

**Society for Prevention Research  
27<sup>th</sup> Annual Meeting  
San Francisco, CA**

**Pre-Conference Workshop II**

**Date:** Tuesday, May 28, 2019

**Time:** 8:30 am – 5:30 pm

**Novel Approaches for Public Health Informatics: Integrating and Analyzing Geospatial and Administrative Data** (*To register select: Health Informatics*)

**Underwritten by:** NIH/Office of Disease Prevention

**Organizers/Co-Chairs:**

- Elizabeth M. Ginexi, Ph.D., Health Scientist Administrator, National Institutes of Health
- Mildred M. Maldonado-Molina, Ph.D., Associate Professor, Health Outcomes & Biomedical Informatics, University of Florida

**Presenters:**

- Andrew J. Curtis, Ph.D., Professor, Department of Population & Quantitative Health Sciences, Case Western Reserve University School of Medicine
- Hye-Chung Kum, Ph.D., Associate Professor, Health Policy & Management, Texas A & M University
- Jaclyn Hall, Ph.D., Assistant Research Scientist, Health Outcomes & Biomedical Informatics, Faculty, Institute for Child Health Policy, University of Florida

**Description:**

Surveys, interviews, and experiments traditionally have been the primary means of data collection for prevention science research studies. We now have data from a number of sources that, while often collected with purposes other than research in mind, can be repurposed to inform studies of human behavior with unprecedented levels of fidelity and precision. These data sources include geospatial or geocoded sources of data and web-based electronic administrative or commercial record keeping databases, such as digital health care administrative data from patients, providers, and insurers. Advances in automated data processing including algorithms from machine learning, text mining, data mining, artificial intelligence and natural language processing, and innovations in computational modeling all present possibilities for utilizing these data sources for public health monitoring, surveillance, or even as platforms for targeting at risk populations with tailored health messaging and personalized real-time interventions and digital experiments.

The technological advances enabled by the creation of geospatial and administrative data streams and context-aware systems and computational modeling are now the focus of an exciting line of *public health informatics* research. Interdisciplinary digital health intervention research which brings together computer scientists and prevention researchers offers enormous potential to improve public health with broader, more rapid intervention reach for individuals and populations. Even traditional data collection approaches, such as collecting interviews can now benefit from new mobile geospatial technologies and natural language processing (NLP) technologies and associated software. This preconference workshop will convene experts to showcase innovative methodological approaches to curate, integrate, and analyze geospatial and administrative data. In addition, various research challenges, such as issues of data privacy, data sharing between collaborators, and more general

deficiencies and limitations will be addressed. Following didactic morning and afternoon tutorials, workshop attendees will participate in a moderated roundtable dialogue to discuss how these data methods may be applied and refined for future prevention science studies.

This workshop will be accessible to prevention scientists with and without exposure to geospatial and administrative data methodologies. The workshop will be educational, but very interactive with ample opportunity for open discussion and for participants to ask questions about incorporating these methods into prevention science research.

Specific learning objectives:

- Understand the unique capabilities of complex geospatial and digital administrative data for public health inquiry and how they complement and extend, but do not replace traditional methods.
- Understand what types of prevention research questions may be appropriate for studying with integrative administrative or geospatial data.
- Become familiar with and understand the main considerations and challenges inherent in the design, delivery, and data analysis involving such data sources.
- Participate in interactive discussions regarding the gaps and opportunities for future prevention science studies in the digital age.

This workshop will be accessible to prevention scientists with and without exposure to geospatial and administrative data methodologies. The workshop will be educational, but very interactive with ample opportunity for open discussion and for participants to ask questions about incorporating these methods into prevention science research.

**Target Workshop Audience:**

Prevention Scientists, Prevention Researchers, and Graduate Students studying prevention science who are either designing longitudinal studies aimed at evaluating the effects of prevention strategies on selected outcomes or currently conducting data collection for these types of longitudinal studies. Having a basic background in research methodology is sufficient but not required.

**Materials to be provided to Attendees:**

- Copies of the workshop slides
- A comprehensive bibliography for further reading on RSD/ASD
- An RSD/ASD checklist to guide practical / operational decisions
- Access to the web site for an NIH-funded education program on RSD, containing additional resources and discussion forums

**Workshop Agenda:**

8:30-9:00 Welcome & Introductory Session

*Leveraging Geospatial and Administrative Data for Prevention Research:*

*Data Types & Sources, Infrastructure/Expertise Required, Examples of Existing Projects Funded by the National Institutes of Health*

**Elizabeth Ginexi, Ph.D.**, National Institutes of Health &  
**Mildred M. Maldonado-Molina, Ph.D.**, University of Florida

9:00-10:30 Methods Overview Session: Geospatial Data

*Contextualizing Geospatial Data to Study Health Vulnerability in Community Settings*

**Andrew J. Curtis, Ph.D.**, Professor, Case Western Reserve University School of  
Medicine, Department of Population & Quantitative Health Sciences

10:30-10:45 Morning BREAK

10:45-11:30 Applications Session: Geospatial Data

*Research Projects Utilizing Contextualized Geospatial Data to Study Health in  
Community Settings*

**Andrew J. Curtis, Ph.D.**, Professor, Case Western Reserve University School of  
Medicine, Department of Population & Quantitative Health Sciences

11:30-12:30 Lunch BREAK

12:30-2:00 Methods Overview Session: Administrative Data

*Population Informatics: Applying Data Science to Big Data about People to Advance  
Prevention Research*

**Hye-Chung Kum, Ph.D.**, Associate Professor, Health Policy & Management, Computer  
Science & Engineering, Texas A & M University

2:00-2:15 Afternoon BREAK

2:15-3:00 Applications Session: Administrative Data

*Research Projects Utilizing Administrative Data Linkages to Study Chronic Disease  
Management using Remote Monitoring and Cancer Surveillance*

**Hye-Chung Kum, Ph.D.**, Associate Professor, Health Policy & Management, Computer  
Science & Engineering, Texas A & M University

3:00-4:30 *Moderated Roundtable Dialogue* - Interactive discussion with the full panel of speakers  
and workshop attendees to discuss prevention applications

**Moderators:**

**Mildred M. Maldonado-Molina, Ph.D.**, University of Florida

**Jaclyn Hall, Ph.D.**, University of Florida

4:30-4:45 Concluding Remarks

**Mildred M. Maldonado-Molina, Ph.D., & Elizabeth Ginexi, Ph.D.**

4:45-5:00 Complete feedback survey & Adjourn

**Co-chairs**

**Elizabeth M. Ginexi, Ph.D.** is a Health Scientist Administrator at the National Institutes of Health (NIH) Office of Behavioral and Social Sciences Research (OBSSR) where she is focused on the application of innovative research methodologies, measurement, and analytic approaches to advance behavioral and social sciences research. Dr. Ginexi is an Applied Social Psychologist with expertise in family- and community-based etiology, prevention, and treatment research; policy interventions to target population-level health behavior; and quantitative analysis, statistical and computational modeling. Prior to joining OBSSR she served as a Program Director in the Tobacco Control Research Branch (TCRB) at the National Cancer Institute (NCI) from 2010-2016 and as a Health Scientist Administrator in the Prevention Research Branch (PRB) at the National Institute on Drug Abuse (NIDA) from 2003-2010. Prior to NIH Dr. Ginexi was a Senior Study Director at in the Substance Abuse Research Group at Westat. She began her research career as a Postdoctoral Fellow and Research Associate at the George Washington University, where she helped with data collection, data management, and analysis on longitudinal field studies involving family based mental health etiology studies and preventive interventions.

**Mildred M. Maldonado-Molina, Ph.D.** is Associate Professor in the College of Medicine, Department of Health Outcomes and Biomedical Informatics at the University of Florida (UF). She is affiliated with the UF Institute for Child Health Policy, and Director of the Family Data Center. Dr. Maldonado-Molina's research program focuses on reducing health disparities in alcohol and substance use among youth and improving maternal health and child birth outcomes. She has published over 75 peer-reviewed studies and is currently the PI in four projects. She is also Co-Investigator in two NIH projects, including a P50 Center grant from the Pennsylvania State University entitled "Center for Complex Data to Knowledge in Drug Abuse and HIV Behavioral Science" and an R25 project entitled "A Master Course on Power for Multilevel and Longitudinal Health Behavior Studies". From 2009-2014, she was the PI of a five-year study from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) entitled "*Alcohol Contextual Influences: Health Disparities and Mortality*". She has also been principal, co-principal investigator, or co-investigator in nearly 30 externally-funded projects. She is past Director of Education and Training at the Department of Health Outcomes and Policy, and a member of the Interdisciplinary Program in Biomedical Sciences and the MD/PhD Advisory Boards in the College of Medicine at the University of Florida. Her work has been recognized by several scientific organizations, and she received the 2009 New Investigator Award from the National Hispanic Science Network and the 2009 Early Career Prevention Network (ECPN) Award from the Society for Prevention Research for her contributions in the area of statistical applications in the drug use and prevention field. In 2011, 2013, and 2015 she received the UF College of Medicine's Exceptional Teacher Awards for outstanding teaching accomplishments; and in 2013, the Distinguished Alumni Award from The Methodology Center at Penn State for significant science contributions to the field of drug use.

**Speakers**

**Andrew Curtis, Ph.D.** is professor, Case Western Reserve University School of Medicine, Department of Population & Quantitative Health Services. His work employs geospatial technologies and geographic information system (GIS) analysis to support neighborhood scale intervention strategies designed to

reduce health disparities. In 2005 after the landfall of Hurricane Katrina, he and his students were part of the academic team that helped with geospatial support for search and rescue operations in the Louisiana Emergency Operations Center. His geospatial recovery work continues in New Orleans and the post-tornado landscapes of Tuscaloosa, Alabama and Joplin, Missouri. In the summer of 2012 in Haiti he developed fine scale field mapping strategies to assess water risks in urban Cholera hotspots. Examples of his health-related GIS consultation include the Department of Public Health in Los Angeles County, diabetes clinics, non-profit organizations and ground-level community groups. He is also recognized as a leader regarding spatial privacy.

**Hye-Chung Kum, Ph.D.** is an Associate Professor in the Department of Health Policy and Management (HPM) with joint appointments in the Departments of Computer Science and Engineering, and Industrial Systems Engineering at Texas A&M University. She is a data scientist cross trained in computer science (PhD in datamining) and Welfare Policy and Management (Master of Social Work) at University of North Carolina at Chapel Hill. For over 15 years, she has been conducting research on methods and tools to use big data about people to answer questions about population health and well-being with a focus on using integrated government administrative data. She is the founder and director of the Population Informatics Lab and have collaborated on diverse areas of research in computer science (information security, information privacy, natural language processing (NLP), Human Computer Interaction (HCI), machine learning), health services research, health informatics, welfare policy, nutrition, epidemiology, and ELSI (ethical, legal, and social implications). She has over 50 publications in diverse journals and her research has been funded by NSF, NISOH, PCORI, North Carolina Division of Social Services, and Texas Health and Human Services Commission. She is a 2018 presidential impact fellow of the Hagler Institute of Advanced Study and a Royster Fellow.

**Jaclyn Hall, Ph.D.** is an Assistant Research Scientist in the College of Medicine, Institute for Child Health Policy at the University of Florida. She is trained as a spatial ecologist and has almost two decades of experience conducting research on spatial questions using geospatial technologies. She is experienced in using environmental and demographic datasets to examine how humans interact with the natural environment. Dr. Hall contributes to a variety of health policy related projects, including geocoding, mapping, creation of new spatial data sets, cluster analysis, study group delineation. In addition to facilitating the spatial analysis needs of projects and research of the Department of Health Outcomes and Policy and the Institute for Child Health Policy, Dr. Hall is developing a research program that investigates the pattern and process of homogenization of neighborhoods within heterogeneous cities and the implications for health outcomes and using environmental data to understand health outcomes.