Fei Wan, PhD, University of Arkansas

An evaluation of bias in 2 stage residual inclusion survival models

Unmeasured confounding is a common concern when cancer researchers attempt to estimate a treatment effect using observational database. To address this concern, the two-stage residual inclusion (2SRI) model, an instrumental variable (IV) based method, has been widely accepted as the method of choice in many clinical studies involving cancer patients' survival outcomes. However, a compelling theoretical rationale has not been postulated. We propose a novel two-stage modeling framework to for understanding the bias and consistency in estimating the conditional treatment effect for 2SRI when the outcome is time to event.

Under this framework, we demonstrate that the bias in the 2SRI estimator can be reframed to mirror the problem of omitted variables in non-linear models and that there is a direct relationship with the collapsibility of effect measures. We demonstrate that the 2SRI estimates are generally unbiased for Cox models, only if the influence of the unmeasured covariates on the treatment is proportional to their effect on the outcome. We also propose a novel dissimilarity metric to quantify the difference in these effects and demonstrate that with increasing dissimilarity between the effects of the unmeasured covariates on the treatment versus outcome, the bias of 2SRI increases in magnitude.

Poster Session Chi Wang, PhD, University of Kentucky

A Novel Statistical Tool for Differential Expression Analysis of NanoString nCounter Data

Hong Wang, Tingting Zhai, Craig Horbinski, Hao Wu, Yinxing Liu, Shaoyi Sheng, Jinpeng Liu, Isaac Hands, Eric B. Durbin, Heidi Weiss, Arnold J. Stromberg and Chi Wang

Background: The advanced medium-throughput NanoString nCounter technology provides a novel approach for miRNA profiling and targeted mRNA expression analysis and has been increasingly used for biomarker discovery for early detection, prognosis and prediction. Compared to microarray and RNAseq, it has shown superior applicability to formalin fixed paraffin embedded samples, which are major sources of clinical samples but often have less-than-ideal quality in terms of small tissue amount and degraded RNA. However, to date there are limited and sub- optimal methods for the differential expression (DE) analysis of nCounter data. Specifically, most of the current methods are based on t-

tests, which do not fit the count data generated by the NanoString nCounter system and data normalization procedures of current methods are either not suitable for counts or not specific for NanoString nCounter data.

Methods: We have developed a novel DE detection method by considering a generalized linear model of the negative binomial family to characterize count data and allows for multifactor design. Data normalization is incorporated in the model framework through data normalization parameters, which are estimated from positive controls, negative controls and housekeeping genes embedded in the nCounter system. We have proposed an empirical Bayes shrinkage approach to estimate the dispersion parameter in the model and a likelihood ratio test to identify differentially expressed genes. Our method has been implemented in an open source R package, NanoStringDiff, which is available at http://bioconductor.org/. We also have developed a user-friendly web application, NanoStringDiffWeb, for this package based on the R Shiny web framework, which is available at MarkeyBiostatTools.uky.edu.
Results: Simulation studies demonstrate that NanoStringDiffprovides more accurate and powerful results in DE detection compared to existing methods. Applying NanoStringDiffto a human glioma cell line data set identified 14 DE microRNAs comparing IDH1 mutant and GFP control groups at 1% false discovery rate, the top 5 of which have been validated using Q-PCR. The web interface, NanoStringDiffWeb, allows the user to perform the analysis by a few mouse clicks and requires very minimum user's input.
Conclusions: We have developed a comprehensive and general framework to characterize NanoString nCounter data and to detect DE genes for both simple and complex experimental designs. The NanoStringDiff and NanoStringDiffWeb provide tools for easy implementation by core facility statisticians as well as biomedical researchers with different levels of programming and biostatistics skills.

Session 1, Yuan Liu, PhD, Emory University

Taking the heavy load of routine statistical analyses and generating professional looking report: introduction of a series of high-performance SAS macros in BBISR at Winship Cancer Institute

Yuan Liu, Chao Zhang, Yaqi Jia, Dana Nickleach, Sungjin Kim, Jeffrey Switchenko, Jeanne Kowalski

Biostatistics/Bioinformatics Shared Resource (BBISR) at a Cancer Institute usually is a highly collaborative and fast-paced research facility. For a typical analytical project, it is required to generate descriptive statistics, univariate analysis, and/or multivariable model for researchers to reach a comprehensive understanding about the data on hands before they head to the more advanced methodologies for specific issues. Generating reports for those routing statistical analyses could be a time-consuming and tedious work, especially when dealing with a large database with a big number of variables. To speed up our turn-around time for projects and enhance the communication with the clinical investigators, we have developed a series of high-performance SAS macros that can generate interpretable and professional looking summary report in a timely and reproducible way, and they are also freely available in our website. Recently, they are routinely used in the retrospective research projects based on National Cancer Data Base (NCDB) at our shared resource, and help us process a large amount of projects in a short time period for ASCO and ASTRO deadlines. In addition, using those macros help improve the overall integrity of programming, which enable much easier maintenance and update as projects evolve. As for the entry level biostatisticians, such as master level biostatisticians or students, those macros ensure them an easy integration into research projects and learn and grow as they go. In Table 1, a list of selected SAS macros and their functions are presented, and in our tool box we have created about 15 SAS macros, and we keep them alive by upgrading and adding new ones or new features. Table 2 illustrates a result summary table by %PHREG_SEL, a multivariable Cox model by backward elimination. Figure 1 shows Kaplan-Meier curves and related statistics by %KM_PLOT. We sincerely hope this could be a significant contribution and benefit the community as a whole.

Session 2, Nolan A. Wages, PhD, University of Virginia

A web application for conducting the continual reassessment method

Broad implementation of model-based dose-finding methods has been limited, with traditional or modified 3+3 designs remaining in frequent use. Part of the reason is due to the lack of reliable, easy-to-use, and robust software tools for conducting more efficient designs. With the aim of augmenting broader implementation of model-guided methods, we have developed an R shiny web application for the continual reassessment method (CRM). The application has the ability to generate simulated results, as well as to be employed throughout the study to sequentially recommend a dose for each new accrual based on the current data. At the conclusion of the study, it can be used to determine the maximum tolerated dose (MTD). Therefore, it is able to be utilized for both the design stage and for direct protocol implementation. The web application requires no programming knowledge, and it is free to access on any device with an internet browser. We hope the development of this software will facilitate more efficient collaborations within study teams conducting single-agent dose-finding trials.

Session 3, Shawn Garbett, MS, Vanderbilt University Medical Center

The Role of R Packages in Reproducible Research

A look at why the open source statistical language R and R packages are important for reproducible research. Case studies demonstrating the value added will be explored. Examples and ideas of integrating into research workflows will be provided.

Session 4, Clift Norris, Ph.D., MD Anderson

Design Strategies for the Development of Robust and User-friendly Software

The Department of Biostatistics at the University of Texas MD Anderson Cancer Center has been among the world leaders in developing methods for novel clinical trial design, conduct, and analysis. In addition to innovations in methodology, we have developed computer software for learning, designing, and implementing statistical techniques. All of the tools are freely available to be downloaded or used online with a web browser. In this talk, the main objective is to discuss the strategies and challenges encountered when developing statistical software that is both statistically robust and user-friendly. We will discuss the techniques and trade-offs involved in building software, and demonstrate some of the tools we have published. These are some issues we will address: (1) What makes a graphical user interface 'easy to use'? (2) Why is there always a balance required between Time-of-development and Robustness? (3) Why does software cost so much to develop? (4) How should we developers manage rapid release schedules and increasing workload? (5) How should we increase the visibility and usage of software? (6) How can software be used to enhance collaboration? (7) How do we provide support for software far into the future? Many of the published programs may be found at:
https://biostatistics.mdanderson.org/softwaredownload and
https://biostatistics.mdanderson.org/softwareonline.

Session 6, Michael J. Schell, PhD, Moffitt Cancer Center
Ji-Hyun Lee, DrPH, University of New Mexico

Identifying Statistical Innovations Ready for Practice from the Literature

A significant challenge in growing and maintaining a robust statistical practice is the difficulty of identifying the best statistical innovations. Using the Web of Science, we have identified all 685 articles that have had 25 or more citations in 2014-2015 from 12 statistical journals that were identified as most important to biostatisticians in a survey in 2003. We propose the estimated fraction of theoretical and applied use citations that can be obtained using Web of Science journal categories. For easier access to the articles of particular value to a given practitioner, the papers have been categorized into 7 domains and about 40 sub-domains. The sub-domains contain an average of 17 articles, and a maximum of 37. The list of these identified articles will become available through a sortable Excel spreadsheet.

Session 7, Bin Huang, DrPH, University of Kentucky
Application and Development of Statistical Methods for Cancer Registry Data to Facilitate Population-Based Cancer Research

Bin Huang, Li Chen, Quan Chen, Brent Shelton, Heidi Weiss, Tom Tucker
Markey Cancer Center, University of Kentucky

State center cancer registries can be a valuable resource for various cancer prevention and control research initiatives. As one of the SEER cancer registries, the Kentucky Cancer Registry (KCR) is an integral component of the Markey Cancer Center's Cancer Prevention and Control (CP) Scientific Program and provides unique opportunities to conduct population-based research. We present several motivating research projects from the CP program based on cancer registry data and the novel statistical methods utilized to support these projects. Specifically, flexible spline Poisson regression models were used to generate life tables and improve relative survival estimates; point and interval estimates were generated to develop composite cancer burden indices; a new kernel-smoothed nonparametric method was developed to estimate and compare the distribution of exposure data or biological markers that present with detection limits; correcting the ovary cancer incidence rates by adjusting the background risk population based on oophorectomy prevalence rates; and improving the cancer registry data through ancillary data linkage. Faculty statisticians from the core facility have successfully received funding from the CDC and NCI for some of these methods or these methods have been utilized in funded grants by CP investigators. Furthermore, some of these methods have been recognized nationally and are utilized by other state and national cancer prevention and control communities. For example, the method for the composite cancer burden index has been implemented on the NCI Ci*rank website and methods to create lifetables have been used to create US SES lifetables by SEER. We will discuss these methods and demonstrate the tools developed to facilitate the application of these methods to cancer registry data.

Session 8, Ji-Hyun Lee, DrPH, University of New Mexico
Richard Reich, PhD, Moffitt Cancer Center
Staff Statistician Career Pathways in a Modern Cancer Center

The career of statistician was recently named the fourth best job of 2017 by *US News and World Report* based on factors such as salary and unemployment rate. It was the number one STEM job and the number one business job. In spite of an increasing supply of university trained statisticians, companies report increasing demand for, and difficulty finding of, qualified statisticians (Society for Human Resource Management, 2016). In such a hot job market, understanding the career interests and needs of statisticians are important to recruit, train, and increase job satisfaction in these individuals. The purpose of this discussion is to introduce

topics of importance for non-faculty track biostatisticians working in Cancer Centers. Preliminary data from a survey administered in advance of this meeting will provide a launching point for the discussion. With an initial sample size of 14, this survey pointed to the importance of relationships with immediate supervisors, collaborators, scientific directors, and fellow staff members, respectively. Also, salary/benefits and interesting work emerged as the most important factors for job satisfaction. The importance of additional factors such as promotion tracks and quality training/mentorship also were recognized. This data, coupled with qualitative information from the discussion, could be used to guide the management of biostatistics facilities in Cancer Centers.

Hotels w/Moffitt Rate & Complimentary Transportation to Moffitt's Campus

1. **Embassy Suites USF** (1.2 miles away from conference)
   $169/night (ask for Moffitt rate-not a guarantee, book ASAP)
   3705 Spectrum Boulevard
   Tampa, FL 33612
   (813) 977-7066

2. **Residence Inn Tampa North** (3.5 miles away from conference)
   $124/night (ask for Moffitt rate-not a guarantee, book ASAP)
   13420 Telecom Pkwy N
   Tampa, FL 33637
   (813) 972-4400
   *Will need to provide the conference program for this rate

3. **TownePlace Suites Tampa North Marriott** (3.3 miles away conference)
   $99/night (ask for Moffitt rate-not a guarantee, book ASAP)
   6800 Woodstork Road
   Tampa, FL 33637
   (813) 975-9777

Other Hotels in the Area (These locations do NOT offer transportation to Moffitt)

1. **Wingate by Wyndham at USF** (1.7 miles away from conference)
   Starts at $169/night
   3751 E Fowler Ave
   Tampa, FL 33612
   (813) 979-2828

2. **Courtyard Marriott Tampa North** (4.2 miles away from conference)
   Starts at $158/night
   13575 Cypress Glen Ln
   Tampa, FL 33637
   (813) 978-9898