CCBAR

Chicago Core on Biomarkers in Population-Based Aging Research at the University of Chicago - NORC Center on Aging

Chicago Workshop on Biomarkers in Population-Based Health and Aging Research

Sponsored by the
Chicago Center on Demography and Economics of Aging
and the
National Institute on Aging
June 8 & 9, 2006

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The 4th annual Chicago Conference on Biomarker Collection in Population-Based Health and Aging Research was organized and coordinated by Dr. Stacy Tessler Lindau, Dr. Natalia Gavrilova, and Dr. Lis Nielsen, with assistance from Alicia Fraser, Adelle Hinojosa, Kathleen Parks, Karl Mendoza, and many others.

Financial support for the 2006 Workshop on Biomarkers in Population-Based Health and Aging Research and the Workshop Proceedings came from the Chicago Center on Demography and Economics of Aging through its grant # 2 P30 AG012857, which was awarded by the Office of Behavioral and Social Research, National Institute on Aging.

The following proceedings were prepared from a transcript. Please bear in mind the informal nature of the workshop; participants' presentations and that comments should not be used without permission. Special thanks to Karl Mendoza, Syed Saad Iqbal, Bhairavi Nallanathan, and Ada Gomero for their assistance in the transcript preparation.

ISBN 0-9769978-3-5

Suggested Citation

Lindau S, Gavrilova N (eds). Proceedings of the 2006 Chicago Workshop on Biomarkers in Population-Based Aging Research. NORC - University of Chicago Center on Demography and Economics of Aging. Chicago, IL. 2007.

WORKSHOP SPEAKERS

David Almeida, PhD	Associate Professor, Human Development and Family Studies, Penn State University.
Duane Blackburn, MS	Agency Representative to the National Science & Technology Council, Federal Bureau of Investigation (Investigative Technology Division)
C. Sue Carter, PhD	Professor of Psychiatry, Co-Director of The Brain Body Center, University of Illinois at Chicago
Emil Coccaro, MD	Director of the Clinical Neuroscience and Psychopharmacology Research Unit, University of Chicago; Professor and Chairman, Department of Psychiatry, University of Chicago
Cathy Cohen, PhD	Professor of Political Science at the University of Chicago and former Director of the Center for the Study of Race, Politics and Culture
Joel Frader, MD	Professor of Pediatrics and Professor of Medical Humanities and Bioethics, Northwestern University
Alicia Frasier, MPH	Analyst, Chicago Core on Biomarkers in Population-Based Aging Research at the Center on Demography and Economics of Aging, NORC, University of Chicago
Thomas Fisher, MD	Emergency Department, University of Chicago Hospitals
Natalia Gavrilova, PhD	Research Associate, Department of Obstetrics and Gynecology, University of Chicago; Research Associate, Chicago Core on Biomarkers in Population-Based Aging Research at the Center on Demography and Economics of Aging, NORC, University of Chicago
Markus Heinrichs, PhD	Assistant Professor, Department of Psychology, University of Zurich, Switzerland
Donald Ingram, PhD	Chief of Behavioral Neuroscience Section, Gerontology Research Center, National Institute on Aging, NIH
Todd Lee, PharmD, PhD	Research Assistant Professor, Feinberg School of Medicine, Northwestern University; Senior Investigator, Midwest Center for Health Services and Policy Research
Stacy Tessler Lindau, MD, MAPP	Director, Chicago Core on Biomarkers in Population-Based Aging Research; Assistant Professor, Obstetrics and Gynecology and Medicine-Geriatrics, University of Chicago

Dario Maestripieri, PhD	Associate Professor, Comparative Human Development and Committee on Evolutionary Biology, Biological Sciences Collegiate Division, University of Chicago; Fellow of the Brain Research Institute, University of Chicago
Thomas McDade, PhD	Assistant Professor, Department of Anthropology, Northwestern University**;** Director, Laboratory for Human Biology Research, Northwestern University
Lis Nielsen, PhD	Individual Behavioral Processes Branch of NIA's Behavioral and Social Research (BSR) Program
Ryne Paulose, PhD	Epidemiologist, Division of National Health and Nutrition Examination Surveys, National Center for Health Statistics, CDC
George Preti, PhD	Researcher, Monell Chemical Senses Center
Dima Qato, PharmD, MPH	Consultant, National Social Life, Health and Aging Project (NSHAP), NORC, University of Chicago
Robert Wallace, MS, MD	Professor, Department of Epidemiology and Internal Medicine, University of Iowa; Co-Director, Center on Aging, Carver College of Medicine
Frank Wilhelm, PhD	Professor, Department of Health Psychophysiology, University of Basel, Switzerland; Head of the Clinical Psychophysiology Laboratory, Institute for Psychology, University of Basel, Switzerland
David Williams, PhD	Professor, Department of Sociology; Professor, Department of Epidemiology, School of Public Health; Director of South Africa Initiatives Office, Center for Afro-American and African Studies, University of Michigan
Wendy Wolf, PhD	Director and Co-Investigator of the NUgene Project, Center for Genetic Medicine (CGM), Northwestern University

PREFACE

Sponsored by the Chicago Core on Biomarkers in Population-Based Aging Research and the Behavioral and Social Research program at the National Institute on Aging, the 2006 Chicago Workshop on Biomarkers represents an effort to bring together laboratory and population scientists, who study social behavior from a biological perspective. This innovative series of invited workshops grew, initially, from the National Social Life, Health and Aging Project (5R01AG021487), an interdisciplinary study collecting a broad array of biophysiological measures in combination with survey questionnaire data using a home-based, national probability sample, and lay interviewers. Motivated by the enthusiasm and interest of participants attending that first, small workshop, deriving largely from a single project, the annual workshop series has evolved to address the needs and interests of a variety of projects (e.g. Health and Retirement Study; MIDUS; SEBAS; Social Isolation, Loneliness, Health and the Aging Process; AddHealth, NHANES, Canadian Longitudinal Study on Aging, WHO, etc.) and researchers, many of whom receive funding from the National Institute on Aging. Since 2003, the annual Workshop has fostered an expanding interdisciplinary network of senior and junior scientists actively engaged in biomarker collection in population-based health and aging research in North America and Europe.

While growth is one indicator of success, repeat annual attendance from lead researchers demonstrates the ongoing value of the conference. Diverse attendance from disciplines across the social and biomedical sciences, presents a major, unique draw to the workshop. Every year, we see an increasingly broad range of attendees from across the social sciences (e.g. sociology, anthropology, economics, public policy, political science, demography, psychology) and both clinical and basic science biomedicine (e.g. pediatrics, ob/gyn, internal medicine, geriatrics, otolaryngology, dermatology, infectious disease, cardiology, neuroscience, epidemiology). We have also engaged bioethics participants, as a consequence of an ongoing CCBAR collaboration with the University of Chicago MacLean Center on Clinical Medical Ethics. Finally, in an effort to draw innovation from outside academe, the workshop has also included an outside speaker, or "Translations," series. In 2004, an FBI special agent spoke on the collection of biological data from crime scenes, provoking thought and lively discussion about human subjects' rights and the importance of preventing "biologic contamination" of the scene, or research setting. Additionally, he shared biomarker collection technology and equipment used by the FBI, but unfamiliar to most researchers. In 2005, a NASA speaker, the first U.S. physician-astronaut, spoke on not only the technical and technological challenges of collecting biological data from astronauts in space, but the sensitivity of such data, where aberrations from normal could prove career-threatening. This year, a representative from the White House Office of Science and Technology Policy presented an overview of biometric technologies and elaborated on how the government applies them in its daily activities. These memorable sessions have been positively reviewed and have resulted in translation of technology, materials, and ideas from fields typically beyond the scope of most academic researchers.

Each year, the Chicago Core on Biomarkers in Population-Based Aging Research (CCBAR) at the NIA NORC-University of Chicago Center on Demography and Economics of Aging, publishes the Proceedings of these workshops. They are distributed to Workshop attendees and NIH colleagues and are posted on the CCBAR website at http:/biomarkers.uchicago.edu. We have found that these Proceedings provide an excellent reference, particularly for investigators new to this field. We thank each of the individuals who shared and edited their presentation for publication in the Proceedings and for the very engaged participants, whose repeat attendance and intellectual involvement, fuels our enthusiasm for continuing this Workshop series.

Stacy Tessler Lindau, MD, MAPP
Lisbeth Nielsen, PhD

INTRODUCTION

Lis Nielsen and Stacy Tessler Lindau

NIELSEN: I'd like to welcome everyone to this Fourth Annual Biomarker Workshop sponsored by the Chicago Core on Biomarkers in Population-Based Aging Research and the Behavioral and Social Research program at the National Institute on Aging. Stacy Lindau and Alicia Frasier and I have been working on this program for about a year now, and we're very excited about what we have ahead of us in the next day and a half.

The Behavioral and Social Research Program at NIA supports social and behavioral research and research training in the processes of aging at both the individual and the societal level, and our emphasis is on the dynamic interplay between processes of aging at the individual level and individuals' changing biological, biomedical, social and physical environments. Much of the research that we support looks at multilevel interactions between all of these different levels of analysis. In fact, interdisciplinary research is heavily promoted by our program and viewed as critical for advancing behavioral and social research on aging. The goal and the challenge of interdisciplinary research is to integrate the conceptual and analytical strengths of disparate scientific disciplines in order to create hybrid disciplines that can eliminate gaps in terminology, methodology, and approach and ultimately broaden the scope of scientific investigation.

One way that BSR tries to contribute to these efforts is by bringing to the table, in meetings such as this, individuals from a variety of disciplines: psychology, sociology, epidemiology, genetics anthropology, economics, neuroscience, to share knowledge, discover common ground and develop a common language, with a goal of fostering collaboration and with the potential to push the behavioral and social science of aging forward. This workshop represents an effort by the program committee to bring together laboratory and population scientists studying social behavior from a biological perspective.

Stacy Lindau and others present here today have, at previous meetings and in other contexts, emphasized the challenges associated with bringing the biomedical laboratory into the population area of research. This is a continued challenge as new biomarkers are discovered and need to be validated and will continue as this field progresses. But for this meeting we're shifting the focus a bit, and the emphasis is more on how we bring the laboratories of the psychophysiologist, the social neuroscientist, the geneticist, and the ethologist to bear on their insights on population-based aging research.

So our goals are twofold and relate to our overarching programmatic themes of linking biology and genetics with the behavioral and social sciences, including efforts to advance methods of data collection and models of analysis.

First, the goal from the perspective of the population scientist - where the emphasis has been to identify pre-disease pathways and improve our understanding of the psychosocial factors that translate into health outcomes over the life course - is to take a look at methods from

laboratories that offer opportunities for considering alternative measures of phenomena like allostatic load or cumulative physiological risk.

The more expensive and time-consuming measures of the psychophysiology laboratory, for example, allow for exploration of dynamic processes and systems related to, say, stress reactivity in the sympathetic nervous system or HPA axis. It may be that the earliest signs of dysregulation in these systems are actually in their reactivity and recovery parameters that are only captured using the fine grain measures of the laboratory and less obvious perhaps in basal levels of biomeasures that sometimes can be collected more readily in population-based surveys. There are a variety of methods for bringing the laboratory and the field together, and many examples exist in currently funded BSR projects. One way is to embed experiments in surveys either in the home, as the NSHAP study here in Chicago has done with measures of sensory function, or as the Health and Retirement Survey has done with measures of cognitive performance.

Another way is to bring survey respondents into the laboratory for more detailed medical and psychophysiological assessment, as in the MIDUS study out of Wisconsin led by Carol Ryff, where participants are recruited from a large population-based sample to undergo a variety of stress reactivity and affective neuroscience investigations in the lab.

A final bridge between the laboratory and the field relates very much to the second goal of the workshop. This is to look at research opportunities from the perspective of the laboratory scientist. From this perspective, these kinds of meetings represent an opportunity to consider how to refine laboratory methods to make them more population-research friendly. They also prompt consideration of how to integrate measurements of well-being and social behaviors into laboratory paradigms, so we can test how laboratory findings translate into real world behaviors and outcomes of relevance to the aging population. I really look forward to this day and a half of exciting presentations we have lined up for you, and I'll turn the podium over to Stacy who will introduce the first speaker.

LINDAU: Good morning. Lis, thank you so much. It's been a pleasure to plan this meeting with you this year. As you know, last year's workshop was planned together with Northwestern, Thom McDade in particular, who we also welcome here today. It's amazing to see how the workshop has grown. There are many familiar faces, but also many new ones and I'm delighted to welcome newcomers as well as those of you who are returning this year.

We are going to start by introducing ourselves to each other. Please say who you are and where you are from. This will give everyone a chance to appreciate this phenomenal interdisciplinary group which is hard to find in other workshop settings. This will also give everyone an opportunity to know with whom they may want to sit at lunch or talk during a break.

(Introductions omitted from this record.)

In welcoming everyone to the workshop, I'd also like to acknowledge some of the individuals who played an important role in planning. Alicia Frasier, an epidemiologist, works with the Biomarker Core. She really deserves the lion's share of credit for making this workshop happen today. Of course Natalia Gavrilova, Lis Nielsen and I were also very much involved, but the logistics fell largely on Alicia, and she's done a great job. We also had a terrific amount of support from Precious Johnson, Amelia Karrakar, Adelle Hinojosa, and Kathleen Parks, all of whom are at NORC and Karl Mendoza who works with me.

As we begin, I want to share an article with you from the New York Times published last December. This article describes population-based collection of biological measures as it has been done with a Kenyan tribe that has been "poked and prodded," so to speak, by anthropologists over many, many years. The story provides a poignant view, from the vantage point of a remote people, on the kinds of research practices we have been discussing at this workshop. A weathered cattleman describes a researcher plucking hairs from his head: "I thought I was being bewitched…I was afraid. I'd never seen such a thing before." He and others describe the feeling of being prodded by foreign people, looking foreign in every way, wearing university sweatshirts that they'd leave behind as they returned home, and what that meant to the tribe over time. The article also describes the value of this research, what scientists have learned from being able to reach out and, literally, touch, individuals living in remote groups.

While some of us are in the business of going to remote areas and using very much the same methods (in fact, introducing some of the methods) that others of us use in population research in the developed world, others of us are much more comfortable or experienced in the clinical setting using more invasive approaches. We ask questions about social and cultural factors, in an attempt to understand how these affect illness and/or medical care of our patients. We work in mobile clinics in an attempt to broaden clinical health research to populations who may not or cannot access clinical settings. Or, we are going to the population setting to address questions of health at a demographic level and we're bringing biomeasurement with us.

But what do we do with the information once we collect it? A goal of this workshop has been to promote methods and models for analytic integration of the data we are collecting, increasingly in collaborative, multidisciplinary teams. There is good evidence that the data we are collecting resonates across disciplinary boundaries, but I have continued concern about the analytic integration of these data towards advancement of health. The hope has been that by bringing social, behavioral and biomedical scientists together we might learn something more or understand phenomena more richly than if we tackle problems alone.

Reflecting on our own project, the National Social Life, Health and Aging Project, we are actively facing these issues. We have a wonderfully diverse team and are anticipating, any day, delivery of the

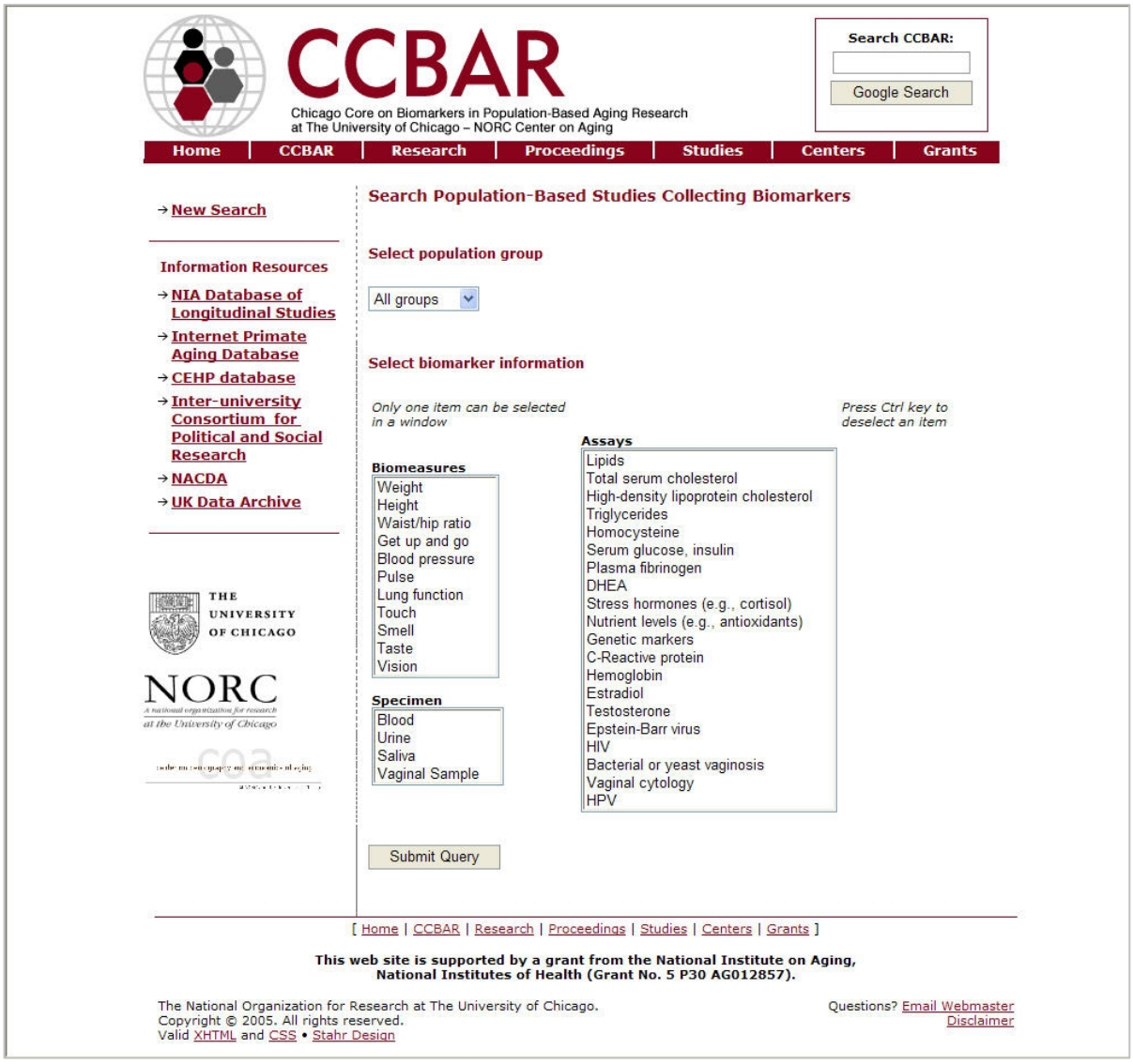
Wave I data set. Will we dive in to the data together? There is a natural tendency, and, to some degree, a need, to go off into our own disciplinary areas of comfort to start working with the data rather than face the challenge of trying to take integrated approaches to understanding and analyzing the data. We have to be purposeful about working across disciplines to answer some of the important questions that really beg for a multilevel, multifaceted perspective.

Thom McDade and I had the opportunity to present our thoughts on some of these issues last week at the National Academy of Sciences. It seems to me that there is general agreement that technological advancement, a push toward minimally invasive techniques for collecting biomeasures, is an area of critical interest. Our best practices require not only applying the technology that we use to collect biological data but also being deeply familiar with the underlying biophysiological processes that justify the use of these measures. This does not mean that each of us needs to be a master of physiology or biomedicine, but advancing the frequency and quality of our collaborations with life scientists is, at least, important for communicating the value of such data to the people who participate in our studies and for making sure our data collectors also understand, believe, and can communicate the rationale behind biological data collection.

Ethical issues related to the collection of biological data in the population setting echo those pertinent to the collection of non-biological data. Some of the issues push the boundary of can easily be extrapolated from traditional interview-based research. These include the destiny and future use of biological specimens, sharing of results with study participants, sharing biological data across studies or with other investigators, prioritizing use of limited biological specimens, and adapting to changes in technology that can complicate many of these issues. In terms of realizing the maximum public value of our work, we are faced with not only how to share, but how to know and build upon what each other is doing.

Relevant to this last point, I want to briefly introduce a project under way at the Biomarker Core which aims to provide a web-based mechanism for sharing information related to biomeasure collection across studies. Natalia Gavrilova, among other talents, is a skilled computer scientist and deserves 99.9% of the credit for creating this prototype. Alicia Frasier, with help from Precious Johnson, has also been instrumental in developing the substantive content for the database. This is a work in progress for which we are seeking funding – we'll talk to our friends at NIA about that another time - but aims to provide an easily searchable database that will allow researchers to see, by population group, which studies have collected which biomeasures. We hope that this will become a way of rapidly sharing information for the purposes of promoting collaboration across studies and for continuing to advance the knowledge and science around population-based use of biomeasures.

To further highlight Lis Nielsen's comments, I'd like to share this diagram as an example of what we mean by integrated health research. This framework summarizes the way I think about the work we are doing with the NSHAP project and may resonate for many of the sessions that will happen here over the next day and a half. IN particular, an area of interest for the NSHAP project is thinking about how social factors influence health and health outcomes over time. We're interested in the biological and physiological mechanisms through which this occurs. One of the things that makes our particular study unique is that we are collecting a fairly comprehensive, albeit superficial, panel of measures related to sensory function. How is it, for example, that individuals perceive their world, process information from contacts with other individuals, either consciously, subconsciously, or unconsciously? What happens in the brain? We don't yet have a portable functional MRI that fits in our wheelie bag, so we can't really know what's happening inside the brain with NSHAP data (but maybe one day!). And then, ultimately, we want to know how these social stimuli influence biology and, therefore, health.



As we work with this model, I'm very much aware that there's concern about not only how social factors get under the skin to influence health and health outcomes, but what does it mean to think about how biology affects behavior. I think it would be an understatement to say that this is a sensitive and controversial area. We will take this head-on as part of our discussion tomorrow in a session on a proposal to measure biological aspects of nihilism.

I want to briefly update you on plans for our 2007 workshop. This is intended to be a global workshop on trends in integrated health and aging research that will bring together representatives of many of the world's studies on aging, particularly those working at the interface of social and biomedical approaches to health. This will be held at about the same week next ear and some of the agenda topics are listed here. Next year's workshop will help us look together at possibilities of international collaboration and comparative analyses across national and cultural boundaries.

I'm now going to introduce Dario Maestripieri who has graciously agreed to give the keynote address. I think he was a bit skeptical of me when I first approached him upon Martha McClintock's urging, but I have found him to be open-minded and, even enthusiastic, about being here today.

I want to put this in some context. Dario studies non-human primates. What does this have to do with our work. This is a clip from the Kids National Geographic magazine that my son received in the mail back in January. It is a piece on whether animals love each other. One part of the story was about gorillas at the Brookfield Zoo, here in Illinois. It described the death of Babs, the gorilla and how the other gorillas interacted with her as she was dying. The researchers at the zoo laid Babs' body to rest in an open area where the other gorillas could visit her and made observations about the kinds of interactions, possibly emotional interactions, they observed. Emotion, personality and behavior are typically assessed via self-report in human health research, but I'm guessing it's hard to get a self-report from non-human primates. What can we learn about how to objectively or biologically observe such human traits and behaviors by thinking about the work of non-human primate researchers? I'm optimistic about the possibilities of learning from research on non-human primates as a way to thinking about how to measure, in an objective or non-self-report way, things that are hard to get at like emotion and affect and behavior. Dario's talk will motivate us in that direction.

Are there opportunities for translation from the kind of work and observations that are done with non-human primates? I think so. Dario's talk, as it's titled, Population-based Non-human Primate Biobehavioral Research sounds really, really familiar except for the non-human primate part.

Dario is an Associate Professor of Comparative Human Development and a faculty member in the Biological Sciences Collegiate Division at the University of Chicago. He's a fellow of the Brain Research Institute also at the University of Chicago and received his Ph.D. in psychobiology at the University of Rome LaSapienza, somewhere I'd like to go. His research interests include behavioral biology and behavioral development from a comparative and evolutionary perspective, and I actually have a picture here of a recent publication, 2006, Proceedings of the Royal Society, which got quite a lot of press attention. Fascinating work, and I encourage you to talk with Dario more about that. Without further ado it's my great pleasure to introduce Dario Maestripieri.

Social and Biological Influence on Rhesus Monkey Development and Health Across the Lifespan: Challenges and Rewards of Population-Based Bio-Behavioral Research with Nonhuman Primates

Dario Maestripieri

Outline of the Talk

- Part I: Challenges and rewards of population-based bio-behavioral research with rhesus monkeys
- Part II: Social and biological influences on development of group-living rhesus monkeys
- Part III: Environmental influences on variation in biomarkers of aging between and within rhesus monkey populations

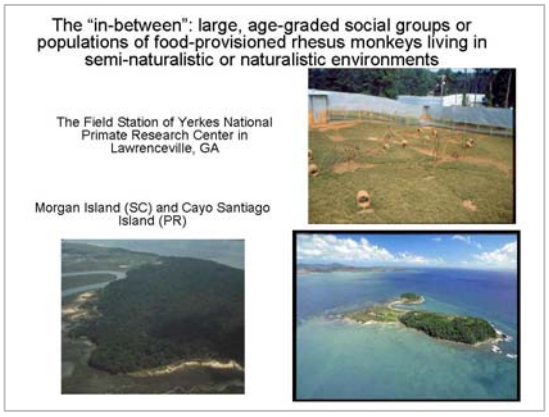
Rhesus monkeys are not as genetically similar to humans as chimpanzees are, but they are estimated to share 95 to 96 of their genetic material with humans, so they are a lot more similar to people than, for example, mice and rats are. Rhesus monkeys also share some basic life history traits with humans including a long period of slow growth, late age of reproductive maturation, and slow reproductive output in adulthood (rhesus females produce one infant at a time with interbirth intervals of one or two years). This long gestation length, long life span, and low reproductive output are characteristics that are also shared with other nonhuman primates, whereas rodents have very different life history traits. There are also similarities in behavior and cognition between rhesus monkeys and humans, and also obviously in health and pathology.

Rhesus monkeys are also excellent animal models for biomedical research because they're available in large numbers. They're not an endangered species. They're widespread in the Asian continent, and they're also found, thousands of them, in research facilities pretty much everywhere in the U.S. and other countries. One of the reasons for that is that they survive and breed very well in a wide range of environments. Researchers have tried to breed all kinds of primate species, and by far rhesus monkeys have been the most successful. They're also relatively easy to handle, at least compared to other primates such as chimpanzees, and for those who are interested in developmental and aging research, their lifespan is shorter than ours. They typically live about 20 to 30 years, so they allow researchers to do longitudinal studies of development and aging that will be difficult or impossible to do with humans.

Most biomedical research with rhesus monkeys is done in research laboratories. What I mean by a laboratory is a research facility in which rhesus monkeys are often housed in unnatural housing conditions. In the wild, rhesus monkeys typically live in large social groups of up to 100 individuals. These groups have a strong matrilineal structure. Females spend their entire lives in groups surrounded by their kin. Males leave their group at puberty and immigrate into another group. In the lab, rhesus monkeys are often housed in small cages, either by themselves or in pairs; and when they're housed in groups, these groups tend to be small, and they have somewhat of unnatural composition in terms of age- and sex-classes of individuals or in terms of presence or absence of kin.

In some research laboratories rhesus monkeys also have an unnatural rearing history, which means they're separated from their mothers at birth and they're either hand-reared or they are reared with peers (with same aged monkeys) in a small cage. When these monkeys grow up, they develop these very strong attachments to peers and often show behaviors that you would not see in the natural environment. Rhesus monkeys are very resilient to environmental perturbations, so they survive and breed well in all these conditions. However, they do develop a lot of behavioral and physiological alterations which might affect, for example, some biological measures that we might want to collect from these animals for behavioral or biomedical research.

One alternative to do research with monkeys in the lab is to do it in the wild. The wild looks a lot prettier than the lab, but it has its own limitations, in terms, for example, of how much you can manipulate the animals and what kind of biological measures you can collect from these animals. So it's probably not the best setting for doing biomedical research.



There's another alternative to the lab and the wild, which is what I call the in between. This consists of research done in large age-graded social groups or populations of food-provisioned rhesus monkeys living in semi-naturalistic or naturalistic environments. I will give you three possible examples of this. One of them is the population of almost 2,000 rhesus monkeys living at the Field Station at the Yerkes National Primate Research Center in Lawrenceville, Georgia where I've done most of my research. Another is the population of over 3,000 rhesus monkeys living on Morgan Island, an island off the coast of South Carolina. The third one is the population of over 800 rhesus monkeys living on the island of Cayo Santiago, off the coast of Puerto Rico.

The monkeys living on these islands are food-provisioned but free-ranging, whereas the monkeys at Yerkes are captive. What all these populations have in common is that the monkeys live in very naturalistic social environments. They live in large social groups, which have a structure that is very similar to the social groups you would see in the wild. The females in these groups have a matrilineal structure. The females remain in the group surrounded by their kin for their whole life. The males are taken out of the group, removed at puberty, to simulate the process of male migration in the wild, where males migrate freely between groups. The monkeys that live in these situations look very behaviorally and physiologically healthy. I think they provide an excellent opportunity for doing behavioral and biomedical research.

Why should we do research with these large populations of monkeys? Well, there are some possible issues with doing research in the lab versus taking a population-based approach. I'm going to just give you three examples. These are just a few of the issues.

One is an example of a phenomenon that was studied both in the lab and in these large populations, and lab studies provided very different results from the findings that were obtained from the more population-based studies. This is an example that has to do with the study of the hormonal regulation of female sexual behavior and the nature of female sexuality. The second example I want to discuss has to do with the function of maternal attachment. This is a case in which laboratory studies have resulted in theories of maternal attachment, which in my opinion don't have a lot of applicability outside the lab in which they were developed. Finally, I want to talk about aging. Aging is another possible example of a phenomenon that looks different in the lab and in free-ranging monkey populations.

So let me start with the first of my examples. There's been a lot of interest in studying hormonal regulation of female sexual behavior in monkeys. Early studies were done back in the '60s and '70s. They were done in the lab using female monkeys that were individually housed in a cage. These female monkeys were paired with a male, so male and female were together in a small cage, and they were observed every day of the female's menstrual cycle. Some experimental manipulations of hormones were also done with these monkeys. One of the results of these studies of individually-housed monkeys was, for example, that the females mated with the male every day of the menstrual cycle. It didn't seem to make a difference whether the female was having her period or was close to ovulation. Females did not seem to encourage or reject male sexual behavior under these laboratory conditions. Changes in concentrations of steroid hormones across the menstrual cycle did not seem to affect female sexual behavior or sexual motivation.

The conclusions of these studies were that: female sexuality in rhesus monkeys seems to be uncoupled from hormones, and females seem to be sexually passive and always available to satisfy male sexual desires. Now, I have to add that not only were these studies done in the lab, they were also all done by male researchers, so there might be a little bit of a bias and wishful thinking in these conclusions.

When studies of sexual behavior were conducted at the Field Station of the Yerkes Primate Center or in the free-ranging population on Cayo Santiago, these were the findings of this research: First females mated mostly or only during the periovulatory period of their cycle. Females actively solicit or reject male sexual behavior depending on the phase of the cycle they're in and the identity of the male. They're interested in sex only when close to ovulation, they like some males but dislike others.

Peaks in female sexual motivation and behavior are strongly linked to peaks in estrogen concentrations at mid cycle. So the conclusion of these studies was that female sexual desire is actually under tight hormonal control, and females are in charge of sexual activity in rhesus monkeys.

So the results and the conclusions are very different from these different sets of studies. If we hadn't taken a population-based approach to this question, I think we would have accepted findings and conclusions that are not entirely accurate.

The second example I want to discuss is about the nature and function of maternal attachment. There has been a lot of interest in studying what maternal attachment really does for monkey infants for a long time and a lot of studies were conducted between the '50s and '90s using individually-housed monkeys.

In many of these experiments, monkey infants were separated from their mothers and then followed developmentally. A lot of these studies led to the formulation of a psychobiological theory of attachment according to which the primary function of maternal attachment is to facilitate the development and regulation of the infant's biological rhythms. In other words, what was observed in these studies was that these infants who were separated from the mothers at birth and grew up without a mother exhibited a wide range of dysregulations of physiological rhythms, for example, in terms of body temperature regulation, respiratory rhythms, and cardiovascular function. So the conclusion was that the primary function of attachment to the mother is to facilitate the development of the infant's developing rhythms.

However, this theory is entirely based on the study of rhesus monkey infants separated from their mothers at birth, reared by hand and housed individually or in pairs in small cages. This theory does not take into consideration any information or any data whatsoever about what monkey mothers really do for their infants and with their infants because monkey mothers were never included in

these studies. They were never studied. So the theory is entirely based upon the presumed effects of mother's absence rather than on the observed effects on mother's presence. If you instead study mothers and what they do, for example, in large populations of monkeys, you'll see that mothers do a lot more than just helping the infant develop regulation of biological rhythms.

This approach is similar to the approach one would take, for example, to figure out what the function of money is and just study people who don't have money. For example, you just study homeless people. Homeless people clearly don't have any money, so that study can tell you something about the function of money. A lot of homeless people, however, are also mentally ill, and so were many of the monkeys in these laboratory studies. So some of these studies might have demonstrated the effects of mental illness on the infant's behavioral and physiological development and not the function of maternal attachment.

The third example has to do with aging; and as I said, this is just a potential case of a phenomenon that I think could benefit from a population-based approach. Captive rhesus monkeys in research laboratories in the U.S. seem to have a median lifespan of 25 years and a maximum life span of 40 years. In the free-ranging population of rhesus monkeys on Cayo Santiago, however, the median lifespan is 15 years, and the maximum lifespan is 25 years. This is actually a conservative estimate. On average monkeys don't live as long as this.

This difference potentially raises a number of questions, which could be addressed with a population-based approach. For example, do aging rates differ in lab-housed and free-ranging monkeys? Do free-ranging monkeys experience the same aging-related disorders as the lab-housed monkeys and at the same age? Are the causes of death the same or different? Do all the findings of aging research obtained with lab-housed monkeys also apply to free-ranging individuals? Would you like to know the answer to all these questions? Yes? So do I, so that's why I think it's worth addressing some of these questions using a population-based research approach.

Taking a population-based approach to research with human primates involves a lot of challenges, the first of which is funding. It's not easy to obtain funding for this type of research, for example, to do research in one of the research facilities I have just described. Now, without offending anybody in this room, it seems that NIH doesn't fully appreciate that value of animal models in population-based research. Even though all these facilities and their rhesus monkey populations are supported by NIH or other governmental agencies, these facilities are not supported as research facilities. They're supported as breeding facilities. In other words, it saves everybody a lot of money to let the monkeys run around on an island and reproduce on their own instead of keeping them in a lab and having to clean their cages every day. It's hard to convince funding agencies that these breeding populations also provide great research opportunities. As a result, all these facilities are largely underutilized in terms of research; and in some cases, researchers are not even allowed to step on the island where the monkeys are and do any research because they might interfere with the breeding program of the colony.

There are also obvious logistic difficulties involved in conducting research in places like Morgan Island or Cayo Santiago. First, you have to find the monkeys. Then, you have to follow them to collect behavioral data. You have to catch them if you want to take biological measures. This entails a lot of investment of time and resources. You have to train people to do this type of work. You also have to train the monkeys, for example, for capture, to allow researchers to obtain the samples they need. And finally there's also the challenge of getting the research published because this is not viewed as mainstream research. Sometimes you get a lot of resistance from reviewers and journal editors about publishing this type of work. So there are many challenges in conducting population-based research with rhesus monkeys, but I think there are also many rewards.

What are these rewards? One of them is the ecological validity of the findings, which in my opinion is priceless. You're studying monkeys in an environment that's very similar to the environment in which they evolved. You're pretty confident that your results, your measures, are not going to be an artifact of some really artificial environmental situation. Conducting primate research with a population-based approach also increases the extent to which the findings can be extrapolated to humans. Humans live in free-ranging populations. They don't live as inmates in jails. This research also provides the opportunity to understand biology / environment interactions in a way that would be difficult to do in the lab. For example, this research provides the opportunity to understand the effects of social environmental variables on behavior, physiology and health. Finally this research provides the opportunity to study phenomena at a population level, which is something you can't really do in the lab. For example, you can examine the distribution, the maintenance, and the transmission of particular phenotypes and genotypes within a population. There is some value, I think, in modeling these types of phenomena with animals and particularly with primates.

Now I want to give you an example of how my collaborators and I have tried to take this population-based approach to primate research, and to describe to you the results of a long-term project that was just completed at the Field Station of the Yerkes Primate Center of Emory University in Atlanta. This is a project in which we looked at a wide range of social and biological influences on development in group-living rhesus monkeys and addressed a number of different questions such as: What are the effects of exposure to naturally occurring variation in parental care and to infant abuse early in life on behavioral and neurobiological development? Are normal and abusive patterns of parental care transmitted across generations from mothers to daughters? What is the relative contribution of genetic and experiential factors to the intergenerational transmission of normal and pathological parenting? And finally, what are the neuroendocrine and neurochemical mechanisms underlying the effects of early experience on the intergenerational transmission of normal and pathological parenting? And specifically, what is the role of brain monoamine systems, and in particular, serotonin, in this process?

Because a lot of the data that I'll be presenting have to do with serotonin. At this point, I have to make a brief digression and give you some background information on serotonin and its relationship to impulsive and aggressive behavior in both humans and nonhuman primates. In biological psychiatry, there's been a lot of interest in serotonin and its relationship to impulsive aggression since the '70s. In the '60s and '70s, there were some studies that found reduced levels of serotonin in the brains of individuals who committed suicide compared to those who died from a violent death but did not commit suicide. Subsequent studies found reduced levels of the serotonin metabolite 5-HIAA in the cerebrospinal fluid (CSF) of depressed patients who attempted suicide by violent means.

Then there was a study showing that lithium treatment of prison inmates increased serotonin and suppressed impulsive aggressive behavior. There were also studies finding that violent offenders prone to impulsive aggression had lower concentrations of the serotonin metabolite in their CSF, and finally there is also some evidence showing that impulsive aggressive patients show a blunted response to the fenfluramine challenge. This is a pharmacological challenge in which the rise in plasma prolactin induced by acute administration of a serotonin agonist, the fenfluramine, is used to assess central serotonergic activity.

Taken together, these findings suggest that there is this relationship between low serotonin and impulsive aggression. Many of these findings have been replicated and expanded with rhesus monkeys. Research done by my collaborator, Dee Higley, with rhesus monkeys showed that rhesus monkeys with low levels of serotonin metabolite in the CSF exhibit high rates of impulsive aggression, have more wounds and scars and higher mortality rates from aggression, exhibit reduced amounts of social interaction, and engage in risk-taking behavior; for example, they take long leaps

from one branch to another (and sometimes fall and break their neck). They also have high mortality in infancy and emigrate from their group earlier.

Research done by Lynn Fairbanks with vervet monkeys shows that males who have low levels of serotonin metabolites in their CSF have a lower latency to approach an intruder, are more likely to achieve high dominant status, show blunted responses to the fenfluramine challenge, and fluoxetine treatment of these males reduces their impulsivity.

These findings replicate and expand those of human research, but also raise the question of where this variation in serotonergic function comes from. For example, there seem to be strong individual differences in CSF concentrations of serotonin metabolites, but where do they come from?

One answer to this question is that these differences are genetically inherited. There has not been a lot of work on this in humans, but there was a study with twins done back in the '70s that showed that the concentrations of monoamine metabolites in the CSF tend to show evidence of heritability. But I think by far the best evidence of heritability comes from a study of baboons that also took a population-based approach. Jeff Rogers and collaborators at the Southwest Primate Center in Texas measured CSF monoamine levels in almost 300 baboons. This is a population that has been completely genotyped, so their pedigree is fully known. These researchers used variance components methods to estimate heritabilities of CSF monoamine metabolite levels and multivariate analyses to estimate both genetic correlations and environmental correlations between metabolites.

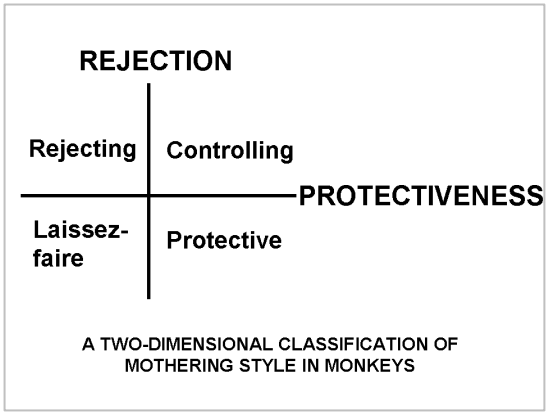
These researchers found evidence of high heritability for each metabolite, and both genetic and environmental correlations between metabolites but could not identify the source of these environmental correlations. So they suggest in the discussion of their paper: "It is conceivable that social experience during development, for example, the style of maternal behavior received as an infant might affect these neurotransmitter systems"

This is what the project I am going to present was all about. But before we get there, there is another piece in this genetics story that I have to tell you, and this has to do with the serotonin transporter gene polymorphism. Many of you might be familiar with this already. For those of you who are not, let me just briefly explain it to you. The figure here shows a map of the promoter region of the serotonin transporter gene in humans and rhesus monkeys. In both species this region is polymorphic, which means there are two alleles in the population, one short and one long, that differ in humans by a length variation of 44 base pairs, while in rhesus monkeys, the difference is 21 base pairs between the short and the long allele. This polymorphism is functional, because in vitro studies have shown that the short allele confers lower transcriptional efficiency to the serotonin transporter gene.

Why do we care about this? We care because studies have shown that this polymorphism in the serotonin transporter gene is associated with anxiety-related traits. This study published in Science in 1996 looking at over 500 individuals showed that this serotonin transporter gene polymorphism accounts for 3 to 4 percent of total variance and 7 to 9 percent of inherited variance in anxiety personality traits. This might sound like a small percentage to you, but in fact it's a large percentage of variation that is explained by this genetic polymorphism. Now, these findings have generated a lot of interest and last time I checked this article had been cited 1,065 times, which is pretty much twice as much as all my publications put together.

So now let's go back to our project. We were interested in looking at maternal behavior as a possible source of developmental variation in the serotonergic system, and in looking at how infants develop who are exposed to variable early experience.

Let me first give you some background information about maternal behavior in infant development in rhesus monkeys. Rhesus monkey females typically produce one infant at a time either every year or every other year. They don't get any help from males, from the fathers. Males don't even know they're the fathers. Mothers do everything on their own. Infants are typically weaned within a year or so. For the first few days or weeks after an infant is born the infant is in almost continuous contact with its mother. Then at some point the infant will break contact and start exploring the environment on its own. At times, it's the mother that breaks contact with the infant and tries to encourage the infant to be independent. Initially, infants spend very brief periods of time out of contact with the mothers. They don't walk very far from them. Mothers are very vigilant, and some mothers are so anxious that they don't allow their infants to move away. They restrain them by pulling them by their legs and their tail. In this slide, this baby just wants to go and play with this other baby; but this mother thinks there's something wrong with it and so she's pulling the baby by the leg. So there are anxious mothers. At some point, as infants get a little older, they spend more and more time away. When they try to go back to their mom, they find a little surprise. They get rejected by their mothers. Mothers hold their infants at a distance with an arm, as shown in this slide, or in other ways keep their infants from coming back and making contact; this is how infants are gradually weaned. Maternal rejection is a normal behavior that is functional for the acquisition of infant independence.



A TWO-DIMENSIONAL CLASSIFICATION OF MOTHERING STYLE IN MONKEYS

These patterns of maternal behavior that I have just described are something that you see in all rhesus monkey mothers, but there's a lot of individual variation. Monkey mothers who live in the same group or in the same population differ dramatically from one another, for example, in how frequently they make or break contact with the infants, how often they groom or hold the infants in their arms, and whether they restrain or reject their infants. If you analyze maternal behaviors with the principal components analysis, what you see is that behavioral variation tends to fall along two main dimensions, or factors, that are called maternal protectiveness and maternal rejection. What this means is that maternal behaviors like mothers making contact, restraining, approaching, and grooming the infant, all tend to be positively correlated with one another; whereas behaviors such as breaking contact, rejecting and leaving the infant will tend to be positively correlated with one another. These two dimensions are independent from one another so that mothers can be high on one factor and low on the other, or high or low on both factors. The interaction between these two dimensions results in four different types of parenting style. Mothers who score high on both dimensions are classified as controlling. Those who score low on both factors are classified as laissez-faire, and those who score high on one factor and low on the other are classified as either protective or rejecting. These individual differences in maternal behavior are very stable over time. Interindividual variability in maternal behavior also includes a small subset of mothers who exhibit violent and potentially harmful behavior such as this shown in the slide. For example, in the population of almost 2,000 rhesus monkeys at the Field Station of the Yerkes Center, about 5 to 10 percent of all mothers show this type of behavior, which I call abusive.

In addition to dangling and dropping their infants on the ground, abusive mothers also drag the infants by their tail, or their legs. They sometimes just grab their infants and toss them up in the air, or pin them down with their hands or step on them with their feet. The consequences of maternal abusive behavior may range from some superficial bruises or scratches to serious injury and death. This slide is too dark, but there's a rhesus monkey mother holding a dead baby here. These abusive mothers are otherwise perfectly normal individuals. You can't tell them apart from nonabusive

mothers, for example, when they don't have a baby; and even when they have a baby, they don't abuse these babies all the time. In fact, again this is a dark slide, but this is an abusive mother holding a baby with perfect posture. Abusive mothers know perfectly well how to be good mothers. But here's the same mother two seconds later dragging the baby by its tail. Abusive mothers alternate long periods of competent maternal care with short bouts of abuse.

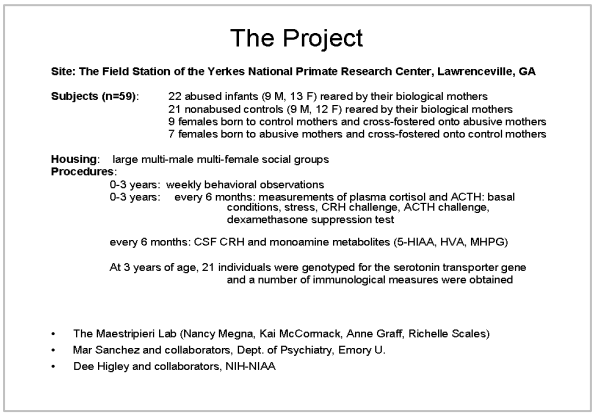
We've done a lot of work on infant abuse in rhesus monkeys, and I just want to summarize some of our main findings because they're relevant to the data I will be presenting later. By examining this phenomenon in this large population of monkeys at the Yerkes Field Station, we have discovered that 5 to 10 percent of all infants born in a given year are physically abused by their mothers. The sex of the infant, the birth order and the health status do not seem to affect the probability of infant abuse. We looked at animal records for over 30 years, a data set of over 3,000 individuals over a period of five to seven generations, and these records show that infant abuse runs in families along the maternal line. It is concentrated in some matrilines. It never happens in others. Within these matrilines, it is most likely to be exhibited by closely related females, for example, mothers and daughters or pairs of sisters. For example, we had a family in which there were five sisters that were all abusive and no one else in that group was.

Abuse begins as early as the first day of infant life and usually ends by the time infants are three months old. Abused infants are not neglected so there seems to be a separation between abuse and neglect. Abusive behavior is limited to a female's own offspring. In other words, abusive females don't go around and abuse other females' infants. They just do it to their own offspring. Mild and severe abuse differ in frequency but not in the pattern of behavior. Abusive behavior is very different from any other pattern of maternal or aggressive behavior, so it's very easily and effectively identified. Mothers are very consistent in their abusive behavior. Abusive mothers abuse most, if not all, of their infants that they have over the years. They're consistent in both the rates and the patterns of abuse that they use across different infants.

If these abusive mothers are induced to adopt an unrelated infant, they will abuse this unrelated infant with almost identical rates to those that they used with their own biological offspring, suggesting that abuse appears to be a stable maternal characteristic. Infants don't seem to play a significant role in the occurrence of abuse, for example, because infant abuse begins as early as the first day of life when infants show little or no independent activity. There are no obvious differences in the physical or behavioral characteristics of abused and nonabused infants; but we did find some differences in the acoustic structure of infant cries between abused and non-abused infants, which may somehow contribute to the occurrence of abuse. We have also discovered that abuse is sometimes preceded by stressful events, which suggests that some of the abusive mothers probably have trouble with emotional regulation.

Abusive mothers in general tend to have rejecting parenting styles. They reject infants at high rates. They are also less likely to show nurturing responses to the cries of their infants. They are generally very interested in infants, also those of other females. Some abusive mothers have high levels of anxiety. They are somewhat socially isolated within their group because they are approached by other individuals less frequently. They are also somewhat more aggressive towards other individuals. Parity, age, and dominance rank do not differ significantly between abusive and nonabusive mothers.

This shown in the slide is just one piece of data I want to show you, just to illustrate that the rejection rates of abusive mothers are much higher than those of controls. Control mothers begin rejecting their infants when infants are about three to four weeks old, and then they steadily increase their rate of rejection; whereas abusive mothers begin rejecting the infants essentially the day of birth or the first week; and although the rejection declines as the infants age, at the end of the third month of infant life, abusive mothers are still rejecting their infants at higher rates than the controls.



This is a description of our project. The project was done at the Field Station of the Yerkes National Primate Research Center. This was a developmental longitudinal study in which we followed 59 infants from birth through their first three years of life. Of these 59 infants, 43 of them were reared by their biological mothers: 22 of them were reared and abused by abusive mothers, and 21 were reared by controls. There were also 16 females that were cross-fostered at birth between abusive and non-abusive mothers. In particular, nine females were born to control mothers and cross-fostered onto abusive mothers. Seven females born to abusive mothers were cross-fostered onto control mothers. They all lived in their large social groups. We did weekly behavioral observations, and then for the first three years, every six months we caught the monkeys, obtained blood samples, and measured plasma cortisol and ACTH in basal conditions, in response to stress (a novel environment test), in response to a CRH challenge, in response to ACTH challenge, and in response to dexamethasone suppression test.

One main interest in this project was to look at the development of stress reactivity in the offspring of these mothers, but I'm not going to show you any of the hormonal data today. Instead I will concentrate on data for CSF monoamine metabolites.

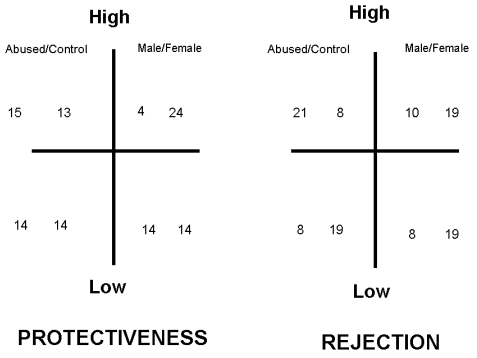
Every six months we also did CSF taps on these infants. We measured CSF CRH and the three monoamine metabolites, the metabolites of serotonin, dopamine and norepinephrine. Finally, at three years of age we genotyped a subset of individuals, 21 of them, for the serotonin transporter gene, and we also obtained a number of immunological measures. This was a collaborative effort, which involved my lab, my collaborator at Emory University, Mar Sanchez, and Dee Higley and his group at NIH.

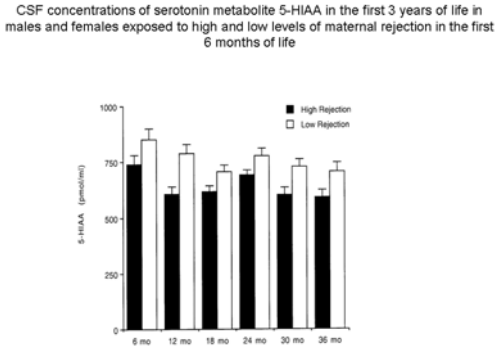
Let me show you some data. These are data showing that individual differences in maternal behavior are stable over time. What you see in this table is six different measures of maternal behavior: making contact, restraining, breaking contact, cradling, and grooming; and these are correlations between these measures, for example, during the first and the second month of infant life, the first and the third month, first and fourth month, so on and so forth.

As you can see, most of the correlations are highly significant suggesting that individual differences in maternal behavior tend to remain stable over the first six months of infant life. Individual differences in CSF levels of monoamine metabolites in the offspring are also very stable over time, over the first three years of life. In this table you see the CSF concentrations of the monoamine metabolites, and the correlations between the measures obtained at 6 and 12 months of age, 6 and 18 months, and so on (they were measured at six-month intervals). For all three measures there are very highly significant correlations suggesting that individual differences in these monoamine metabolite levels are consistent over time. This result replicates many of the findings obtained by Dee Higley and others suggesting that these variables are almost like trait-like characteristics of individuals. They tend to be stable even into adulthood.

Then we analyzed maternal behavior of all the mothers, including abusive and non-abusive, in our group with the principal components analysis, and we split our sample in half, around the median value for scores of maternal protectiveness and maternal rejection. Then we compared behavioral

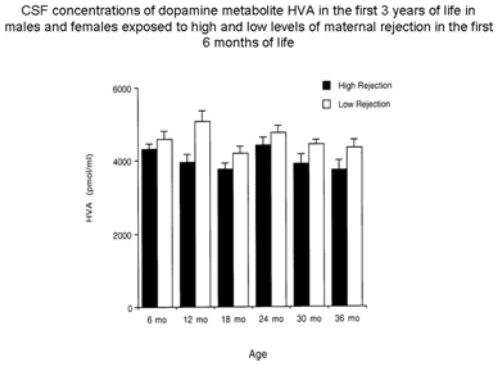
and biological measures in the offspring that were reared by mothers with high protectiveness with those who were reared by mothers with low protectiveness and offspring of high rejection mothers versus offspring of low rejection mothers.

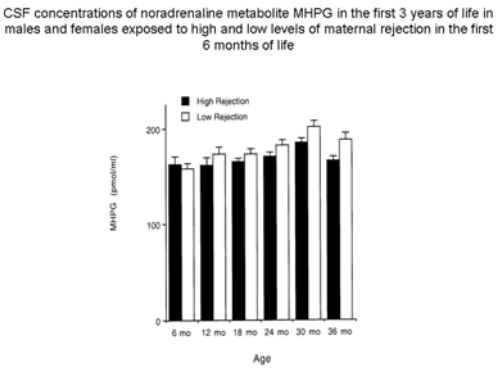




We found no significant differences between the offspring reared by high protectiveness and low protectiveness mothers. All the differences we found were between offspring reared by mothers with high rejection and mothers with low rejection rates. This figure shows the CSF concentrations of the serotonin metabolite 5-HIAA measured at six-month intervals in the first 3 years of life, in infants that were reared by mothers with high rejection rates and mothers with low rejection rates. These are only infants reared by their biological mothers, only 43 infants.

As you can see, there is a significant difference across all ages, whereby infants reared by highly rejecting mothers have lower CSF levels of the serotonin metabolite than the infants reared by low rejection mother. This figure shows the same data for the dopamine metabolite, HVA, and again there is the same difference. The infants of high rejecting mothers have lower HVA in CSF. The differences concerning the norepinephrine metabolite MHPG are also the same direction, but not as strong. So, in general, being reared by a rejecting mother seems to result in having lower CSF levels of these monoamine metabolites.

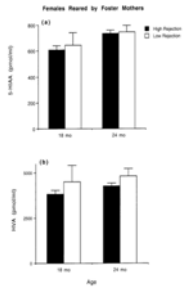




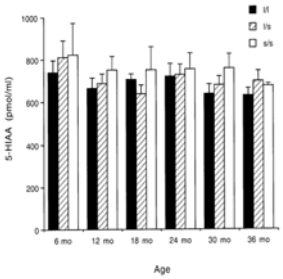
The question is: Do these differences reflect genetically inherited similarities between offspring and mothers, or are these really the effects of early experience? To answer this question, we looked at the data from our cross-fostered females, although we only had two data points for the cross-fostered females: CSF measures were only obtained at 18 and 24 months of age for these individuals. But again, if you compare cross-fostered females reared by high rejection and low rejection mother, you see that the difference is in the same direction. CSF serotonin metabolite levels are lower in the

offspring of highly rejecting mothers. This is the figure for the dopamine metabolite, and again difference is in the same direction suggesting that it's really an effect of early experience and not the result of genetic similarities between mothers and offspring.

CSF 5-HIAA and HVA in the second year of life in cross-fostered females exposed to high and low levels of maternal rejection in the first 6 months of life

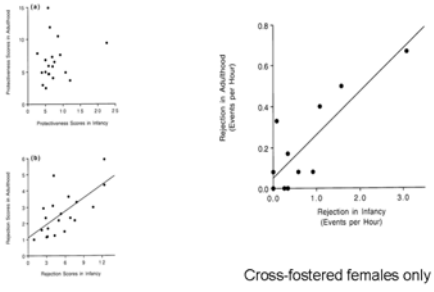


CSF concentrations of serotonin metabolite 5-HIAA in the first 3 years of life in individuals with the long and the short allele for the serotonin transporter gene



To further explore the possibility of genetics effects we compared the CSF levels of serotonin metabolite among individuals that differed in the serotonin transporter gene genotype. These are comparisons involving individuals that are homozygotic for the long allele, those that are homozygotic for the short allele, and then the heterozygos individuals. There seems to be a trend toward a difference among the groups, but the difference is not statistically significant. So the three groups of individuals that are genetically different for this particular gene are not significantly different in the serotonin metabolite concentrations. This result is consistent with our conclusion that differences in CSF serotonin metabolite leves in the offspring might be the effect of their early experience. We then waited until our female infants were old enough to give birth, which in rhesus monkeys occurs between three and four years of age, and we wanted to see whether there were similarities in the maternal behavior of the offspring and that of their mothers.

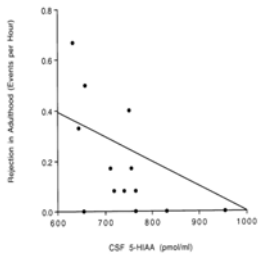
Intergenerational transmission of maternal behavior:
Maternal rejection rates in infancy and adulthood



Cross-fostered females only

Cross-fostered + noncrossfostered

Intergenerational transmission of maternal behavior:
CSF 5-HIAA correlates with maternal rejection rates of adult cross-fostered females

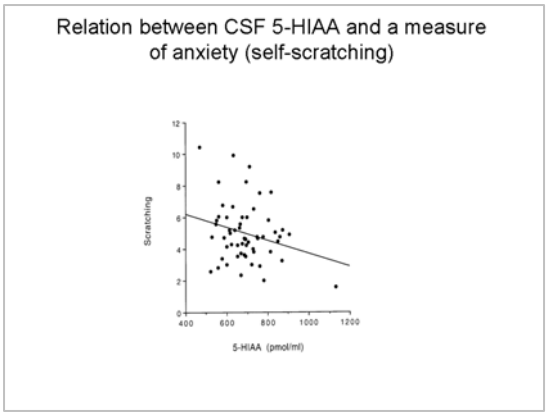


We found no similarities for maternal protectiveness, so there was no significant correlation between maternal protectiveness scores of daughters when they gave birth and those of their mothers. However, the maternal rejection rates of the daughters matched those of their mothers. This data set includes both cross-fostered and noncross-fostered females. If you just look at the cross-fostered females, you see the same significant relationship. This result suggests there is intergenerational transmission of maternal behavior, particularly maternal rejection, which is probably mediated by early experience and not by genetic factors.

This is another set of results showing a relationship between CSF levels of serotonin metabolites in adulthood in these females and their maternal rejection behavior in adulthood. The females who have lower CSF levels of 5-HIAA are also the females who exhibit higher levels of maternal rejection with their own offspring. So low serotonin might be one the physiological mechanisms that are responsible for the intergenerational transmission of high rates of maternal rejection.

The next question we wanted to address was that of intergenerational transmission of infant abuse. We looked at females who gave birth and measured the proportion of these individuals who displayed abusive parenting with their own offspring. The figure you're about to see compares the proportion of individuals who were abusive in these four groups of individuals: Females born to abusive mothers and reared by them, females born to controls and reared by abusive mothers, females born to abusive mothers and reared by controls, and females born to controls and reared by them.

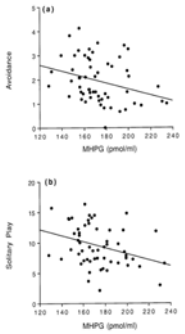
What the results suggest is that there is a strong effect of experience on the intergenerational transmission of abuse. In other words, about 50 percent of females who are reared by abusive mothers regardless of their birth condition exhibited abusive parenting with their firstborn offspring, whereas none of the females reared by control mothers, regardless of their birth condition, ever abused infants.



The next figure here compares CSF levels of the serotonin metabolite in abused females that either became or did not become abusive mothers. The top graph here compares the serotonin metabolite levels in these two groups of females that were born to abusive mothers and reared by them; and as you can see, the females who became abusive mothers in adulthood have lower levels of serotonin metabolite than the females who did not become abusive mothers. The bottom graph compares serotonin metabolite levels in these two groups of individuals among the cross-fostered females. This comparison involves a very small sample size but the difference is in the same direction. The crossfostered females who became abusive mothers had lower levels of the serotonin metabolite than those who did not.

CSF levels of monoamine metabolites were also correlated with other behavioral measures. For example, self-scratching in monkeys is a good marker of anxiety; and this figure shows that there is a relationship between CSF levels of serotonin metabolite and scratching behavior suggesting that individuals with a low level of CSF 5-HIAA tend to be more anxious than individuals with higher CSF 5-HIAA. CSF concentrations of MHPG, the norepinephrine metabolite, were negatively correlated with social avoidance of other individuals and with solitary play. So individuals with low CSF levels of this metabolite engaged in high levels of avoidance of others and in high rates of solitary play.

Relation between CSF MHPG and avoidance and solitary play



Conclusions of study

- Exposure to variable rates of maternal rejection in the first 6 months of life results in long-term alterations in the development of brain monoamine systems
- Both maternal rejection rates and abusive parenting are transmitted across generations with non-genetic mechanisms
- Long-term changes in serotonergic function are one of the mechanisms through which early experience affects the intergenerational transmission of maternal rejection and abusive parenting
- Experience-induced changes in brain monoamine systems also affect other aspects of emotional and behavior functioning later in life.

So what are the conclusions of the study? I think we've demonstrated that exposure to variable rates of maternal rejection in the first six months of life can result in long-term alterations in the development of brain monoamine systems; and that both maternal rejection and abusive parenting can be transmitted across generations with nongenetic mechanisms. We suggest that long-term changes in serotonergic function might be one of the mechanisms through which early experience affects the intergenerational transmission of both maternal rejection and abusive parenting. And finally our results suggest that experience-induced changes in brain monoamine systems may also affect other aspects of emotional functioning and social behavior later in life, for example, anxiety and social avoidance.

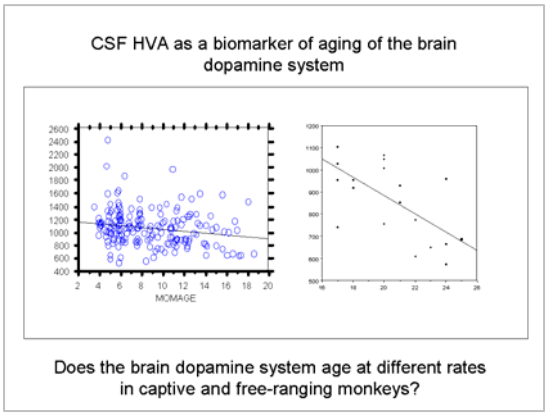
Now, I would like to conclude my presentation by saying a few words about some work in progress that has to do with environmental influences on variation in biomarkers of aging in rhesus monkey populations. As many of you know, rhesus monkeys have already been effectively used in many areas of aging research including, for example, neurobiology and cognition, skeletal and reproductive aging, dysfunction of the endocrine and immune systems, metabolic syndrome and diabetes, and also studies of cardiovascular disease. To my knowledge, however, rhesus monkeys have not yet been used for population-based research on aging, for example, research that investigates naturally occurring variability in aging and aging-related disorders between and within populations. They have not been used for research that investigates the possible contributions of genetic, social and behavioral factors to differences between individuals and population in aging and aging-related disorders.

Why do we care about all this? People interested in studying aging in nonhuman primates have put together this database, the Primate Aging Database or PAD, with the support of the National Institute of Aging, the National Center for Research Resources, and the Wisconsin National Primate Research Center. This primate aging database, which is now available online, contains over 400,000 data points for body weight, blood chemistry and hematology variables for healthy, non-experimental subjects across time and across research facilities. This website states that this database can generate normative data for a large number of primates across research settings because these data come from a variety of research facilities in the U.S. However, to my knowledge, this database contains no data on biomarkers of aging in free-ranging or semi free-ranging primate populations so we don't know whether these data also apply to free-ranging primates. So this database may be used to generate normative data for captive monkeys but maybe not for monkeys in general.

If we are not sure whether our aging data from captive monkeys can be extrapolated to free-ranging monkeys, how can we be sure that data from captive monkeys can be extrapolated to "free-ranging" humans? As I mentioned before, rhesus monkeys in captive lab research facilities tend to have a median lifespan of 25 years and a maximum lifespan of 40 years. However, on Cayo Santiago the median age of rhesus monkeys is 15 years and maximum lifespan is about 30. At this moment there

is a population of about 850 rhesus monkeys on Cayo Santiago. The oldest female I believe is 24 years old, so it doesn't even reach the median lifespan of captive monkeys. These monkeys are food provisioned. They have no predators, so what is going on? Does the aging rate or the aging process differ in a free-ranging rhesus population on Cayo Santiago when compared to captive rhesus monkeys?

I recently came across this paper. Actually this was given to me by my student Christy Hoffman, who is here, so thanks, Christy. This is a paper by an experimental gerontologist, Steve Austad, about the uses of intraspecific variation in aging research, and why it's important to look at variation in aging between populations. In this article, Steve Austad argues that researchers may take advantage of variation in aging within a species to investigate the nature and mechanisms of aging. For example, he suggests that a promising approach is to identify naturally occurring slowly aging populations to contrast mechanistically with a reference population. He reviews evidence that in rodents these population differences in aging rates and lifespan have already been identified. As far as I know, this has not been done with nonhuman primates. What Steve Austad is really talking about, for the most part, is populations that differ genetically in aging, but I think that studies of intraspecific variation in aging can also shed some light on the influence of environmental variables on aging rates and lifespan. So for example, we can ask the question: Is the aging process accelerated in free-ranging monkeys due to the greater cumulative effects of physical, energetic, ecological and psychosocial stressors experienced by free-ranging monkeys when compared to captive individuals? Maybe not, but this is an empirical question.



We recently started a project to address this question and have developed a number of biomarkers of aging among free-ranging rhesus monkeys on Cayo Santiago. We followed the guidelines and the advice of Don Ingram and collaborators in this paper in which they argued that for measures to be considered valid biomarkers of aging, they should show some evidence of significant cross-sectional correlation with age, significant longitudinal change in the same direction as the cross-sectional correlation, and significant stability of individual differences over time. I'm not going to go over the list of all the biomarkers, but I will give you one example. We would like to use CSF concentrations of the dopamine metabolite, HVA, as a biomarker of aging of the brain dopamine system.

I already showed you data indicating that individual differences in CSF HVA are stable over time. These are data from captive rhesus monkeys showing that between the ages of 2 and 20 years there is a slow but steady decline in CSF HVA in relation to age. If you look at data for the age range between 16 and 28 years, at some point particularly over 20 years of age the decline becomes much sharper. These are data from captive monkeys, so the question is: would we see the same pattern in free-ranging monkeys?

One of the questions that we want to address with our current project is: Does the brain dopamine system age at different rates in captive and free-ranging monkeys? Another question has to do with the effects of social variables on interindividual variation in aging and health. We know from human studies that people of low socioeconomic status and without social support are generally more vulnerable to aging-related diseases than individuals with more financial resources and support. However, in human studies social factors often co-vary with a number of other variables, for example, diet, physical activity, smoking, and alcohol and drug use, all of which affect aging and

health. We'd like to address this question with the rhesus monkeys on Cayo Santiago because they all have a similar diet and similar levels of physical activity, and they don't engage in smoking, alcohol or drug use, at least as far as we know. So some of the questions that we want to address are: Are monkeys of low social status and without social support more vulnerable to aging-related diseases? What are the neuroendocrine mechanisms underlying the effects of social variables on vulnerability to aging-related diseases? Therefore, we would like to take an approach similar to the one we took for the other project and investigate not only social and behavioral influences on behavior and health but also the neuroendocrine mechanisms that might mediate their effects.

I'd like to conclude this talk with some general statements. One of them is that currently and in my opinion unfortunately there is little or no use of animal models in population-based biomedical research. However, nonhuman primates can be excellent animal models for population-based research. In my opinion, the potential rewards of primate research are far greater than the challenges, so there is more need for primate research out of the lab and into the real primate world.

I'd like to end this by acknowledging all the research collaborators who have helped with the research done at Yerkes, Dee Higley, Mar Sanchez, Kim Wallen, all the staff and students who have helped over the years, particularly Christy Hoffman who is here and who has become our aging expert in the lab, and all the funding agencies who have made this research possible. And thank you for your attention.

McDADE: Hey, Dario. That was great. I think that's a wonderful demonstration of the value of the population perspective and integration of the social and biological; and it reminded me of a conversation I was having with a primatologist, Agustin Fuentes, who you may know, and he does work in Indonesia among other places, but he really looks at the interface between humans and nonhuman primates socially and a little bit biologically too.

So I wonder if you could even take your model of studying primates in naturalistic settings or seminaturalistic settings in the case of Yerkes to a truly naturalistic setting where they are really interfacing with humans in an intimate way in urban centers and other developed countries where actually the transitions of rhesus monkeys in particular are very similar to transitions that humans are going through in terms of lower levels of physical activity, increased levels of high calorie dense diets; and we're seeing metabolic syndrome, same sorts of things, so why not do the same thing but in Jakarta.

MAESTRIPIERI: Absolutely, let's do it. In addition to what you've already mentioned, the studies looking at interactions between monkeys and humans also offer the opportunity to look at the transmission of infectious diseases. Now that there is a great amount of contact between humans and monkeys in some countries, there are also more opportunities for the transmission of disease. A lot of macaque monkeys, for example, carry a virus which is harmless to them but is lethal to humans, the herpes B virus. More and more often we hear of people getting scratched or bitten by monkeys, and there are all kinds of other pathogens that can be transmitted. This type of research would offer the opportunity to address some of these issues.

INGRAM: Very nice presentation, Dario. And of course we support from our laboratory the value of rhesus monkeys as an incredible model of human aging. But I did want to address your overemphasis on the value of natural population to study in aging.

I have studied lots of rats in addition to monkeys; and if we study them in their natural environment, we would get a lot of contamination of the concept we consider aging. They have lots of parasites, have lots of diseases, lots of injuries. In the lab we're going to really see the manifestation of aging in a protected environment. True, it might not duplicate a lot of their social situations and other

situations they examine, but the lab allows the aging process to manifest itself without being accelerated by these extraneous factors that we consider non-aging like disease injury.

Now, as regards to Cayo Santiago, I think the same issue applies. Your estimation of median lifespan of 15 years, I don't know about the life table analysis you did, but it may be even generous there. I mean it may be much shorter. Yes, a conservative estimate. I hesitate to tell you that I had estimates of less than 10 years of median age that may include tons of infant mortality; and I've been there and it's a natural environment, but in my estimation it's an overcrowded situation which again raises lots of issues of the type that you just mentioned in terms of social -- psychosocial acceleration of aging processes. I'm all in favor, but we better look at that more carefully because I think that most of the deaths which you're using in your analysis of mortality are due to injury and disease, and we have to be very, very careful about that. Just because they live a shorter life, we can't call it accelerating aging.

MAESTRIPIERI: Absolutely. We don't know whether it's accelerated aging. As I said, it's an empirical question. And I agree with you about the difficulties of doing research with natural populations. I would include those difficulties among the challenges of this research. However, I would disagree with you about labeling the variables you mentioned as contaminants or extraneous factors. I think the behavioral, social, and environmental variables you mentioned are integral components of health and aging processes. We live in an environment that is contaminated by all these things. So I think for research with animal models to be valuable, it might help to be able to look at these issues in an environment that is similar to the environment in which we live.

INGRAM: We're not in much disagreement and I look forward to further discussing it. I'm totally supporting your ideas.

MAESTRIPIERI: Great. Thanks.

WILLIS: I had actually a related question and comment. The work in this area that I have seen is from Jim Vaupel and colleagues who sort of did the experiment of saying, well, we don't know what the length of life is of creatures in the wild, and we could ask the question: How long could we get creatures to live in controlled conditions, and did this with the flies and so on and got much longer lifespans.

Now, kind of one of the intriguing features if we looked at your research from that point of view is reverse -- take the reverse of what you did is you showing what is happening with people -- or with monkeys in an at least seminaturalistic setting, and then you put them in jail and let them live in these very controlled conditions; and they seem to live longer, but the quality of their life, the nature of the sexuality of the females and so forth seems radically altered. And I guess it would be interesting to understand also what the interconnections are between different activities in these controlled environments. Do we really want to live that long?

MAESTRIPIERI: Well, we don't know, and basically, as you said, we don't know what's going on in one environment versus the other. I'm not really advocating one setting versus the other. I'm really advocating comparative research. We need to compare phenomena in different settings.

WILLIS: Is there, for example, a correlation between this very unregulated kind of sexual behavior and the mechanisms that might be responsible for it and the length of life, or are they just unrelated?

MAESTRIPIERI: We don't know, but that's an interesting question to link reproductive activity to lifespan and longevity and aging-related processes. Absolutely. Great question.

TEMPLE: Thanks a lot. I was curious about the feed forward process about intergenerational transmission of abuse. Was there any ideas about what might be the precipitating factors that initially set up that feed forward, and is there any way that that's strangely adaptive for those like maybe initial sets of individuals?

MAESTRIPIERI: Yeah, we think an abusive mother was dropped from the sky or came with an alien ship and was dropped into the population. Just kidding. No, we don't know. There could have been an environmental trigger for infant abuse, and then the phenomenon could have been maintained in the population through these experiential and neuroendocrine mechanisms. But I don't think this phenomenon is adaptive at all.

In a natural environment rates of abuse would be a lot lower than in captivity; for example, I saw infant abuse in the Cayo Santiago population but it's very rare. So, it's not an artifact of the captive environment at Yerkes. It's just less frequent in the wild.

TEMPLE: Any indication it's related to food shortages?

MAESTRIPIERI: No.

McCLINTOCK: I just wanted to underscore what I think is really interesting about your work, and I think you're headed in this direction, which is the idea of what are called epigenetic effects - so you were looking at the short and the long version of the serotonin, you know, pathway there.

How much do you think the abuse effects, I heard you say neuroendocrine, are in fact mediated by either, you know, methylation or any of the other various ways there are of changing regulating gene expression within the ring?

MAESTRIPIERI: Well, all the work done by Michael Meaney suggests that there are maternal behavior-induced changes in gene expression that affect both endocrine and behavioral outcomes in the offspring, so something similar might be going on here. But there might be different mechanisms, for example, responsible for triggering abuse versus controlling its transmission and maintenance in the population.

For looking at mechanisms that trigger abuse, doing research in the lab may be the best approach; but if you're interested in what keeps the abuse in the population or how abuse is transmitted, you need to take the research into the larger populational context.

HEIMAN: Thanks very much. I actually have two questions, so you could not answer one if you want. The moms that abuse, I'm wondering if they are having lower serotonin levels. Let me just cut to the chase. If you treated the moms with abuse with serotonin reuptake inhibitors, might you see differences in their behavior? We have depressed moms here as kind of a model and, therefore, the related question is when the infants of these moms are born, are they born with lower metabolites, dopamine, serotonin, and so on, not to say that they're there for encouraging the mom to abuse them; but if you have them, they may be picking up signals with abused moms and they're more irritating, anyway.

So that was one thing; and then if you could comment on protective moms, and really that's fascinating too, isn't it? And some sort of things around the genetics and passing on of protective behaviors since they might be antidotes, though I know they can occur in the same set.

MAESTRIPIERI: These are four or five different questions. The answer to your first question is that we don't know much about the moms because we focused our research on offspring development.

Ours was a longitudinal developmental study, so we have some data on first-time moms; but these are very young individuals. We have never really collected biological measures from all abusive mothers or done pharmacological manipulations. The only pharmacology manipulation we did with the mothers involved the opiate system. We were testing the idea that maybe abusive parenting was due to deficits in the endogenous opiate system and its relationship with maternal bonding. We did not find much there. But we did not pharmacologically manipulate the serotonin system.

The other question was whether infants are born with low serotonin metabolite levels. We don't know. The earliest measurement we took was at six months of age which is pretty early, and it's sort of difficult to do CSF taps with infants that are very young. And, well, the data from the cross-fostered animals suggest that infants who are reared by particular types of mothers develop these neurochemical profiles, and not that they're born this way. So an association between maternal behavior and the offspring neurochemical profile seems to occur regardless of the possible genetic similarities between mothers and offspring.

Regarding the protectiveness question, that was an interesting finding that protectiveness doesn't seem to have much of an impact on offspring development. It actually replicates similar findings from other primate studies. I don't know really what to say about it other than protectiveness is a very important aspect of parenting, and it's also related to, for example, availability of social support in the group, and that's one variable that we haven't looked at systematically. So it's possible that the offspring of protective mothers also have support from other relatives and perhaps if we included these variables, then a more general measure of protectiveness and social support might show some interesting relationships with biological data. That's just an idea.

NIELSEN: I'm going to be providing you all with a general introduction to what we have planned for this session today and turn over the session to our moderator, Emil Coccaro, who will bring all the speakers up to the podium and guide the discussion afterwards. So we are really excited about this particular session at this workshop because it's been envisioned as a way to bring the core of social neuroscience research to the attention of population scientists in order to stimulate a dialogue about how biomarkers of social behaviors might be included in population studies and how population research can inform research questions in the laboratory.

The three presentations in this session illustrate laboratory approaches to core social neuroscience questions regarding the biological mechanisms underlying fundamental social behaviors such as trust, bonding and sexual behavior. A common theme among the presenters in this session is an interest in chemical mediators of social behaviors such as neuropeptides like oxytocin or pheromones and how these influence social behavior, well-being and mental health. I'll briefly introduce our three speakers and our moderator before turning over the session to the first speaker.

Sue Carter is professor of psychiatry and co-director of the Brain and Body Center at the University of Illinois at Chicago. Doctor Carter's research focuses on the neurobiology of social monogamy in prairie voles, including pair-bonding and biparental care. Her work with animal models is aimed at understanding the role of neuropeptide hormones including oxytocin and vasopressin in mental health disorders of humans. Her current work explores how early social experiences or hormonal experiences influence subsequent social behavior and neuroendocrine responses to stress.

George Preti is an adjunct professor at the Monell Chemical Senses Center, and his research focuses on the nature and origin of human odors, particularly those from the underarm and oral cavity. His team has identified the chemical structures of many human odorants including so-called modulator pheromones in underarm secretions and has explored the role of human odors as social signals. In addition his lab has been generating metabolic profiles of human secretions to identify biomarkers of disease, individual identity and stress - work which I believe he will talk about today.

And finally Markus Heinrichs comes to us from the University of Zurich where he is assistant professor of clinical psychology and psychotherapy and led a research group on the psychobiology of social interaction. His research focuses on the psychobiological mechanisms of stress and the role of social support, partnerships and bonding and the neuropeptide system oxytocin in protecting against stress and anxiety.

Some of his most recent work involves the intranasal administration of oxytocin in humans to manipulate trust behaviors in economic games, and this particular study, which was published in Nature last year, represents the kind of integrative work at the crossroads of biology, social science and economics in the emerging field of neuroeconomics.

Our moderator today is also working in the field of social neuroscience: Emil Coccaro, who is chairman of the department of psychiatry at the University of Chicago. His work on impulsive aggression and intermittent explosive disorder in families and twins combines genetic analysis, biomarkers and pharmacological research to try to understand personality traits of hostility and aggression in their natural and extreme clinical variants. And his work has recently received considerable media attention. Yesterday he was on the Today Show discussing some of his work on impulsivity. I'm going to turn this session over then to Doctor Coccaro, who will moderate it for us, and I look forward to hearing all the talks.

SOCIAL INFLUENCES AND BIOLOGICAL PROCESSES: STRESS, BONDING, ANXIETY, AND DEPRESSION

Social Bonds, Oxytocin and Health

C. Sue Carter

CARTER: Good morning. Thank you all, including of course the organizers. I suspect that Martha has something to do with me being here, so my thanks to whomever blame or credit goes. First I want to start by thanking also a number of collaborators who are not here. I'm going to go very quickly through a lot of data. Particularly I wish to acknowledge Margaret Altemus because I'm going to talk about work we did on humans together, and also my colleagues at University of Illinois, Chicago who have helped with brand new data about the hormone oxytocin which is present in human saliva.

My talk deals with social bonds, so let me just quickly say what that is. I don't know. No one's ever seen one, but we can detect the consequences of social bonds and perceived social support; and we can use measures like behavior, endocrine responses, autonomic responses or even indices of health. I think it's important to understand that social bonds can be studied from the point of view of causation or consequences or, in terms of natural history - that is when and where behaviors or processes occur. Measures of biological processes, especially neurobiological measures, or biomarkers can of course be very useful; but we are often working with dynamic systems, which can be the bane of a scientist's existence because things are always changing. Abstract concepts like social bonds are most easily studied and understood in the context of their functions which include survival and reproduction. Social support and safety promote survival. Reproduction requires access to mates. Care of offspring is also helped by social support. And of course all of these eventually lead to increased fitness or genetic survival.

The next picture is taken from the work of Berkman, et al. (1992), and provides ultimate proof that social support matters. The dependent variable is living or dying after a myocardial infarct. People with more social support were more likely to survive. I'm sure there are better data out there now, but I like this because it's simple; and over time and age you find that, especially for men, living alone was dangerous, as shown in the red bars.

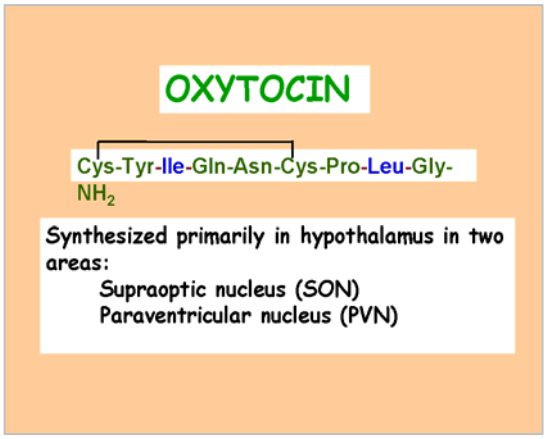
So back to my question: What is a social bond? Well, although we can't easily measure social bonds, we can at least probably agree that they are important. Social bonds do exist, and whatever these things are, they are not limited to humans. In some species, including a small rodent called a prairie vole, you will see that attachment or social support is more than a metaphor. Those are babies hanging on to the mom with little milk teeth. Titi monkeys entwine their tails together, leaving no question about who they like. They only sit with one partner. Physical presence and selective social behaviors are a very strong biomarkers of social behavior. The whole issue of who you want to be around, who you're willing to touch, who you're close to is central to social bonds.

We always have to keep in mind, however, that at least in the world of physiology, almost everything was done originally with the individual as the unit of analysis; and almost no attention was given to how animals were living and what was going on in their lives. I think we can say with great assurance that most living organisms cannot survive alone, and they of course cannot reproduce alone, with the exception of very primitive asexual organisms. The mammalian nervous system was also designed to work in a social environment. Social behavior is necessary for physiological and behavioral homeostasis. We know this by what happens when you take away the "other." The result is increases in drug abuse, food abuse, mental dysfunctions, illness, and any number of other things. So isolation is usually bad.

Now back to our original question: What are social bonds? We again are sort of pushed backward to saying, well, social bonds are good for you. They're associated with better mental and physical health. But even though we don't know what they are, we can ask when they happen. Social bonds form in late pregnancy, especially around the time of birth and during lactation, including parental-young interactions in the postbirth period. Social bond formation may also be associated with sexual behavior, as well as when there's a need for others.

There were two reasons I began to work with this oxytocin; one was the fact that oxytocin was released during and related to many of the behaviors that I was studying including social bonding, maternal behavior and sexual behavior. The second motivation came from personal experience: the birth of my first son on July 27, 1980. During the birth, I was forced to take oxytocin by the obstetrician. The alternative was a C-section. Or worse, it was explained, that if I did not accept the oxytocin my baby might die. Those were the choices I was given. At the time I became very interested in this compound, and that was over 25 years ago.

What is oxytocin? Oxytocin is a nine amino acid compound configured as a tail and a ring. Just to remind you, there are lots of different classes of biological molecules. Neuropeptides are relatively small peptides. However, at least when compared to steroids, they act quickly. However, in comparison to neurotransmitters (such as dopamine), the effects of neuropeptides tend to last longer and are more pervasive in their actions.



Oxytocin is synthesized primarily in the hypothalamus in two small areas, and when stained properly you can actually see these molecules. Useful models for studying the neurochemistry of human social behavior, and especially events dependent on oxytocin, are natural experiments including birth and lactation. Birth is a complex, rather quick event; lactation is much more chronic, and gives us something we can study in humans, and thus lactation is better understood biochemically or neuroendocrinologically than birth. Lactation is the defining feature of mammals, and oxytocin we know, is essential for normal lactation, but interestingly birth can occur even in animals (at least mice) that do not synthesize oxytocin.

Historically breast-feeding was the only viable method for nourishing newborns; for example as shown here, a woman who did not nurse her baby had to find a wet nurse. It is not common historically for anyone, but the mother to nurse the baby except in the upper classes. But now we're involved with this huge experiment that started roughly World War II when it became possible to easily feed your baby a formula. So women began to do this experiment, lactate or not? No one told

them it was an experiment. However, it was. Most women did not realize that the decision regarding how a baby would be fed was also a decision affecting the mother's nervous system and her endocrine system. The fact that some women breast feed and other do not also gave us a way of studying the effects or consequences of relatively high or lower levels of oxytocin.

So one of the questions I like to ask is: Are breast-feeding women really different? And the answer is yes. And we can compare these women to get indirect information regarding the nervous system as well as the consequences of social bonds, including social support.

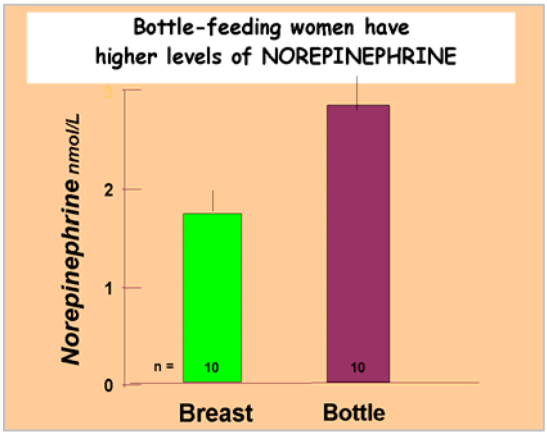
All of this keeps bringing me back to oxytocin. I want to show oxytocin to you. The dark spots in this brain slice are oxytocin stained with an antibody. The brain region you are seeing is the paraventricular nucleus of the hypothalamus (PVN), which looks like a heart. The PVN is right above the posterior pituitary and fibers from the PVN transport oxytocin to the posterior pituitary. Cells and fibers for oxytocin can be measured and stained. Oxytocin also has a "partner in crime" known as vasopressin, or more specifically arginine vasopressin. A lot of things about oxytocin can be better understood if you realize there's a natural antagonist working in the brain to turn the system on and off. Sometimes they even do the same thing, but mostly they're working in opposite directions.

Oxytocin is made in the brain (PVN), released into the blood supply at the posterior pituitary, but also released into the brain where it works on oxytocin receptors to influence both behavior and physiology. And here's a slide of vole brains showing the distribution of oxytocin and vasopressin receptors. This is just to show you that there are differences in where they are located between oxytocin and vasopressin receptors and that these patterns are species specific. This picture is particularly useful since it shows that in the prairie vole the prefrontal areas of cortex are heavily labeled, the bright colors being those areas with large numbers of oxytocin receptors.

To summarize, oxytocin is both released during and facilitates birth, lactation, sex, parental behavior and safe social interactions. In addition what I think makes this chemical very interesting is that it's very ancient. The ancestral form appeared before the evolutionary split that led to vertebrates and invertebrates. Oxytocin is the most abundant peptide in the hypothalamus, at least based on messenger RNA. There's only one known receptor. This is extremely unusual. With serotonin you could have 15 to 22 different receptor subtypes. Vasopressin has at least three receptor subtypes. Now, as I mentioned, lactation has major effects on the maternal brain and also the HPA axis. This is a Tintoretto painting called the Milky Way. The woman is expressing one function of oxytocin: milk ejection. Out of her breasts is coming the milk, which, according to legend, caused the Milky Way.

We have been trying to study lactating women for several years. This is a preliminary study done here in Chicago very recently, in which my colleagues and I were able to measure salivary oxytocin levels. (I have been working with a number of colleagues in the college of nursing, mainly Rosemary White-Traut and her students. In that study samples of saliva were taken. This study was done in each subject's home. Salivary samples were put into chilled ice and then the person's refrigerator and then into a minus 80 freezer. Salivary samples were taken at 30-minute intervals pre-nursing, during nursing and post-nursing. As you can see the oxytocin differed across these different times. The amount of oxytocin is rather low. It's in the under 10 picogram per mil level; But these data were extremely reliable. We thought there would be more oxytocin present at the time of the nursing, but the highest levels were reliably found in the pre-nursing sample. Possibly this is because the women were thinking of nursing when asked to take the first sample of each set. So they were probably primed and really releasing the hormone before they actually nursed the baby, which would have made nursing easier. By the time they got the baby to the breast, they then gave us a second sample, and then 30 minutes later. Each woman -- there were 12 women - gave five sets of three samples. Unlike corticosterone and cortisol, there was no indication of a circadian rhythm. There were

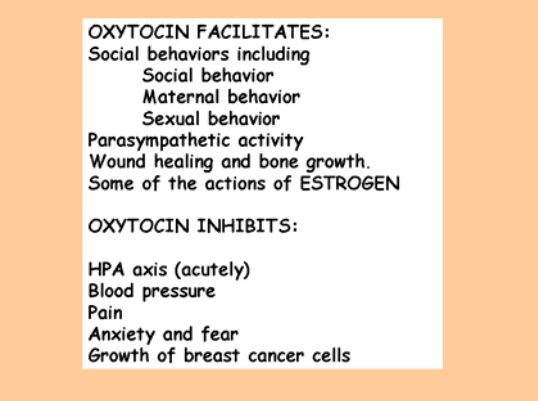
individual patterns, and it didn't seem to matter what time of day it was taken. These data are preliminary, but are highly statistically reliable.

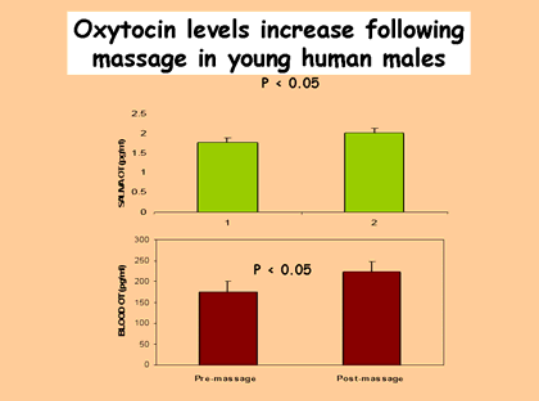


In studies done at the NIH, thanks to Margaret Altemus' good clinical skills, we were able to do a study of lactating women and use them as kind of a model for oxytocin, comparing them to bottle feeding women. I'm just showing you a few examples. These studies, which are published, involved repeated measures. These were women who had an IV catheter, which, once in place allowed noninvasive sampling. I'm showing only a few high points of these studies. Bottle-feeding women had higher levels of norepinephrine. Bottle feeding women had higher systolic blood pressure. It's not that they have "high blood pressure," it's just higher. But it was higher than the control women who were in the follicular phase, and this has been replicated since by several laboratories. Notice that the control and the breast-feeders are very similar on blood pressure.

So what's happening here - and this is true in the next data example, which shows heart rate - bottle-feeding women may be a model for chronic stress, since they went through pregnancy and birth and then did not having the opportunity to experience lactation. This is a fatally flawed design since women were not randomly assigned to these groups. However, thanks to Markus Heinrichs we know that within women who are nursing you get many of the same changes seen in bottle-feeding women just by waiting a longer period of time since the last breast-feeding episode. We also stressed the women in our study using exercise and looked at hormones such as cortisol. Cortisol was not really different between groups before they exercised, but after they exercised, the women who were bottle feeding had higher levels of cortisol. ACTH and vasopressin were even more elevated than cortisol in women who did not nurse; both were lower in breast-feeding women in response to the stressor. Even mitogen-stimulated lymphocyte proliferation is actually higher in the breast-feeding women, in this case probably indicative of better or more enhanced immune system. So lactation, mediated in part by oxytocin -- remember, that's far from the whole story here -- may allow a new mother to manage stress more effectively, buffer between the physiological state of pregnancy and the postpartum period (possibly in part again through oxytocin), allowing even less reactivity to or more appropriate reactions to stressors.

Oxytocin has many newly discovered and interesting functions besides the social ones. For example it may increase parasympathetic activity, wound healing and bone growth. Some of the actions of estrogen are also being affected by oxytocin. Oxytocin inhibits the HPA axis acutely, blood pressure and pain are reduced, as is anxiety and fear, and even the growth of breast cancer cells. I know a number of you are interested in breast cancer; and there's quite nice literature mostly coming out of Italy showing in different cell lines that oxytocin has a powerful anti-proliferative effect on those kinds of cells.

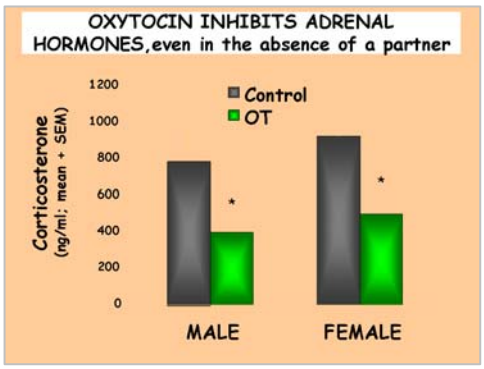




One important question that often comes to mind in these studies is what can we do other than breast feeding to enhance the release or effects of oxytocin. One possible answer, based on work in rats, is "safe" touch or massage. To examine this possibility we also used the same assay described above and in this case in young men, from whom both saliva and blood measurements were taken before and after shoulder massage. Even with small samples - there were 15 samples in each of these bars you can see - a significant increase was measured in oxytocin after massage in both saliva and in blood.

The main focus of my research is neurobiology, which is not easily studied in humans. When I started these studies, which is now almost 25 years ago, and I was interested in an animal that I showed you before, called a prairie vole. Prairie voles are unique because they form long-lasting social bonds. Social bonds are not formed in most rodents, at least not in a way that we humans would call them social bonds, but species that form social bonds are rather rare in mammals. About 3 to 5 percent of mammalian species can be described as socially monogamous.

I was at the University of Illinois Champaign-Urbana when I began to work on voles, so they were literally out in the local fields, and that's how we started working with them. Luckily I had some wonderful collaborators who really maximized the usefulness of this socially bonding creature because they provided us with knowledge of the natural history of prairie voles. Prairie voles do exhibit social bonds in nature, lifelong social bonds. Both males and females have high levels of infant care. You can see the father is combing his hair and holding the baby in his mouth. We could study them in nature. These studies were started by Lowell Getz, who was my collaborator and an ecologist at the University of Illinois. Based on over 25 years of field work we could prove that prairie voles really did live in social bonds and social pairs. And also we found that we could move voles into the laboratory and get pretty much the same thing we saw in nature. Please note that this story has even reached the cover of the New Yorker. This one says under the red "Playground of the Giant Vole, A Vole Is Not A Rat." And that's for sure.

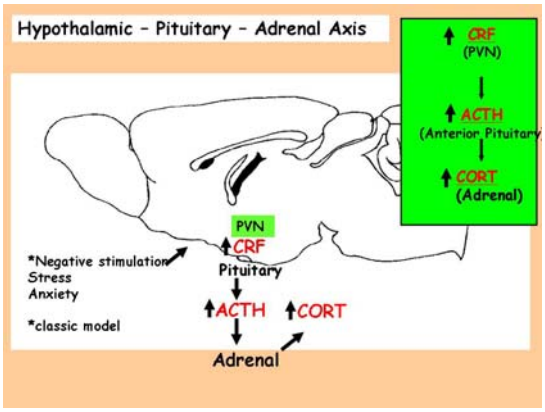


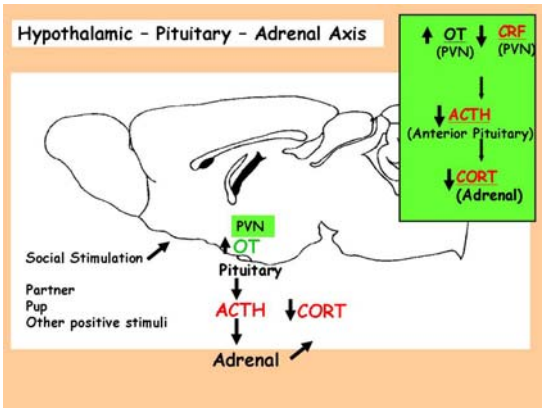
So, do prairie voles have more oxytocin? We finally developed an assay that was sensitive enough to allow us to work with the tiny amounts of blood that voles have. And the answer is yes. Blood levels of oxytocin in prairie voles are higher than in rats. And there tends to be a sex difference. We don't always get it, but usually females have more.

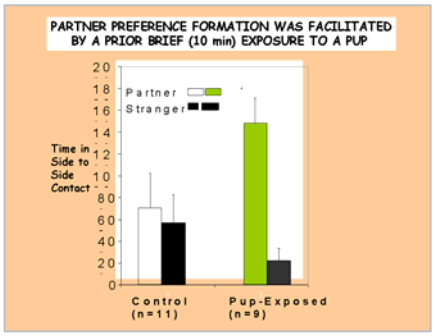
One of the things that we're interested in here is not just the social bond but its absence or the consequence of taking it away. This we can study by isolating the

animals. One predicted finding was that CRF would increase. In addition corticosterone was increased. To our initial surprise, oxytocin actually went up after isolation - not down, as we predicted. I know the data are good whenever they're exactly the opposite of what I predict - which is usually the case. New cell birth (neurogenesis) actually stops when animals are isolated. Animals stopped drinking sugar, apparently becoming anhedonic. Based on a new technology developed by my postdoctoral fellow Angela Grippo, in collaboration with Steve Porges, we also found that prairie voles have a human-like autonomic nervous system and high levels of vagal tone. When voles are isolated, their vagal tone also dropped.

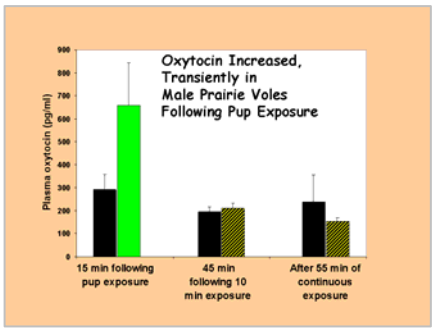
So we began to study the effects of social stimuli on both pair-bonding and the release of oxytocin. I will skim over that work here and just tell you that oxytocin does in fact facilitate pair-bonding as well as parental behavior. This finding led me to the write a paper for Psychoneuroendocrinology in 1998, that Markus Heinrichs mentioned earlier. We found that putting animals together with a member of the opposite sex (a new partner) produced a decline in stress hormones. And then we gave oxytocin in lieu of a partner, and we got exactly the same thing as pairing - a drop in corticosterone. We have all kinds of evidence that this is receptor mediated. So our model here is that in some interesting way the so-called HPA or stress axis is modulated by oxytocin, probably through effects of the level of the brain as well as the pituitary and the adrenal.





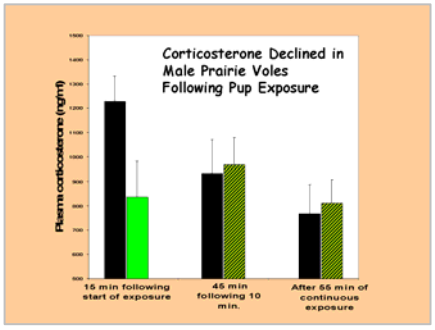


When a boy vole meets a girl this triggers a very complicated neuroendocrine cascade. However, many of the same changes can be seen using as a social stimulus an infant vole. We can show that if you have a few minutes of exposure to baby and then put animals into a pair-bonding situation, they form a new pair bond faster. We were able to measure the oxytocin in this paradigm. Oxytocin in blood increases in response to baby within 15 minutes and has returned to baseline within about 45 minutes. The oxytocin in blood is broken down quickly, but the effects in brain probably last longer. Corticosterone had the opposite pattern of change. We are still studying this model, using it to try to understand the protective mechanisms through which social experiences and hormones such as oxytocin, may allow individuals who have "social support" to more effectively manage stress.



And so I leave you with as many new questions as answers. What are the health consequences of social bonds and perceived social support? There are many. Are these mediated in part by oxytocin? Possibly, maybe even probably, but there's lots left to do. Can oxytocin be used as a biomarker? Possibly. We do not yet have good strong answers to all of these questions. However, neuropeptides can be measured in peripheral fluids and hold potential for indexing at least some aspects of the functions of the central nervous system. I also think asking these questions - especially in the context of a knowledge of the nervous system allows us to ask better questions about behavior and especially about the health consequences of behavioral experiences.

So I want to thank you for your time and attention and welcome you to include measures of the nervous system and the brain as part of your own questions. Thank you very much.



NIELSEN: Thank you for a fantastic talk and a great review of this area of research. I have two questions actually. The first one relates to this finding of decrease in oxytocin during lactation and increase in oxytocin…

CARTER: Do you mean why did oxytocin fall at the time that a woman was actually nursing?

NIELSEN: Actually nursing, and then the increase during social isolation and sort of related to Markus Heinrichs' comment about oxytocin reducing anticipatory anxiety. Shelley Taylor has talked about oxytocin as sort of a social motive. When you're in a state of social isolation, I guess oxytocin levels are high. And I was wondering what your thoughts are on that model and the idea of oxytocin being sort of a motivator of social approach that may not really be related to the social reward or the rewarding aspects of the social interaction. If I can sort of hold that for a second, I have one other short question about whether there's anything known about the role of oxytocin in maintenance of relationships over the long term and whether -- as I understand from your work there's a lot of knowledge about the early establishment of social bonds - say, in long-term monogamous pairs the mechanisms perhaps become shifted to perceptual or habit kinds of behaviors. Do we know anything about that?

CARTER: No, not really. Even from maternal behavior, which is much better studied, it was said for a long time that only the induction was oxytocin-dependent, but Cort Peterson, who started that whole story, has changed his mind, and if you do interfere with this hormone, you can at least start to see maternal behavior not being maintained. The pairbond maintenance-oxytocin relationship question - I don't think anybody has done that. We do not have the right biochemical tools at present to study the long-term effects of oxytocin. Zuoxin Wang and his colleagues have worked with dopamine, which is another piece of this story, and there's a shift between initiation and maintenance that's based on D2 versus D1 receptors; but with oxytocin the problem is to maintain the chemical. Remember, I said we had to put it into the nervous system through a cannula, and to prove it's really there we have to sacrifice the animal and look at where it was pretty quickly after the cannulation. Obviously none of this has been done in humans.

So we don't have a good model. Until somebody produces a long-lasting agonist or antagonist that will pass the blood brain barrier, and I think they're being made now, we can't ask those questions

properly, so I'm sorry. I would love to know the answer. Of course they're wonderful questions, but we don't know.

NIELSEN: If I could ask a follow-up on that about the social motive role of oxytocin.

CARTER: I assume you are referring in part to the work of Shelley Taylor. Dr. Taylor and I have been in communication for a long time because she became interested in the prairie vole story, especially some of our findings on sex differences in the response to stressors, and tried to apply these ideas to humans. In her recent studies of humans, she has results that seem to mirror those we found in voles. Contrary to her prediction (and ours) Shelley anticipated, I believe, that women with less social support would have lower levels of oxytocin. In a published paper that appeared about two months ago in Psychosomatic Medicine (Taylor, et al, 2006) done in older women, she found positive relationships between blood levels of oxytocin and the degree of anxiety and other negative life events. I wasn't expecting isolation to be associated with high levels of oxytocin in voles. She was presumably not expecting isolated humans to have lower oxytocin. However, I think both Dr. Taylor and I would have to assume that oxytocin and the processes that it regulates help the individual cope with negative events. Whether this is true or not remains to be proven. However, the same rationale has been used to explain the actions of cortisol. Cortisol is not necessarily a bad hormone. It may be released during stress to protect you. However, it is important to differentiate between acute and chronic experiences or exposure to hormones. The effects may be very different. Whether this applies to oxytocin or not is unknown.

FRASIER: I want to introduce our luncheon speaker. Just to give you kind of a background, in the past couple years, for luncheon speakers we've tried to bring in somebody who is outside the realm of biomarkers altogether so it's kind of fun and interesting, something you might not know a lot about. Two years ago we had an FBI representative from Chicago come and speak about biological collection within the idea of forensics and crime scene investigation and last year we had a NASA physician who was considered the first physician in space and collected biomeasures and biological data in space and dealt with it that way. So this year we wanted to introduce the topic of biometrics. I'm sure you've been introduced to this in movies and just hearing about it in the news in terms of how it will be implemented kind of in our daily lives to come, which is hopefully what Mr. Blackburn will tell us all about. I wanted to just give you a little background. Everybody has his bio in your folder, but just to tell you, he is the FBI's Agency Representative to the National Science and Technology Council at the Office of Science and Technology Policy in the Executive Office of the President. So he works with Homeland Security and national security-related issues and is an expert on biometrics, and we're really happy to have him here with us. (Applause.)

Overview of Biometrics and United States Government Biometric Activities

Duane Blackburn

BLACKBURN: Thank you all for being here and for putting up with the adjustments we've had to make at the last minute. I am excited to be here. I think there's a few things that our communities can work on collaboratively, and we'll get into that in a little bit, but first off let's take a little bit of time and just let everybody understand what the heck the Office of Science and Technology Policy is so that you have an idea of the perspective I'm trying to come from as well as what biometrics is. It's going to be a somewhat jovial presentation as we go forward. I don't like to be too dry, but I'm not the best humorist, so I get some help in the presentation as well. As we go throughout, we'll have a few pop quizzes, mainly to keep your attention because it's right after lunch but also just for some discussions, and then there will be a final exam at the end, so just to make sure that you do pay attention.

First off, what is the Office of Science and Technology Policy? Well, it's one of those weird offices that hardly anybody has ever heard of, but in essence it is the President's science arm. So the director of OSTP is the President's science adviser, and there's about 30 of us that work in the office. You can see hopefully the four main areas that we tend to work on, first and foremost to advise the President and others within the Executive Office of the President on science and technology issues while they're forming policy. We also take a look across the different government agencies and identify and prioritize different S&T issues that we want to work on for specific administrations and then work to coordinate those activities across the individual government agencies. So we're trying to really just break down the stovepipes that are inherent in federal government because of the way that agencies and departments are set up and the way the budget processes work. And then we also work with the private sector and represent the U.S. in international negotiations on technology issues. The website, if you're interested in more about it, is OSTP.gov.

Now, what is biometrics? Biometrics is one of those terms that gets thrown around a lot and it means different things to different people. We, as the federal government, have chosen to define it this way. First off, it can be described either as a characteristic or as a process. As a characteristic, it's a biological and behavioral characteristic that can be used for automated recognition, so recognition of who that individual is. Now, the key word in there is "automated," so DNA at this point is not considered a biometric because it takes so long to do the analysis and make the match. The other key component on this that's not specifically spelled out but I want to make you aware of as well is that we're interested in recognizing humans. Some people use biometric-like things to identify chickens and cows. That's all fine and good. I prefer to eat them. So we don't deal with those so much. Sometimes when people talk about biometrics, it's also talked about as a process, so I want to do a biometric comparison of someone, and that's just really taking the same things and turning it from a noun into a verb.

Biometric tasks. This is important to keep in mind as you go throughout. There's really two tasks that you have to keep in mind. The first one is when someone makes a claim. So I can walk up and

say, "Hi. I'm Duane Blackburn," and because I was just introduced, hopefully you-all will go, "Okay. Yeah, that's that idiot Duane Blackburn again. All right. Fine." Rather than me coming up and saying, "Hi. I'm Alex Trebek." Hopefully you would know that I'm not Alex Trebek. But that's still an example of verification. Identification, on the other hand, is if I don't make a claim and you're just trying to figure out who I am. So you would take my face image, my fingerprint, my iris, submit them to a database, make scads and scads of comparisons, and hopefully the system will tell you who I am, or at least who I was enrolled as. I guess that's another important thing. Biometrics isn't necessarily who you are. It is who you were originally reported to be. So if you're enrolled incorrectly, it will always tell you that you were someone else. But each of these tasks has different statistics, so if you wanted to go in and look at biometrics from a more scientific standpoint, each of those has different statistics for verification of the two sets of identification, which we won't get into the specifics here on, but there's some reference material that I'm going to tell you about here in a little bit. So just make sure that you're reading the right things into the research that you're doing, otherwise it doesn't make a lot of sense.

You'll also hear a term called "recognition." That's pretty much a generic term, biometric recognition. It doesn't necessarily mean verification or identification. It's just kind of a conglomeration of both of those terms. So if you're talking in a very general sense about using biometrics, people say biometric recognition rather than verification or identification. Biometric systems themselves consist of multiple subcomponents. You can see them on the screen here. First off, you take a presentation. I'm going to step away and just yell for a little bit. You start off with a presentation. For some biometrics that can be rather complicated with the fingerprint or if you're trying to get the retina. However, if you're using something like face recognition or iris recognition, it's a picture. That's as complicated as it gets. Then it goes through all these different sub-processes that really don't mean much to anybody unless you're trying to operate the system. Then you get into the signal processing, and this is where the magic equation happens, where you've taken your image and you've turned it into something that's electronic and now you can compare it to the previously enrolled ones. And then you can see that there's all sorts of other issues that are going on as well, such as, data storage, or dealing with verification or identification, there's some variances there. But in essence this is how just about all biometric systems tend to operate.

Let's talk about the individual modalities a little bit. Fingerprint. This is the most widely used. Some people say that fingerprint biometrics has been around for 100 years or so. That's not entirely true. Remember that biometrics is automated. So they have been doing fingerprint recognition for over 100 years, but not as a biometric, as an automated computerized process. There are various ways of doing it. The traditional way is having someone take your fingerprint, you know, take your arm and roll your fingerprint in ink and then roll it back over again, and this has been typical for several years, just trying to get those nice ten rolled fingerprint images. Over the last five years or so, we've gotten rid of the ink part, and so you can now have that guy break your arm but without getting your fingers all gooey and sticky. So that's a step in the right direction. You also have some latent fingerprints. These are the types of things that you hear about in the press where we picked up a fingerprint off of an improvised explosive device in Iraq that didn't blow up. It's just one of the things that, as you're merging with stuff and talking about it, you leave your fingerprints. And this stuff - duct tape - is the absolute best. I didn't even try to have this prop here, but it worked out great. As you see, you have to unroll this duct tape and you really have to stick your finger on it to go. I have a perfect fingerprint on the underside of that. So all these morons making all those improvised explosive devices are leaving their fingerprints on everything that they do, and we catch a lot of bad guys that way (indicating). There's a difference between rolled, flats, and slaps. We talked about rolls a little bit. That's the big, hairy, mean guy who's breaking your arm and you're saying, "All I want to do is coach Little League" (indicating). (Laughter.)

See, I'm not the best humorist, but I try. So that's the traditional. We also get into flats. You have four down, four down, and then two down. So that's the newest iteration of trying to get ten-fingerprint systems. That's much nicer because anybody can do that. That's pretty simple. And then slaps is a little bit different. That's just individual fingers, putting it down. Most of the commercial systems that you buy - that you log into your laptop or something like that - are just a single fingerprint slap (indicating). How many fingers you need is going to vary by application. The FBI requires ten rolled fingerprints traditionally. Over the last couple of years they transitioned to ten flat fingerprints. US-Visit is two slaps. In general, ten rolled fingerprints is always going to give you the best accuracy and the single slap is going to be the least accurate, but you don't always need ten-fingerprint accuracy for all your applications. There are multiple standards out there to be aware of. That's pretty boring, so I'll skip it.

Technology evaluation. The government has run a few different evaluations of fingerprint technology. You can see the latest one that's up here, 2003. Three systems vastly outperformed all the others. It's actually fairly embarrassing. The National Institute of Standards and Technology offers free from its website a fingerprint recognition matcher so that people can do research, and half the commercial systems out there performed worse than what you can download for free. But the top three matchers were significantly better than all the rest. Additional fingers greatly improve accuracy. Bad data in; get your bad data out. I'm sure it's the same with everything that you-all deal with here.

First quiz time. Watch this and we'll talk about whether this is actually possible or not. (Video clip is shown.) All right. This is the first quiz, so I'll make it a little easy on you. This was an identification application. He made no claim to who his identity was and he stuck one thumbprint on the sensor and it let him in. Now, we're going to do a show of hands. Do you think it's possible or not possible? How many think that's possible? (A show of hands.) All right. Not possible? (A show of hands.) All right. I'm a nice guy, so you're both right. It's a little of both for this first one. Doing something like that, it would be possible to have a single fingerprint in identification if it's a very controlled setting, such as if you only have 50 people or 20 people trying to get access to that door. There's no way you would be able to get a single fingerprint over millions and millions of people and have it identify who you are. The other part that is not quite accurate is how he stuck the fingerprint on. If you noticed, he got the very tip of the fingerprint rather than where the actual data is. So it was a little bit of a trick question there, but you both got it right the first time. That's good. The tests are going to get harder as we go on. (Laughter.)

Let's move on to a different one. This is face recognition. Now, this one has been around since about the 1980s. Actually into the last 18 months or so, the algorithms that were existing, even commercial products, were still based on the original algorithms that were written in the '80s, and so what we're really seeing now is transitioning from very low-resolution images to high-resolution imagery, and that's really gaining a significant increase in performance. We're also starting to see some nice 3-D laser scans and that type of thing of faces that's really trying to get some improved accuracy going. Is it better than humans? Well, that really depends. Humans are very good at recognizing people that they know. So if I put up a picture of a Hollywood celebrity, most of you are going to know who that is. If I put up a picture of me and left the room, you may not know who I am, even though you've seen me. There are some other psychological studies that have gone on. Humans individually are pretty good at recognizing people that are within their own race or within races that they typically deal with and are not very good at recognizing those that they aren't necessarily associated with on a regular basis. So that really makes some of our applications pretty difficult to do. If you're not using automated face recognition but you're just trying to take a picture on that driver's license and compare it to that person real quick, it's not a very easy job to do. (Video clip is shown.) So human face recognition isn't all that easy to do. So, even though automated face

recognition may not be 100 percent accurate in all the tests that we're doing, you can see that it's still going to be beneficial for real-life operational conditions.

Let's take a look at a few different tests. These tests are four years old. We're currently doing the next version of this, so we'll have some updated data for you. Again, three systems were better than the rest. I have no idea why three keeps popping up. It just happened. There's some interesting data for this audience. Older people are easier to recognize than younger people. You get more distinguished as you age, I guess. And performance decreases as the time since enrollment image increases. Let's take a look at those in a little bit more detail. This is identification on a database of 37,000 people. You can see the age bends and you can just kind of see that gradual progression up: as you get older, your face becomes more distinguished and it's easier for the automated systems to track and try to figure out who you are.

Contrast that to the difference between enrollment and application. The example I like to use is, if I took her picture right now and then took another picture of her, the system is going to work pretty good. However, if I grabbed her middle school yearbook picture and compared it to her face right now, it's going to be different. So that's that time difference that we're talking about. You can see, again, a gradual trend as days go by. This is days. The first column is from zero to 60 and the last column goes up to about 1,100 days in between that original enrollment. It's a fairly steady, noticeably statistically relevant trend going down there. So that's one of the issues that we have to face with face recognition technology.

Some of the other issues we face with face recognition are a little bit more difficult. Performance decreases automatically 50 percent just by going outdoors. So if you've got 80 percent, you're doing pretty good, wherever you're measuring that 80 percent, and if you go outside, it's going to be 40 percent. Performance decreases about 70 percent going left to right and up and down as well. So that's why you've heard in the press that the State Department always wanted those nice frontal images on the passports, because if you get an image like this, it's not going to do that much good for the automated face recognition the State Department is doing (indicating). It's not as accurate as fingerprints, or is it? Well, that depends on which community you ask. The fingerprint people say, "Heck, no; face recognition is bad, is terrible." And the face people say, "We're doing pretty good, but you're really comparing apples and oranges," and their point is, fingerprints, even on the newer sensors you have very controlled lighting and distances when you're capturing that image; however, with the face, you just throw any old random face image in and you expect to get good results out. So really the more accurate apples to apples comparison would be comparing latent fingerprints to face recognition, and those are very similar in accuracy rates.

Test time. (Video clip is shown.) One hundred percent confirmation. All right. A show of hands. Technically possible? (No hands.) Not technically possible? (A show of hands.) Broke several laws of physics? (A show of hands.) (Laughter.) There we go. We've got a few of those. It's really that last group. Name some things that you saw that were wrong. You're not going to be wrong. Trust me. It's pretty easy.

AUDIENCE MEMBER: Moving.

BLACKBURN: You're moving. Okay. That definitely causes image quality issues.

AUDIENCE MEMBER: Outside.

BLACKBURN: Outside. That's very difficult.

AUDIENCE MEMBER: Film camera to capture thermal images.

BLACKBURN: Yes, thermal. That's pretty good. They used film for thermal. The other interesting thing was they used thermal imagery, that you can see through a car, which is your first breaking of the law of physics, and that somehow that thermal image from this angle became a visual image from this angle, which I don't understand how they do, and, again, they got the 100-percent confirmation on identification. Interesting plot line there, but not a very good, useful application of face recognition technology. Keep that in mind for the final exam.

Okay. Iris recognition. It's iris, not Irish. I have a slur a little bit. Some people think I'm talking about Irish versus Scottish. It's iris, that little colored portion of your eye. It's not the retina, it's not the back of your eye, it's not the sclera, the white part. It's that colored portion of it. They illuminate that using near-infrared lighting, which is just barely outside the range of visible light. In fact, some humans can actually see it. Most traditional camera systems will pick up that illumination. You might have heard about the laser in the eye myth for iris recognition, you know, scanning your eye with these lasers. No. They're taking a picture. It's just that they're using some illumination that doesn't hurt your eyes. How close you have to be is a question. A couple of years ago you had this little wand system that you had to hold in your hand and it had a little mirror in it, just like the back of my watch, and you had to sit there and you had to really line up your iris so that it could see it. Nowadays you can just kind of walk up to a system and you kind of hold it 12 to 18 inches away and it works pretty good. A company actually demonstrated last year a walk-through portal. As you're walking through a portal, at a distance of about three feet, if you kind of look in the general direction, so I'm walking and I say, well, I'm going to look at that pole, it's going to be able to capture me. If you think about it, that's pretty easy. You're in a boring airport. You put a nice poster there. It gets everybody's attention. What's that? They got your iris.

So you're starting to get that greater distance out of iris recognition in there (indicating). There are a few other issues to deal with with the iris. The actual color of the iris has a big play in it. If you have very dark irises, it's not going to pick you up very well. Iris systems tend to love me. I kind of have dark irises, but not that much, but I have very large eyelids, if you hadn't noticed already. My eyes kind of bug out is what my wife says. That's a good thing for the iris. Contrast that to some Asians who generally don't have their eyes naturally open as much, and that causes significant issues, because now, instead of looking at the whole iris, you have those corners to deal with. But, still, it's very accurate (indicating).

Next pop quiz. (Video clip is shown.) Show of hands. Do you think that's technically possible? (A show of hands.) All right. Not possible? (A show of hands.) I'm going to vote down the middle on this one again. When I first made the slide a couple years ago, I laughed at it and said that's not technically possible. Nowadays it might be. Probably not. You're going to be able to get that distance, again, with that walk-through portal that I talked about. If the database is fairly small, it's going to work. Iris is definitely on the par with fingerprints as far as accuracy goes. Now, that lady who walked through with her eyes closed, if you noticed that, you're not going to get that one. But it's possible. I would not think that this is going to happen all the time, but it could happen every now and then.

AUDIENCE MEMBER: Wouldn't it also make a technical difference between something here where the system is actively searching for one individual as opposed to the other scenes where you have random people walking through a mall and it's supposed to hold them up in real-time and give individualized things for any one out of a sample of 300 million people?

BLACKBURN: The question is coming from an audience member who has watched the movie, but it is a good question. They probably operated this more for verification, so as opposed to comparing it to a huge database, they were looking for one specific individual, and yes, that would make it much,

much easier to recognize and to find that individual. Good observation. Nobody has ever thought that up before.

AUDIENCE MEMBER: How about identical twins?

BLACKBURN: They have different irises, different fingerprints. Some have similar faces. Some individuals are two-faced to begin with. But twins are generally not an issue. They have different biometrics.

AUDIENCE MEMBER: How fast are the systems that you're talking about doing those kinds of searches?

BLACKBURN: The speed of the systems, that's going to vary on a number of factors. If you're just trying to do a verification, you walk up and say, all right, I'll swipe my card and then say, "I'm Duane Blackburn," and put down my fingerprint, and by the time you put down your fingerprint, it has made the capture and made the analysis. Picking up people in Baghdad and Afghanistan, getting their fingerprints- and, of course, you've got to go over the satellite and back down to West Virginia and do the matching, which takes a couple of minutes, and then send it back over, you're talking about –

AUDIENCE MEMBER: How fast are those systems in West Virginia? I mean there are some very, very fast data systems.

BLACKBURN: Oh, yeah. I forget the exact numbers, but it's on the order of magnitude of a couple of minutes for a criminal check and a couple of hours for a civil check, somebody checking to coach a Little League or soccer team. That's comparing a database of about 40 million individuals in a couple of minutes. So it's very fast.

All right. A couple others we're going to blow through real quick. There won't be a quiz on these, but we'll go through them real quick. Hand geometry. This has been around since the 1980s. It's verification only. There's not enough data in it for identification purposes. The way this works is you stick your hand in that box and it has a mirror in it and it takes a picture from this view and from this view. It can figure out the lengths between your digits and the widths of your fingerprints and make a verification claim. So you have to say, "I'm Duane Blackburn," stick your hand in, and it will make that claim. It has been around since the 1980s. They have really never changed the design. A couple years ago when face recognition started saying, "We're making 3-D faces," they decided that they (hard images) were 3-D as well, but the designs are the same. It's very fast, very easy. It's soldier-proof. It doesn't take much training to say, "Stick your hand in." There's little pegs and stuff, so it's a little bit, but it's not too difficult at all. It's quick. It's easy. If it's not a very ultra-high-security area. It works perfectly fine (indicating).

Voice recognition. There's a difference here, voice or speaker versus speech. Speech is the words that I say. Voice is the resonance that's coming out of my vocal tract. This has worked fairly well. The picture towards your lower left is actually on the border. Up until 9/11, what was then INS had a system where people could cross the borders without going through security checks, basically making the claim and then speaking into the phone, saying, "I'm Duane Blackburn," and they would do that voice comparison and let you through. That has been replaced with other systems by now. The picture on the right is an attempt that we had done five years or so ago of voice recognition for Title III wiretap cases.

The Air Force has been using voice recognition for several years. If someone says, "Roger," and they're flying around in an airplane, they want to figure out which pilot is saying, "Roger." Now,

that's a pretty controlled environment. Ideally no one besides your people is going to be flying your airplane around, so you have a very small database that you're trying to deal with, but it worked fairly well for that. For looking at the wiretap cases, for our purposes, trying to figure out who a drug guy was talking to, it didn't work at all. The channel variances just absolutely killed us. If you enroll on a microphone and then say it from a cell phone and then you're trying to compare that to a landline phone, there's channel variances that just totally trash the audio system, the signature of it. So that didn't work at all. And then the different transmission mediums, in addition to the different microphones, cause all sorts of issues. But it works for certain applications. It's still very heavily used in prison systems. Telephone time is something that's bartered amongst prisoners and bartering is not encouraged in prison systems because it's a lack of institutional control, so they will monitor phone calls to make sure that they haven't dialed in and punched in their access code and then handed it to their buddy. So it works very well for those controlled situations (indicating).

Now, there are a few others that are out there in various stages. This is actually called dynamic signature. It's looking at your signature, not necessarily what it looks like but the pressure that you use and the speed and the loops that you make for your signature, and that's what those curves are showing. It works fairly well. It's going to be for verification purposes only (indicating).

Ear biometrics. Really there are no biometrics for the ear right now. There are several people over the last 20 years who have proposed to use the ear as a biometric. Forensics people love ears. They absolutely love ears. Remember when I said earlier that the State Department wanted that frontal face image, well, the FBI forensic folks hated that because they get more information out of the ear than they do from the front of the face. So they lost capabilities in that tradeoff. But there have been proposals here and there for people to automate that process. It just hasn't kicked off, but my prediction is that it will at some point.

Retina is another one that kind of comes and goes. I think right now there actually is a commercial product. This one suffered very heavily from the laser-in-the-eyes routine. It didn't use lasers, but it did take a picture of the back of your retina, and if you go to your eye doctor and get that blinding flash of light that makes you walk around staggering for a couple of hours, they're taking a picture of the back of the retina, which is actually where I stole this image from. So it suffered from the laser-in-the-eye myth. Early on you really had to stick your eye down on top of the thing, kind of like the old camcorder systems. Remember that - that little plastic thing you had to stick your eye on? Now, if you're standing in line and you see some slimy guy and then you're next, you're not going to like that. So people didn't tend to like that. The latest systems are a couple inches away, so that's pretty good. There were also some issues with this. You can tell an awful lot about a person by looking at an image of the retina. That's why your eye doctor takes a look at it. In addition to eye health, you can tell really if you're on any type of drugs or alcohol dependency, stress. There are a lot of privacy concerns with the retina, so I don't know if that one is ever going to become commercially viable.

Iris recognition at a distance. This is an iris taken at a distance of 10 meters. There's not enough data in there yet. So that's more of a sensor and collection issue.

This one is gait or body biometrics. I'm going to have to walk away from the thing here. It's the way that you actually walk. A lot of people look at the gait just from the waist down and the way that you move on things, and other people try to look at your entire body. You maybe swing your arm a little bit. Of course this is going to vary a little bit whether you're carrying a laptop on this side or on this side; or if you're wearing high-heeled shoes, it changes your gait. It's certainly a potential biometric system that people are very heavily doing research on right now (indicating).

Facial thermography. We laughed at this earlier, but people have been doing a lot of studies on this. There was a commercial product five or six years ago. It didn't go too well. I don't know if anybody

is trying to do that right now or not. It's actually looking at the heat that resonates off of your face. They say that these patterns are consistent. I don't know if I buy it yet or not, but some people are working very heavily on it (indicating).

The last one is the one that gets most of the jokes. I was a little disappointed that the joke was already taken this morning. Body odor. I kid you not. There are people actually doing research on body odor as a biometric, but, again, research I'm glad someone else is doing, not me.

Here are some background references if you want to find out a little bit more information about biometric technology. First off I want to tell you about the Biometrics Catalog, biometricscatalog.org. It works for both the English and the European spelling versions of the catalog. This is the government's one location where we really try to make a library of anything that we know of that's publicly available about biometrics. We place it in the catalog. You don't have to register. You don't have to pay a fee, nothing. It has got research reports, an updated newsroom every morning, commercial products, research and evaluation reports. It's a very good resource to try to find information on it.

Our White House-level working group on biometrics recently announced the release of a bunch of what we're calling foundation documents, trying to provide the public a foundational understanding, similar to what you're getting today, on biometrics technology, and these are available in the catalog- but basically Biometrics 101, essentially what we just ran through and the frequently asked questions and glossary. Those are all written at a very introductory level. Ideally our parents and our children will be able to understand those documents. It's taking a very technical item, dummying it down an awful lot, really taking off some of the really hoity-toity scientists and making it understandable. Then we also have some others that are kind of in the middle range on statistics, testing and evaluation, and then some modality-specific papers, so a five-page paper on face recognition or a six-page paper on fingerprints. The other document that I think you all would be very interested in, actually because you brought it up some this morning, and I just called it "Privacy" here, but the full title is "Biometrics and Privacy: Building a Conceptual Foundation." The reason I think it would be interesting for you-all is because your privacy concerns are very much related to the privacy concerns that we have in biometrics in that the privacy concerns are that informational type. So what we tried to do with this document is study biometrics, study informational privacy and how to use different theories in your research, and then combine the two, building that conceptual foundation so you can have privacy-protected biometric systems. So I think you'll be able to get a lot of information out of how to do that for some of your studies. Even if it's not completely related or related at all to biometrics, it will still be an interesting read, so I recommend doing that. Again, those documents are all on the biometricscatalog.org.

These are some examples of federal activities. I talked about the voice system that's no longer there. That's because you now have the US-Visit. All visitors coming into the U.S. have a fingerprint check done at this point, and it's two fingerprints right now, so one finger, the other finger. We're in the process of changing that to be ten fingerprints so that we can better match or compare the visitors to the FBI's bad guy database. Right now it's kind of done by sneakernet and UPS, but we wanted to make an automated system. So in the future, the first time going through you're going to have ten fingerprints as opposed to two going through. That has been a very successful system. Some people say, "Well, you haven't caught any terrorists." Well, terrorists read the newspaper. They're not going to come through the airports, but they were before. But what we are catching is a whole lot of international fugitives, rapists, murderers, child-molesters. We're catching a lot of bad guys that we don't want in the country. So it's a good thing. Registered Traveler and TWIC, and that's the Transportation Work Identity Card credential. (They keep changing the name.) These are other biometric card-based systems where you can pay a little fee and volunteer to be on the list. They do a little bit of background checks. So you get kind of expedited through the security at airports. With

TWIC we want to make sure that people who have access to our critical infrastructures, primarily the transportation ones, have been cleared and vetted. They'll have a little card in there that they'll have to use, kind of swipe, and say, "I'm Alex Trebek," and then enter their fingerprints, and they will make that comparison before it allows them into the system.

DOJ. The big daddy in there is the FBI IAFIS system. We talked about that earlier with the speeds and stuff that's out in West Virginia. They are also potentially looking at incorporating face recognition into the next generation of IAFIS. Right now there is no national face recognition capability similar to what we have with fingerprints. I'll just skip ahead a little bit. The State Department does use face recognition right now. So when you apply for a passport, you submit your face picture and they compare that to all the other face pictures to see if you're trying to get a second passport when you shouldn't.

NIJ. We talked about some of the stuff that's going on in jails, and there's a video clip here of something that we had funded at a school using iris recognition technology. This is not a federal mandate. Some people like to say it is. This is actually something that they applied for a grant for and it was awarded on a peer-reviewed, competitive process. It's a pretty interesting system. They're most concerned about ensuring that people that get into the school are supposed to be in the school and it's not someone trying to come in and try to take a kid that they're not supposed to take. (Video clip is shown.) That was the vendor's statement. He wants it to become that. But it's pretty nice. This particular school system didn't have the issues where they have maybe a parent who doesn't have official custody of the child coming in and swiping it, but that has been a national problem, and they had some interest in ensuring that it didn't happen there. And, again, it was a competitive, peer-reviewed grant, and a lot of different school systems have been up there to view that. It has also helped out technologically. They had some issues with outdoor iris images, just the way that they placed it, facing towards the south sun, so you've got that mirror in there and the sun is back there and you're not going to get that nice image. So it has been a very beneficial project for us.

These are some other federal activities. DOD has what they call the ABIS, Automated Biometric Identification System. It's essentially exactly the same thing as FBI's IAFIS, except they're also collecting face, iris, and DNA. They're not matching against them yet, but they're collecting those. They are overseas, primarily Iraq and Afghanistan, at this point collecting the information. Their servers are co-located within the FBI facilities in West Virginia, so that data is very much intertwined. The State Department, we talked about the face recognition. Also international, you have the e-passports. So the next time you go to try and get a new passport, it's not going to be the same old paper-based approach. You're still going to have the papers and you'll get it stamped, but there's going to be a microchip actually inside the cover that includes electronically all the information that's on that first page of the passport, so your name and all the other information. The idea is that you would take that out, open it up, stick it on the system, and then this can be able to download from your passport your face image, take a picture of you as you cross the border, and be able to make that comparison to make sure that you are the person that was actually assigned that passport. HSPD-12, Homeland Security Presidential Directive No. 12. Now, this is an ID card, to use the British term, "scheme." U.S. term: "plan." The idea here is that we want to make sure that all federal employees are properly vetted and that I don't have to carry around 20 different ID cards anymore because that makes my neck hurt at the end of the day. I should be able to have one ID card and go to any federal facility and be able to get proper access, and we're working on that right now. Those cards are due to roll out this fall.

Here's some non-U.S. Government examples. The first one everybody likes to hear about is Disney World. For several years they've used a variant of the hand geometry. It's two-finger geometry, so just two fingerprints on their season pass-holders. They really mainly did it for convenience so you don't have to sit there and have someone analyze your card. You just kind of walk up, stick your

card in and stick your fingerprint in, and walk on through. It's a really rapid system. There was initially a lot of negative feedback on it, but pretty much all the users said, you know, "Whatever. This gets me into Mickey's house faster. I'm all for it." This past season they have instituted it for all pass-holders. This was a little bit more of a cost-savings mode for them. They had seen a lot of instances of someone going in and then chucking their card over the fence and someone else using the card to come back through. So now if you leave the facility and want to reenter, you have to not only present your pass but also do that verification check.

Retail. Over 6 million U.S. citizens have voluntarily signed up for a little system that's primarily used at grocery stores. The concept here, they did it, again, for convenience and cost savings, but instead of paying for your card and going through and swiping your card and having to sign all this type of stuff, you just kind of walk up and stick your fingerprint on the sensor and you're done. It automatically downloads the checkout amount directly from your bank account. So it's easier for you because you just slap your finger down, but it's saving the retailers money, because every time you use your credit card, there's like a fee that the cards issue, like 30 cents or something like that. However, when you do a direct deposit from one account to another, the fee is like 4 cents. So they're saving a lot of money by doing that, and it's easier for you.

Elections. Philippines and Haiti, and I forget exactly which modality each of these is using, but to vote, you do a biometric verification first. So you claim your identity, and if you don't match, you don't vote.

National ID systems. The United Arab Emirates and the UK have been working on these for several years. Now, UAE really has gotten off without a hitch. UK, there's been a lot of negative feedback in the press on that one. So we'll see how that actually goes. The Haj pilgrimage in Saudi Arabia, that's pretty interesting. If there's a Muslim in the room, I'll apologize first off if I get this wrong, but there's a pilgrimage that comes into Saudi Arabia once a year, and they're very interested in making sure that all the people can come in and worship, but they're equally interested in making sure that they all leave. So they take your iris coming in and going out and then they can have an idea of if you didn't leave. It's kind of similar to what we're doing with fingerprints in US-Visit, except they really only do it during the Hajj time of the year.

Financial services. That's something that you're going to really start to see over the coming years. We talked about the Piggly Wiggly's and that type of thing, the grocery stores, but we expect to see biometrics being used in financial transactions much more in the future, ranging from the hoity-toity selling of stocks and trades to just internet online banking as it comes forward. It's an added layer of the security onion that people are interested in using. That's really the bank's primary responsibility, to protect your money. It's not to grow your money. It's to protect your money. So if they can figure out some way to make biometrics cost-effective, they're going to implement it. So that's the overview.

AUDIENCE MEMBER: What kinds of requirements are placed on these different agencies regarding what they do with these databases? For instance, Disney World, in establishing this database, are they required to set up any kinds of practices regarding ethical concerns, release of information?

BLACKBURN: One of the privacy concerns with all of these different biometric systems is it's going to vary significantly. The federal government has several rules rooted in the Constitution, also the Privacy Act of 1974 and a few other executive orders. We're very regulated in what we do. So we have privacy impact analyses, it goes out for public comment before final decisions are made, and it's very difficult to actually share data. So you have to really show a need to share that data back and forth between government systems. Now, something such as Disney World -- I shouldn't even say

Disney World, but nongovernmental applications - there really aren't rules. There's no laws in there. It's dependent on the agreement that you make with them. So to enter Disney World, you're agreeing to give up your fingerprints, or your finger for the two-finger geometry. To use that Piggly Wiggly or whatever grocery store system thing, it's a convenience to you, but you sign a form that stipulates how they're going to use that data and that becomes the agreement. Other places, some of the stuff that's going on in the Middle East, it's: "Here, you're doing it. Tough." So it varies considerably. But for government systems, it's very regulated. We have detailed processes on what we can and can't do.

FRASIER: Along those lines, say, either Disney World or Dominick's, the grocery store, can the federal government or local governments come in and subpoena that information to use in criminal investigations?

BLACKBURN: I'm not a lawyer. The question is, could the federal government or any other government subpoena that information. Well, anybody can subpoena anything. Whether they get it or not is a different question. I can't answer that. I really don't know.

LINDAU: I was interested in the signature technology. A few years ago I was doing some work in health literacy, and after enrolling about 500 people, I made the observation that the people who were scoring low on the literacy test took a notably long time to sign their name. So we looked at it statistically. The study was not designed to test that outcome, but it did look like there was a statistically significant association between taking more than five seconds to sign their name and low literacy. Since then, we have tried to find technology that would allow us to better capture the signature characteristics, and I just wondered what your familiarity is with signature technology and whether you've learned anything other than identification about the individual by capturing that signature and those signature patterns.

BLACKBURN: The question really is, is there other information that you can get from the signature verification. I haven't studied it myself, but I would imagine, sure. For example, you could probably learn in my case, looking at my signature, that recess was the first thing right after handwriting in elementary school. So my signature is like (utterance) and I'm done. But we could certainly try to put you in touch.

LINDAU: Is that commercially available right now?

BLACKBURN: Yes, that's commercially available. There's actually two different variants. The one that I showed you is the traditional one. There's a newer one that has become commercially available that really doesn't have the same type of electronic sensors. It's a metal pin on a metal surface and it just listens. So there's different ways of doing even that at this point.

AUDIENCE MEMBER: Ten digits is successful because it's ten indicators as opposed to two. So across these modalities, are there any two or three that when you put multiple indicators together you get more reliability and more validity? Have you experimented with, you know, which two of these sort of get it down to 99.9 percent, or is that something that has been ignored?

BLACKBURN: Can you put multiple modalities together to improve your score? Absolutely. Can you put two modalities together and get a worse score? Yeah. So it really comes down to that system integration. How do you put those things together, how do you do the combination? There's been a lot of studies on that right now. Actually probably the biggest area of research going on right now is how to properly do that, and most of us concentrate on the big three - face, finger, and iris - on how to do that. Ninety-nine percent accuracy? In certain situations face, fingerprint, and iris can already do that.

AUDIENCE MEMBER: Can you talk a little bit about DNA identification and sort of the accuracy as to the databases around that? I know in the genetics world they're planning these large studies where they're planning to put a lot of gene-type information on the web, and one of the biggest concerns is issues of privacy.

BLACKBURN: The question was if I could talk anything about DNA. No. Sorry. I don't know anything about it. I did a small project in college, but, you know, I was a college student, so I was probably wrong. I really don't know that much about that, but I know some folks at the FBI that you could talk with, so e-mail me later and I can pass that on.

All right. I love these questions, but we're going to start to run out of time, so I want to get through a couple different slides just so we get to the benefit of why I came and then we'll have plenty of time for additional questions later on. First, your final exam. Now, if I did my job, this is going to be pretty easy for you. (Video clip is shown.) All right. Show of hands. Technically possible? (A show of hands.) Not technically possible? (A show of hands.) All right. Most of you are correct. It's not possible at all. You're not going to get a latent fingerprint driving around in a Mach IV, in a supersonic airplane. That's not going to happen. But if you caught his last thing, "Man, I know what I want for Christmas," so do I. That's what we want to try to get to.

So why biometrics? Why are we as an administration putting so much effort and emphasis on biometrics? We see it as the most definitive real-time tool in a large identity governance toolbox. When you're comparing it to things such as tokens and photo IDs and passwords, it's far and away better. It's more convenient. It's more accurate. More secure. It's very difficult to steal a biometric compared to a password which is written on a piece of paper and it's thrown on your monitor. You know, to be honest, it really does beat the alternative of what the future would be without biometrics. (Video clip is shown.) Agree. It beats the alternative, doesn't it? That's good. All right. Some questions for you to ponder, and I don't expect answers at this point, but this is the main reason I came here, because I would like to be able to leverage some of the expertise that's in this room and bring it back to the biometrics community. Are there standard ways of performing and reporting results from age-based studies that we need to be aware of? How do you collect data from the same individuals over a period of years to decades? In biometrics we're talking about collecting this data on thousands of people. How do you deal with the logistics, the funding issues? Are there any existing datasets that you've already collected that we can swipe? How have you handled the consent issue for minors? At this point we just don't do it. We don't do biometrics research on minors. How do you encourage the sharing of data, metadata, and the results to encourage your study by second parties? These are some of the questions that I could come up with my biometrics background, and you might have some ideas that you can share with our community. Our major conference is coming up in September, September 19th through the 21st, in Baltimore. It's the 2006 Biometric Consortium Conference. They are right now receiving suggestions for presentations on that. So if a group of you would like to get together and try to answer these questions, I would certainly look forward to hearing some of those answers presented at that conference or some other venues in the future.

Now, let's finish up. What other questions do you have for me?

AUDIENCE MEMBER: Is there, like, a computer hacker/computer system designer relationship here where, as these technologies are developed, the, quote, "bad guys" try to figure out ways to dupe the technology?

BLACKBURN: Are the bad guys trying to break biometrics?

AUDIENCE MEMBER: I mean you could probably get, like, false fingerprints. I don't know anything about this.

BLACKBURN: Absolutely. It's one of the fun things about working in government. As soon as you make a solution, then someone has broken it. Particularly when you're starting to talk about implementing these in financial transactions in the future, people are very willing to work on this. The most often quoted thing is the gummy fingers example where someone took a very, very low-end fingerprint sensor and essentially got gummy bears, melted the gummy bears, spread it out on a piece of paper, and then singed their fingerprints by sticking in their fingerprint, and then turned it around and stuck it into the fingerprint sensor. Now, again, this was a cheap fingerprint sensor. It wasn't designed for liveness detection, but it could break in. You typically use this in, you know, your child's diary or something, which is actually a commercial product. But there's stuff like that that is happening. Now, a lot of the biometrics systems have liveness detection systems built in, so you can actually, like, sense your heartbeat within your fingerprint. Your iris, even if you do this, that thing is bouncing around like crazy. It's a fairly noticeable pattern. Like if you try to fake that, you're going to be able to notice it. The face as well, you can do all sorts of stuff, like, "All right. Smile for the camera." There's lots of liveness stuff that you can do in the systems. But, yeah, there's definitely people working to hack into the systems. Everybody knows how to hack into the passwords and the PINs and all that type of thing. So it's a better solution, and, implemented correctly in a high-security realm, you're going to have multiple layers of security. So even if you break in and figure out somebody's password, you're not going to have their card or you're not going to have their biometric. Good question (indicating).

AUDIENCE MEMBER: You talked about biosharing issues, and most surveys that we make available as public-use files, we deidentify the data. It's racking my mind. How would you deidentify a dataset containing identification data?

BLACKBURN: How do you de-identify biometric data? That's a difficult question. In the face tests, the large data set was from the State Department, so we had good passport data, but for some of the other tests, we actually had to have people come in and take their picture every three weeks for a year, and, again, we couldn't make that association. It was pretty obvious for some cases, you know, that's Duane, but it's difficult, and the way we did it is we just had a printout of pictures, say, "All right. Thanks for showing up," take 15 minutes, flip through here, find you, and that gives us your identification number. You can't do that with fingerprints. You can't do that with iris. So that's an issue. Another issue, which was someone else's question, multimodal, how do you do that and protect privacy? That's a big concern right now, because if you have an association between three different signatures, you have a face image, you have a fingerprint, and you have an iris, and you want to do that research on multimodal, well, what if you happen to know the person? Again, we're good at recognizing people we know. So if my brother did that, I would be able to see his picture and say, "Hey, that's my brother's fingerprints. That's my brother's iris." Now we have a problem. Can you mix and match them? Sure. What does that do statistically, though? That kind of messes it all up. So huge issues, and we don't have answers just yet.

LINDAU: That question raises a really interesting problem at the interface of our fields, which is that, as we're collecting more and more biological data on our respondents, even if we're, say, collecting DNA data and not yet using it, the degree to which data we collect from our respondents could ever be linked with really personally-identifiable data that the government is collecting is maybe an unanticipated way that the people in our research studies could become identified. We're very cautious about biological data, but I think some of us have the attitude that things like salivary specimens and blood spot specimens and anthropometric measures that we collect shouldn't really be privileged in a way different from the sensitive questions we might ask people when we study them, and yet my anxiety has increased markedly. When I think about the possibilities of what we think is

de-identified data now to potentially become identifiable by linkage with systems like yours or Disney World's or something else.

BLACKBURN: And it gets even scarier the more you think about it, particularly from a research-based perspective, unless you're doing age-based studies and you're just trying to do human studies. And most of you all are in universities. Where do you find your research subjects? It's people who are willing to take $5.

LINDAU: Many of us find them -- they are basically as close as you get out to a random sample of the U.S. population. They're not our patients. They are citizens of the United States of America who may very well be captured in your systems.

BLACKBURN: And I heard someone earlier today say they were from Purdue University. Purdue collects a lot of biometric data for research. I would guess there are associations there.

AUDIENCE MEMBER: That's why I guess I brought up the whole DNA database, because now there's literature saying that you only need like 20 to 25 different nuclear type SNPs to be able to identify someone, and now there are capabilities of people who are going through like CODIS and being able to identify relatives and then going back and asking people to break their codes. So the whole issue of privacy is becoming very big.

BLACKBURN: Good point. Anyone else?

AUDIENCE MEMBER: Just out of curiosity, a number of us here are from the government and we got fingerprinted when we started working there. Who has those fingerprints?

BLACKBURN: Who has those fingerprints? I don't have them. West Virginia has them. FBI's CJIS division in West Virginia has that data. Now, unless you're an FBI employee, it's in the civil file. FBI employees are placed in both. So if I go through somewhere, you know, I'm going to get hit as a bad guy and they'll say, "Oh, he's just an old FBI guy." The FBI keeps track of those. I've been to the facility. It's pretty hard to get into. You're safe.

AUDIENCE MEMBER: Oh, I feel much better now.

BLACKBURN: All right. Well, it's pretty interesting for me. I was kind of racking my brains originally. I was, like, what's the association between our fields? I thought originally it was going to be all these nice questions, and I truly would like you-all to think about those questions that I posed, try to come up with some answers, and get it back, because I think there's definitely some cross-collaboration that could go on between our research fields. I did not expect the body odor to be an association. I've used that slide about 50 times, got jokes every time, but I've never had somebody actually in the room that knew what I was talking about. (Laughter.) I thank you so much for the opportunity. I'll stick around for individual questions, but thank you again for the opportunity. (Applause.)

QATO : My name is Dima Qato. I'm a PharmD, MPH, and I'm working with the NSHAP project.

With the steady rise in medication use, it has been increasingly important to understand the determinants of drug use in the general population, however complicated it is, and while limited research has been done thus far in the role of medication use in the general population, it is an essential to the understanding of health behavior and practices and informing health and pharmaceutical policy and ultimately to improving health outcomes. Today our panelists will address pharmacoepidemiology, specifically on medication use as a key health behavior and complicated confounder.

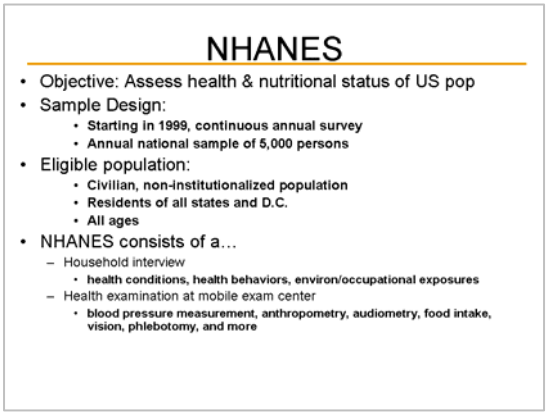
Our first panelist is Dr. Ryne Paulose. She is an epidemiologist in the Division of National Health and Nutrition Examination Surveys at the National Center for Health Statistics of the CDC. She is a graduate from Pennsylvania State University with a doctorate in bio-behavioral health and a minor in gerontology. She has been working with the National Center for Health Statistics for the last five years and is responsible for the prescription medication data.

PHARMACOEPIDEMIOLOGY: MEDICATION USE AS A KEY HEALTH BEHAVIOR AND COMPLICATED CONFOUNDER

Prescription Medication Data and the National Health and Nutrition Examination Survey (NHANES)

Ryne Paulose

PAULOSE: Good afternoon. Thank you, Stacy, for this invitation. I think it's a great opportunity for me to talk about the National Health and Nutrition Examination Survey, NHANES, and to more specifically discuss the prescription medication data that we collect.



Now, this is a very quick overview of a very large and complex survey, but the main objective of NHANES is to assess the health and nutritional status of U.S. children and adults. This survey in some form has been in existence since, I think, the early 1960s, but starting in 1999, the survey became a continuous annual survey. That means that we're out in the field every year collecting data on a national representative sample of approximately 5,000 persons. The persons who are eligible for this survey are residents of all states and of all ages, and we collect data only on the civilian, non-institutionalized populations. We don't sample institutionalized persons. Now, NHANES consists of a household interview which is conducted by trained interviewers using a computer-assisted personal interview. It's usually done at the home, at the participant's home. During the household interview, a number of questions on health conditions, health behaviors, environmental and occupational exposures and many more questions are obtained, and then participants who complete the interview and who also agree to participate are then eligible for the medical examination that occurs at our mobile exam centers, and at these centers we have trained health technicians, phlebotomists, and physicians who assess a variety of health measures that include blood pressure measurements, anthropometry, audiometry and much more.

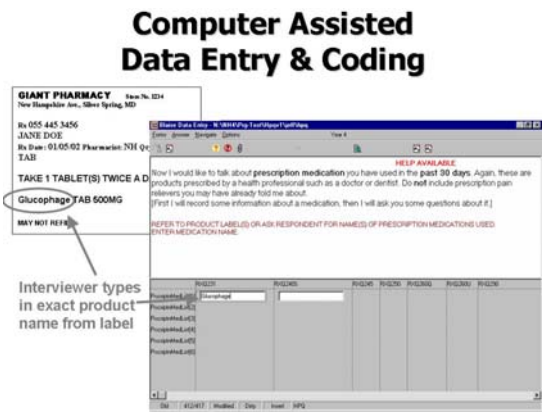
For this presentation I will be focusing on the prescription medication data that's collected on NHANES. Now, on surveys such as NHANES, drug data can be collected in two ways, primarily in two ways. They can be collected by directed recall, where you ask a specific question about a specific drug or about a drug class of interest. This can be through an indication-specific question that asks about medications taken for something like pain, for high blood pressure, for wheezing, or you can also ask a drug-specific question, where you ask about the specific drug of interest, such as aspirin or antihypertensive medications. NHANES currently asks both indication-specific and drug-specific questions.

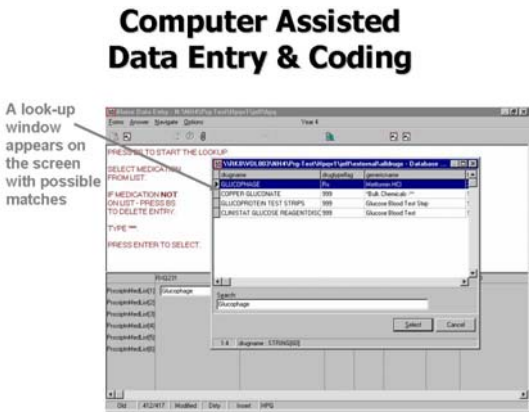
The other way to collect drug data is to conduct a medication inventory method. This is where participants are requested to show or to report to an interviewer all of the drugs that they have used during a specific time period, and then the interviewer would record the information about the drug directly from a drug container. Actually, this type of data collection can be done over the phone, and this is done by researchers at Boston University's Slone Epidemiology Unit. The survey they conduct is a phone interview, medication inventory by phone. But NHANES uses the in-home interview/medication inventory method. The specific questions that we ask during the medication interview part of NHANES is: "In the past 30 days have you used a medication for which a prescription was needed?" So we're asking about a one-month time period of medication use. "What were they?" The interviewer asks if they can please see the containers of all the medications used, and then we also ask for how long they had been using the product and what the main reason for use of the product was.

I want to identify several difficulties of collecting drug data on surveys. I think first, one of the difficulties is that a drug can have multiple names. There are multiple names for the same product. For example, it has a chemical name that describes the atomic or molecular structure of the drug. Then there's also the trade or brand name which is chosen by the pharmaceutical company that manufactures the drug. Then there's the generic name which is assigned by the U.S. Adopted Names Council, and this generally tends to be the official name of the drug. So as a consequence, with survey data collection, participants may report the generic name or the trade name of the drug. Additionally, there could be difficulties with recall where you get incorrect or imprecise information that's collected about the drug, and what we see is where the participant doesn't remember the specific drug they've taken and they tell us they've taken a heart medication or pain medication. Then sometimes on surveys the level of detail may also not be collected, and it's primarily due to time constraints and issues such as participant or interviewer burden.

Now, I wanted to just go over how the drugs can be reported. If you have information at this level where you know the drug is Prozac - it's a 20-milligram strength, 30 capsules total, and the manufacturer is Kaiser - you have enough information to uniquely identify that drug product and assign it to a unique 11-digit NDC code. Additionally with this information you can identify what ingredient this drug is made of, and in this case it would be fluoxetine HCL, and you can further assign a generic ingredient code. Additionally, you can further classify this into a drug class category and a code, and in this case fluoxetine would be identified as an antidepressant. Now, sometimes you only know that it's Prozac, 20 milligrams, and in this case you don't have enough information to assign it to an 11-digit NDC code, but you do have enough information to further identify that it's fluoxetine and that it's an antidepressant. We don't collect drug dose or strength. We only collect information such as Prozac or fluoxetine. We get just the name, brand name or generic name. In this case we still can identify that it's fluoxetine HCL and further classify it into a therapeutic class code, in this case antidepressant.

This sort of data entry and coding can be done manually where you can use existing pharmaceutical resources, such as Physicians' Desk Reference or Drug Facts and Comparisons, to assign a unique 11-digit NDC code. You can also assign a generic code using Mosby's GenRx, and then, you know, classify it manually based on different therapeutic classification systems that exist. But with large surveys and large numbers of participants, manual coding is inefficient. It can be costly and time-consuming. What is favored is to use a computer-assisted data collection and coding system. In this case we would be using a comprehensive drug database that would be uploaded to a computer that the interviewer could use in the field directly, and this is what NHANES does.





Here is an example of one of our data collection screens. During the household interview, the survey participants are asked to report any prescription medication they've used in the past 30 days, and they would show the interviewer a drug container that would hopefully have this level of information on it. What the interviewer would do for NHANES, since we're only collecting the name, they would take this name, Glucophage, and enter it into the computer. If they do not see the container and they can't transcribe the drug name, they would ask the participant to verbally report it and they would record whatever the participant verbally says into the computer screen. Based on what we've been collecting so far, our interviewers are seeing over 80 percent of the drug containers, so we're getting very reliable data. I know some of the other surveys get higher rates of containers seen, but 80 percent is still very good in terms of what we're collecting.

Then once the interviewer has entered the drug name into the computer, what happens is that automatically and transparently to the interviewer, their text entry gets matched against a prescription drug database to identify an exact drug match or similar text matches. In this case, they have entered Glucophage. It's based on a trigram match. A list of matches comes up. In this case, the first entry is an exact match, so the interviewer would select that. What we do is we retain what the interviewers have entered in addition to what they select for quality control purposes, and based on the first four years of data that I have reviewed, we're getting less than a 2-percent error rate in terms of what the interviewers are selecting. Often the reason for an incorrect match is because of drug extensions or drug strengths. New interviewers sometimes have issues where they're entering the drug form and we have to just train them to not enter the drug form because it's not in our database and they're not going to get an exact match. So far, with the last four years of data, we've been getting over a 92-percent match rate, meaning that all of the drugs that the interviewers are seeing, 92 percent or more of the time they are finding an exact match in our database. So it's very good.

Now, what we currently use for data collection: since 1999 through 2006, we have been using First DataBank's drug database. It's Master Drug Database, MDDB. We purchase it annually and we have a license to use their database. We modify it because the database that you license contains, I think it's like 12 files with hundreds of variables, and since we are only collecting a certain level of detail, we subset the file and upload drug names, generic names, therapeutic class codes onto the laptops that our interviewers use. We update the file annually. Although First DataBank provides updates to us quarterly, we only update the file that the interviewers have on their laptops annually, and that would allow us to update the file with any new products that have come into the U.S. drug market.

Before we release the data to the public, we go through several post data collection edits. The first thing we do is convert all of our drug names to their generic names. The reason is because we're not consistently collecting brand names. We don't ask the participant to remember whether it's a brand

name or the generic. We just ask them to report the drug that they have been using. So for consistency, before we release it to the public, we convert everything to the generic name. Here are examples: Prozac would be converted to fluoxetine HCL. Tylenol No. 3 would be converted to acetaminophen and codeine. Then, additionally, during the post-data-collection editing process, we remove drugs that have been incorrectly reported in this section, and what we see a lot is nonprescription drugs being reported because the participant sometimes can't distinguish what's an over-the-counter from what is a prescribed product. If the physician tells them they should be using aspirin, they will report aspirin as a prescription drug, although it only exists as an over-the-counter product. Also what we do is we assign the therapeutic classification codes to our public release files, and this is really to assist the analysts. We feel it's easier for our researchers to analyze our drug files based on certain drug classes of interest. If you're interested in antidepressant drugs, you can use the codes to identify all antidepressants and you yourself don't need to know what drugs fall into that category. So it's to assist the researchers.

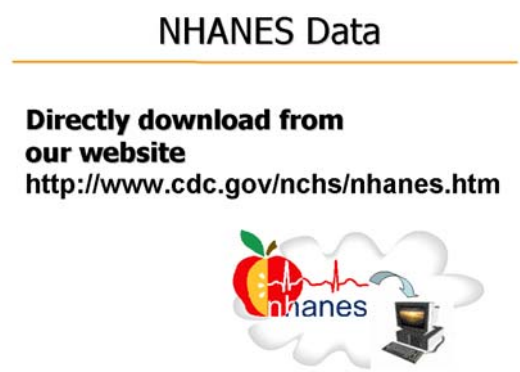
Now, one of the issues with this therapeutic class code is, from 1999 through 2004 we have been using FDA's National Drug Code Directory. They have a therapeutic classification system that they make available free to the public, and I think all of the NCHS surveys that collect drug data have been using this therapeutic classification system. But even though we use that, there are various other drug classification systems that currently exist that researchers can use, and I've just identified a few of these there. But each system classifies a drug into a specific category based on certain criteria. It can be indication for use, the chemical structure, or the drug's effect on the body systems.

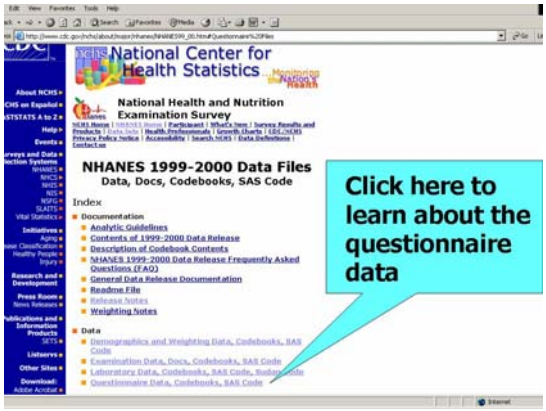
Now, here is an example of two different therapeutic classifications systems. One is FDA's classification system and this other one is AHFS. I'll just subset a couple just as an example for you here. According to the American Hospital Formulary Service, fluoxetine is classified under antidepressants, which is classified under psychotherapeutics, which is classified under a broader category of CNS agents. Now, if you look at FDA's therapeutic classification system, fluoxetine does fall under an antidepressant, but then just under an overall CNS category. So, similar to some extent. Now, if you're interested in, say, ibuprofen, it would be classified as a nonsteroidal anti-inflammatory drug under the American Hospital Formulary Service System, and therefore under the broader category of a CNS agent. But if you look, NSAIDs do not fall under the CNS category based on FDA's classification system, but, rather, they fall under the relief of pain category, and here we see NSAIDs and ibuprofen falling under there. So it's important to know these differences, I think, when you're analyzing the drug data, especially when you're looking for certain drugs of interest.

Now, since 2005, the system that we've been using, FDA's therapeutic classification system, has been withdrawn by FDA and it's no longer available on their website. For the last year, NCHS has been reviewing alternate drug databases that we can use and release to the public. One of the issues for us is that we can't always use a proprietary database because there are licensing issues regarding their therapeutic classification system. Right now, what NHANES has been using, I think we pay about $14,000 to $16,000 annually to use their drug database for our data collection process, but we are not allowed to release to the public the therapeutic classification system that's part of those files because it's proprietary. If we release it, all the drugs that our survey participants report annually, the public can essentially recreate this proprietary database. So they do not allow release of -- well, First DataBank does not allow release of - their therapeutic classification system.

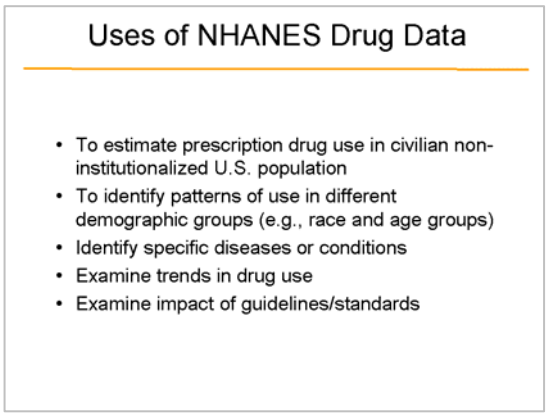
So during 2005, we had a lot of meetings, a lot of discussions of what was available out there. We went through several nonproprietary databases that we can obtain free of charge, but there were a lot of issues. They were not comprehensive enough. We found that the VA system that currently exists only captured about 50 to 60 percent of the drugs that were reported in NHANES, and that's just not acceptable to us, especially since with the First DataBank we've been capturing over 90 percent.

Then we looked into alternate proprietary databases, and what we found by accident in several discussions with different Multum/Cerner representatives is that Multum will allow the government to license and use their product free of charge. It was shocking to us, especially since we've been paying $14,000 to $16,000 annually for the last six years. And then, more importantly, the other thing was could we release their therapeutic classification system to the public with our public-use files, and they agreed. It's absolutely amazing. So what's going to be happening is that any of the drug data released from NCHS, the drug data will all have a common therapeutic classification system and we will all be using it for data collection, coding, editing, and public release.





So, currently what's available to the public: Currently we have NHANES III drug data that's from 1988 to 1994 and we also have 1999 through 2002 prescription drug data from NHANES that's available on our website for the public to use, and expected before the end of the year is the 2003-2004 data. Now, all of these data can be directly downloaded from our website. The data are released in two-year cycles, so if you go to our website, you can just access the specific years of data that you're interested in. And since prescription drug data is part of the questionnaire files, you would click on "questionnaire data." Here is a list of all the questionnaire files that currently have been released to the public. Prescription medications are here. We've released code book, documentation, raw frequencies, weighted frequencies. These will assist analysts just to get a sense of what sample sizes they're dealing with. And then the data are self-extracting Zip files, but they're SAS transport files, so you have to be familiar with SAS. I think we're making available to our data users some alternate ways of accessing and analyzing our data, but currently I think it's only through SAS. Now, I just wanted to mention that the files that we release are drug product level files; that is, the file contains multiple records per person. So it's very important to know that if you're going to try to compute prevalence estimates, you need to learn how to reduce the file to person-level to compute your estimates. These are just a list of the variables. I'm actually not going to go through that.



Lastly, what I wanted to show you is that there are many uses of the NHANES drug data. I've listed a few of these, but I think the major one is to estimate prescription drug use in the civilian, non-institutionalized population. Here is an example where much of our drug data has been released in the *Health, United States, 2004*. Actually the 2004 was a special edition on drugs, so there's a lot more drug data in this 2004 *Health, United States*, for anybody who is interested. Now, this figure

shows that about 40 percent of the total U.S. population is reporting a prescription drug, but, as you would expect, there are clear increases by age. And based on this figure, we can see that 18 or under 18, it's about a 20-percent prevalence of use, and it goes up to about 80 percent, actually more than 80 percent, as expected, in the 65-plus sample. The other thing I wanted to show you is that, in addition to computing just overall prevalence of prescription drug use, you can actually look at specific drug classes. Here is an example of psychotropic drugs where we looked at antidepressants, anxiolytics, and antipsychotics, and based on the NHANES III data, we've shown that about 5.5 percent of adults reported the use of a psychotropic drug. But when we looked at the 1999-2000 data, we found that this number had significantly increased to over 10 percent of adults, and the increase was primarily due to a threefold increase in antidepressant drug use. And this figure, published in the 2004 *Health, United States*, shows that this increase has occurred for men and women across all age groups. So the overall increase in psychotropic medication use is due to the antidepressant drug use. Additionally what you can do with our data is to look at the specific generic drugs. Based on the 1999-2000 data, it shows that the top three antidepressants that were used were the three SSRIs - fluoxetine, sertraline, and paroxetine. So there is a lot that you can do with our drug data. I think that's it. (Applause.)

QATO: Dr. Ryne Paulose is going to be taking questions now.

WALLACE: Thanks for that, and I use the dataset. First of all, I just want to say in the HRS that we actually do self-administered medication inventories, which are not quite as good but reasonably good, and it saves us a lot of time and money just to add that to the menu. My other question is, do you prompt people about medications that they otherwise might not think about, like laxatives or medications administered per rectum or eye medications or skin medications or things like that? Oral contraceptives are another that people forget. They remember to take them, but, you know . . .

PAULOSE: That's a good question. No, we don't currently prompt participants regarding what they have used in the past month, but the directed recall questions that exist as part of the earlier part of the questionnaire where they do ask about the specific drug classes, we do ask about oral contraceptive use. So during the medication inventory method, we do not. But at the beginning of the interview the participants are reminded by the interviewer that they will be asked about medications they have used in the past month, and what the participants usually do is they will collect all their drugs and leave it in a box on the counter ready for the interviewer. So we don't do any prompting. I think it's a good suggestion to do it, but we have not been able to do it, because this is the very end of the survey, it's a lengthy survey, and what the participants get is just a reminder at the beginning that we will be asking about the medications that they have used in the past 30 days.

HAUSER: I have two questions. Are there any social or economic differentials in response to the prescription drug items? Second, this is off topic, but just from a public health point of view, I wonder whether there is some possibility of going into the use of OTC and dietary supplements within NHANES.

PAULOSE: Let me answer the second question. As part of the medication inventory section of the questionnaire, we do ask participants about their use of prescription and nonprescription dietary supplements, vitamins and minerals. I didn't mention it here.

HAUSER: You mentioned the exclusion from the –

PAULOSE: Yes. So we do ask about that. We also do ask participants about nonprescription antacids that they have used in the past month as well, and then the prescription medications. We do not ask about all over-the-counter drugs. There was a period of time, 1999 through 2004, where we asked about nonprescription and prescription pain-relievers that they've used. So over-the-counter

medications would just be too much of a burden to try to collect at this stage for our survey participants and our interviewers, because there are no clear guidelines on labeling. The dietary supplement people have a huge task with what they're getting. Labels change every month. So currently there is no good way to collect over-the-counter drugs accurately with the way that we are doing the survey. So unless it's specific therapeutic classes, like pain-relievers, we don't collect all over-the-counter. Now, your first question?

HAUSER: My first question was about the social or economic differentials in response to the prescription drug items.

PAULOSE: That's a good question. We have not looked at the response rates in terms of -- well, the 80 percent of the time that we're seeing the containers, it's across the board. I mean our non-Hispanic whites, non-Hispanic Blacks, Mexican-Americans, they are all reporting about the 80 percent, but we have not looked at it in terms of data collection. It's a good point, one that I'll have to further investigate.

HAUSER: Thank you.

LINDAU: I thought this was very clear. It was a technical talk. For those of us who are collecting medication data or thinking about doing it, it's extremely helpful. And, as Dima mentioned, we have never had a session like this here. For people who don't want anything to do with medication data, I will say that I think there are so many potential uses of the medication data in the kind of research that many of us are doing, and whether we're looking at medication as a primary predictor of health or health outcomes or not I think is one use of it. But the other is that it's hard to imagine, especially in studies of aging, understanding the relationship between biological processes and social factors without at least controlling for medication use, especially when we're looking at things like blood pressure and inflammatory markers or markers of immune function. So I wanted to know if you can comment on any analyses that people in the NHANES group - or people working with the NHANES data - have been using that speak to the analytic issues.

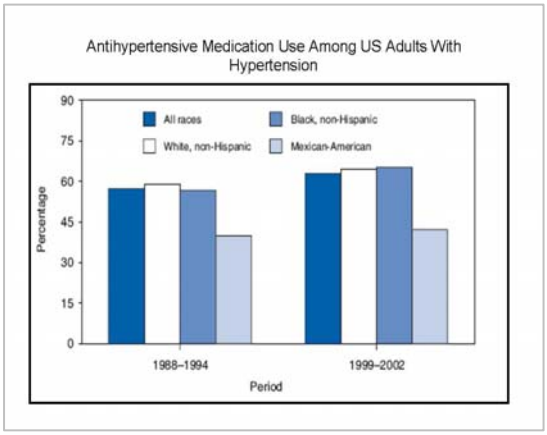
Antihypertensive Medication Use Among US Adults With Hypertension

Qiuping Gu, MD, PhD, MPH; Ryne Paulose-Ram, PhD; Charles Dillon, MD, PhD, MPH; Vicki Burt, ScM

Background—High blood pressure can be controlled through existing antihypertensive drug therapy. This study examined trends in prescribed antihypertensive medication use among US adults with hypertension and compared drug utilization patterns with recommendations of the Sixth Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure.

Methods and Results—Persons aged ≥18 years from the National Health and Nutrition Examination Surveys were identified as hypertensive on the basis of either a blood pressure ≥140/90 mm Hg or self-reported current treatment for hypertension with a prescription medication. In 1999–2002, 62.9% of US hypertensive adults took a prescription antihypertensive medication compared with 57.3% during 1988–1994 ($P<0.01$). Men had the greatest increase in antihypertensive medication use (47.5%, 1988–1994 versus 57.9%, 1999–2002 [$P<0.001$]). In both surveys, antihypertensive medication use increased with age, was lower among men than among women, and was lower among Mexican Americans than among non-Hispanic whites and blacks. Multiple antihypertensive drug use increased from 29.1% to 35.8% ($P<0.001$). Polytherapy with a calcium channel blocker, β-blocker, or angiotensin-converting enzyme inhibitor significantly increased by 30%, 42%, and 68%, respectively, whereas monotherapy with a diuretic or β-blocker significantly decreased. For hypertensives with diabetes, congestive heart failure, or a prior heart attack, the utilization patterns closely followed the Sixth Joint National Committee guideline recommendations.

Conclusions—Antihypertensive medication use and multiple antihypertensive medication use among US hypertensive adults increased over the past 10 years, but disparities by sociodemographic factors continue to exist. (*Circulation*. 2006;113:213-221.)



PAULOSE: Yes. There's actually a couple of other slides. Especially several, like antihypertensive medication use, identifying people with hypertension, we do have our blood pressure data that's collected, but in addition to that, we are looking at the drug data because it allows you to capture some people that might not have -- well, obviously the people that are on antihypertensive medications that have blood pressure under control, you're missing that group of hypertensives. So the way that NHANES drug data gets used, it's to complement the examination data and the self-report data, and it's clearly a wealth of data that I think adds to the information that you collect on the participants. And we definitely deal with looking at hypertension, and people with high cholesterol, we'll look at their drug data. Those are the two that have been published so far. We

look at people who are self-reporting depression and stuff and then we look at whether they're getting treatment. So there's a lot of uses of the drug data that complement the rest of the data that we're collecting.

HALL: I have two questions about the drug data collection. It seems like there would be a lot of rationale for collecting information on OTC medications for indications that also have prescription forms, such as antihistamine allergy medications, sleep medications. It sounds like you are collecting or at least have collected in some samples in the past over-the-counter antacids. Could you speak a little more as to whether there are classes of OTC's that you are collecting?

Dima presented a great review of some of these issues to the NSHAP project about a month ago. It looks like none of the major classification schemes that are being used break down with an indication by pharmacological mechanism, and while the numbers on some of these samples might be too small to get the data for aggregate meta-analyses or for long-term repeated sampling in larger projects, it would seem that there would be a lot of reasons where people might want to distinguish between long-term effects of SSRI versus SNRI versus TCA antidepressants, for instance, versus other (inaudible) antidepressants and so on, and a lot of the other medication categories as well. I don't know whether these classification systems distinguish among the subcategories of nonsteroidal anti-inflammatories or NSAIDs versus COX-2 inhibitors and so on. A lot of that information could be very important to epidemiologists, and certainly the drug companies have no motivation to collect that data.

PAULOSE: Regarding the OTC drugs, I don't think we're going to be adding data collection on all over-the-counter products unless something changes with labeling and stuff.

HALL: Right. I mean my concern was more trying to do analyses on categories that have both prescription and nonprescription category members that could lead to very complicated sets of data if you're not collecting the ones that aren't prescription.

PAULOSE: Right. That's an excellent point. I mean, it's true. What you're able to do is control for some of the factors. You're not going to be able to control for all drugs that the participant may be using and may be affecting their health. So at least controlling for prescription drugs that they have used in the last month is better than not looking at anything at all. But clearly looking at over-the-counter drugs that have effects like those that you mentioned would be very important depending on the health outcome you're interested in. Unfortunately, one of the issues for NHANES is that a lot of the stuff that gets on our survey is directly funded by certain agencies within NIH that have specific interests, and without the funding it's really hard to add a whole section on over-the-counter drugs. The reason that the nonprescription and prescription pain-reliever data was collected was because NIDDK was interested in it and they proposed adding it, and they funded it. So unfortunately the funding stopped and it's off our survey now. But we do have six years of pain-reliever - over-the-counter and prescription - drug data, and it would be great. So if you know of people who have money, tell them to come talk to us. Regarding the therapeutic classification system, there is no perfect therapeutic classification system that's out there that will answer every researcher's questions, but what you can do is take at least what we've given and you further classify it knowing specific drugs you're interested in. The psychotropic drug paper that I published, I had specific ways of classifying the antidepressants. I actually did a crude look at the drugs based on the general classifications that were in the public release file, and then I further reviewed with a couple of physicians that I work with on further classifying it according to the research question I was interested in. So what we do is just provide something at the minimum for researchers to start with, but there is no perfect system out there.

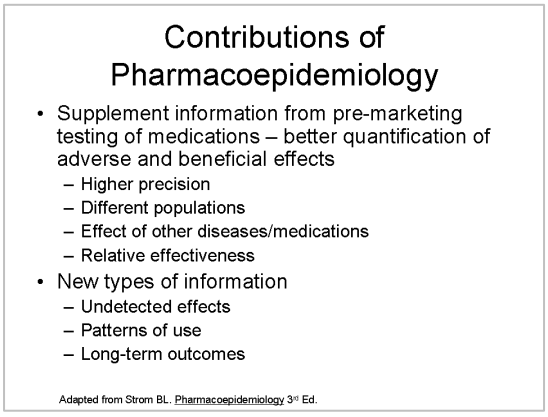
QATO: Thank you very much, Dr. Paulose, for the very informative presentation. (Applause.)

Our next panelist is Professor Todd Lee, who is currently working at Northwestern University and is an adjunct professor at UIC College of Pharmacology. He is a senior investigator for the Midwest Center for Health Services and Policy Research. He earned his PhD in Pharmaceutical Outcomes Research and Policy from the University of Washington and has previously worked as a managed-care fellow with Regence Blue Shield and Premera Blue Cross in Washington.

Pharmacoepidemiology: Medication use as Key Health Behavior and Complicated Confounder

Todd A. Lee

LEE: I have a couple of tangential comments before I get started. NHANES, the field-based interview semis are parked in our parking lot at Hines right now. So they're doing the Chicago area field-based interviews. The second is, I just came back from a trip visiting some in-laws in Iowa and they live on a farm, and when my 3-year-old asked to go out to see the pigs, I'm certainly glad I didn't have any Passion on, because who knows what would have happened. (Laughter.) My talk is going to be a bit different than things you've heard so far today, and I feel like I'm a bit of a fish out of water here. I'm going to give you a walk through my world of outcomes research, specifically how it relates to medications, and try and give you three examples of some issues that we deal with when we're doing pharmacoepidemiology and related studies.



So what is pharmacoepidemiology? It's the study of the use of and effects of drugs in large numbers of people. As you would expect, it combines the fields of clinical pharmacology - how drugs work in humans - and epidemiology population-based studies. How can pharmacoepidemiology inform us? Well, it provides information really in specific sorts of ways. It can supplement information from pre-marketing testing of medications, gives us a better idea of effects, both beneficial and adverse, with respect to medications. We can get higher rates of precision of actual drug effects. We get estimates of sort of real-world effects versus very tightly-controlled effects in clinical trial settings. We get information on different populations, populations that aren't typically included in clinical trials. We can look at the effects of other diseases and medications and how they interact and modify the effects of medications, and then, importantly, we can look at the relative effectiveness of drugs within a class. Also it provides us new types of information. It provides us information on some undetected side effects; we can get information on patterns of use; and, importantly, some long-term outcomes.

So I'm going to focus on three of these things. I'm going to talk about effect of other diseases and medications, on specific medications of interest. I'm going to talk about relative effectiveness, and then I'm going to talk about undetected effects. The title that's given is "Medication Use." Is it a key health behavior and a complicated confounder? Well, how does it fit into each of these two categories? It is definitely a key health behavior. It's a sign of disease. It's a sign of problems. It can be associated with harmful and beneficial outcomes that are important to patients and providers. As was noted earlier, it's an incredibly common treatment modality. Is it a complicated confounder? Yes. It's sometimes very difficult to tease apart the effect of a medication from the other effects

when we're doing population-based research. Trying to determine the effects of multiple medications in multiple patients is sometimes complex. And, importantly, is the outcome associated with the disease process or the medication? I'm battling a sinus infection right now, and is my drowsiness and my lethargy an effect of the restless nights I've been having because of the sleep problems or is it related to the drugs? So that's what we're left to sort of tease apart when we start to look at these medications in patients.

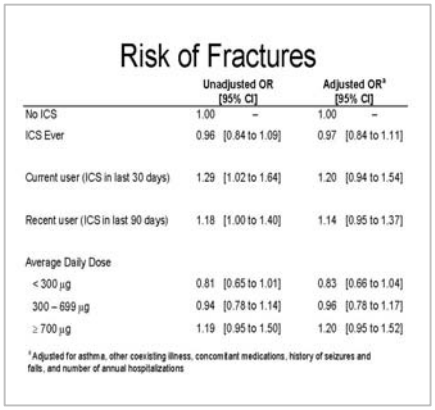
Where do we get our prescription data? Well, NHANES uses self-report and interviews, through direct observation and patient recall. There are other ways to get medication data, and I'm a lot more familiar with these other ways than I am with patient self-report, although patient self-report is incredibly important in a couple of projects I've been involved with, including one where we're collecting drug information from patients throughout the world, 15 different countries, and, as you can imagine, health care systems and medications that are used in all these different countries are incredibly diverse, so we have to rely on patient recall and direct observation of medications that are used. But what I'm mostly familiar with is extracting information from medical charts and records, and the EMR actually makes this an incredibly viable way to collect data for large patient populations, and through dispensing records. When a transaction occurs at a pharmacy, there's a record of that transaction. It's contained in a database and maintained so we can access that information in order to look at prescription drug use.

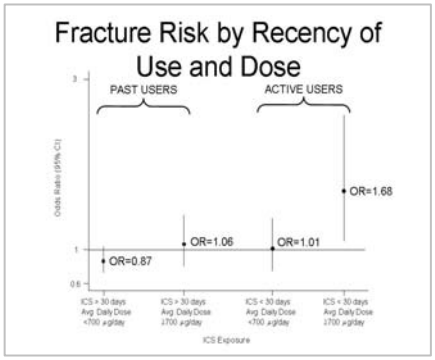
So what are the three key issues that I'm going to walk through in the next several slides? Well, I'm going to talk about how to define exposure. That's usually critically important. The fact that treatment assignment is not random. People get on these drugs for reasons and we have to really consider those reasons when we're trying to make inferences about effects that we observed. And then what about effect modification and how does it play a role with different diseases and different medications.

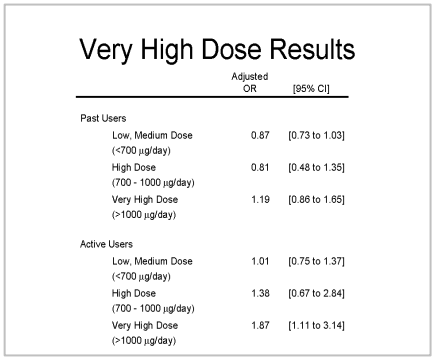
As a full disclosure moment, the studies I'm going to talk about have been funded by GlaxoSmithKline, AHRQ, and the VA. The first example is looking at the risk of fractures in people with using inhaled corticosteroids. Inhaled corticosteroids are a common treatment for patients with respiratory disease. In this case we looked at patients with chronic obstructive pulmonary disease in a VA population to see if they were associated with an increased risk of nonvertebral fractures. We conducted a nested case control study in a cohort of VA patients with COPD. Who were these patients? Well, they were patients with a diagnosis, not incident cases but prevalent cases of COPD. They had been newly treated with COPD medications and they were fracture-free during at least a three-month follow-up period. Our cases were people who experienced a fracture during a three-year follow-up period, and we excluded fractures that occurred within 14 days of a motor vehicle accident. Our controls were all patients from a COPD cohort that didn't have a fracture. We individually matched them on COPD diagnosis, age category, and gender.

Now, exposure. Exposure is incredibly important here. So how do we identify and define inhaled corticosteroid exposure? Well, we grab all of the inhaled corticosteroid prescriptions that happen between when the patients start in the study, their COPD diagnosis date, and prescription date in this case, and when they stop, the end of follow-up or the fracture date. We have to convert it to a similar metric, because there are different potencies with respect to inhaled corticosteroids. And then how do we classify exposure? This is going to make a difference in our inferences. Do we simply look at it as ever/never exposed? Do we look at the recency of exposure, the cumulative dose, or the average daily dose? All of these bits of information provide us different pieces of the puzzle in trying to tease apart the relationship between exposure and event. Pharmacy data is not always easy to deal with, and it's particularly true when you're talking about inhaled products. We're a little bit different than NHANES when we're looking at prescription filling records, that we don't just have a, "Yes, I used that drug." Now we've got lots of details on when drugs were filled, the strength

that they were filled, the duration for which they have been filled. So when we're dealing with inhaled products, there's sometimes some ambiguity around prescribing guidelines, especially with the short-acting beta-agonists and how many inhalations are contained in a single inhaler. It's a much more straightforward question when we're talking about capsules and tablets. Those are much easier to count than actuations in an inhaler.



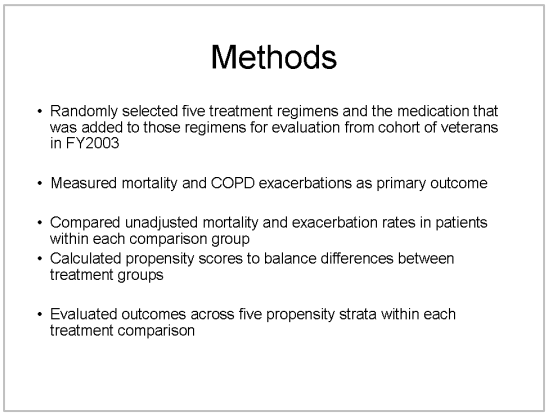




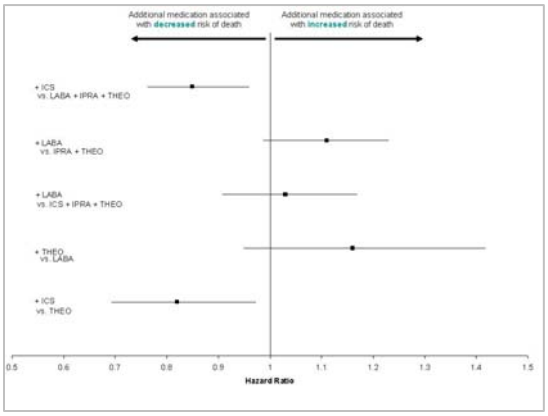
This is just an example of some raw data that we have to deal with. The left-hand column is an identifier, then you've got a prescription dispensing date, you've got the product, you've got the directions for use, and then finally the day supply and the total quantity. Now, this is just an example, and we're talking about 350,000 observations here for a little more than 30,000 people. So we have lots of information. How do we get this into a usable form? We extract bits of information from those variables. We used the VA product variable, for example, to determine the specific product that was used, the dose and the strength that was used, and we used that to then assign a number of actuations to each of the products. The directions variable is used to determine the dosing frequency and the number of doses per day. Now, we can combine this information and get some really meaningful exposure information or we can just simply look at ever/never exposure where we talk about whether or not an individual is exposed to an inhaled corticosteroid product at all based on the presence or absence of that product during the follow-up period. We can also look at cumulative exposure. Now, here is an example where you've got fluticasone, 220 micrograms per spray, a 13-gram container, and the directions are take two puffs twice daily. Now, we use that information - fluticasone, 220 micrograms - convert it to a standard equivalent, in this case beclomethasone equivalents, so we essentially halve the strength of fluticasone, and then take that times the quantity dispensed, number of doses per day, do this for each record, for each individual, and we can get a sense of actual cumulative exposure over the time period. So what do we find? Well, when we're comparing ICS use, inhaled corticosteroid use, to no use, the unadjusted results are on the left and the adjusted results are on the right. When we just look at ever/never use, there's really no association between fractures and the use of inhaled corticosteroid, but when we look at current users, we start to see sort of a signal sort of develop here. We've got a 1.3 increase in risk with respect to inhaled corticosteroid users versus never users in the unadjusted fashion. When we adjust that, the risk is no longer statistically significant, but it's still a slightly elevated risk. When we look at people who have stopped using, they've only used it in the last 90 days, not in the last 30 days, we see that the effect is sort of muted. But now, when we start to look at the dose-response relationship, we can see a clearer picture beginning to develop. From that previous table, we're motivated to look at not only the dose but also the interaction between dose and recency. So on the left we've got past users, and in those past users we've got low dose and moderate to high dose, and then we've got active users, people that were filling prescriptions at the

time of their fracture, and low dose and relatively high dose, and now is when we start to see sort of an interaction between frequency of use and the dose of the medication.

So how about if we further look at that dose response. Well, when we break it into very high users and active users, we really start to see the effect. So the effect is being driven by people who are using relatively high doses of medication and who are actively using the medication. So that's great. We see this association. Well, what's important to know is whether we can draw some inferences from this association and is this relationship causal. Well, I would suggest that, yes, we found that there is a temporal relationship, which gives credence to the conclusion that it may be causal. We saw that it's related to the active use of the medication. The strength of the association, our odds ratio isn't that impressive, so I give that not such an impressive sort of number. We definitely found a dose-response. How about the replication of the findings? Well, that's sort of a plus/minus. Some people have found what we did. Others have found no association. What people have not done is really done this in the age of people that we've done it in. These people were 65, 70, 80 years old. So there may be an age-related phenomenon going on. Is there biologic plausibility? Absolutely. We know that all steroids are associated with decreases in bone mineral density. There have been some associations with inhaled steroids and decreases in bone mineral density. So there is definitely some biological plausibility behind the relationship. And then when people stopped using the drug, the effect sort of went away. People who were past users didn't have the same sort of effect as people who were current and active users. So we sort of have a trend toward believing that this relationship may be causal.



Now, switching gears here, confounding by indication is a huge issue in pharmacoepidemiology. We completed a recent study funded by AHRQ where we saw that theophylline was always associated with worse outcomes in people with COPD. Now, the question that always came up was whether or not we had adequately controlled for disease severity. Was simply getting theophylline an indicator of worse disease and the worse disease was associated with the worse outcomes? We sought to randomly select five regimens in COPD and see if the addition of a medication was always related to worse outcomes. So we measure mortality and COPD exacerbations in these patients. We look at unadjusted mortality and exacerbation rates in each group. We calculate propensity scores to try and get away from some of the sort of differences between the treated group and the untreated group. We then evaluate the outcomes across the populations.



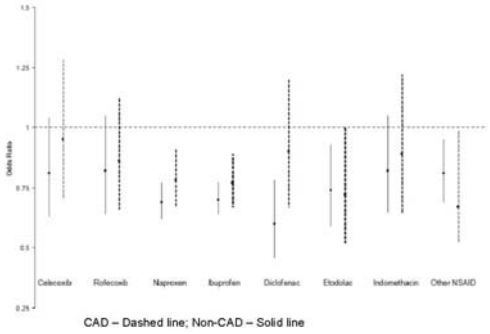
Here is a really busy slide, but I'm going to point you to two things. In each of these groups - Group 1, Group 2, and Group 3 - the column on the right contains the medication that was added to the regimen. So we're comparing theophylline use to theophylline plus inhaled corticosteroid use in Group 1. Well, what you see across the board is that there is an increase in the likelihood of visiting the pulmonologist when you add a drug. Makes sense. And pulmonologists are typically treating more severe COPD patients. So we're definitely worried about confounding by indication here. The same thing with the exacerbations; more exacerbations in the people that have an extra drug. This is Group 4 and Group 5 that we selected. It's the same thing here: more pulmonary visits, more visits to a pulmonologist when an extra drug is added to a regimen and more COPD-related exacerbations during the baseline period. So we've got really good indication here that we've got some confounding by indication going on. People who get the extra drug in their regimen are more likely to be those that have more severe COPD. When we look at the unadjusted outcomes, it sort of plays itself out. Mortality on the left, exacerbations on the right, the risk in the group with an additional medication relative to the group without the additional medication. With respect to mortality, we see no association for lots of the groups where the confidence interval crosses 1. Exacerbations, the group with an extra medication is always related to having a higher risk of exacerbations during the follow-up period. So these are unadjusted. So when we use propensity scores to try and get similar groups and then look at mortality risk, we see that it is actually a drug effect. There is a drug effect here, and it's not simply being that when you get an additional drug in your regimen relative to people without that drug you always do worse. In this case we see those that had inhaled corticosteroids added to their regimens actually have a decreased risk of mortality during the follow-up period, whereas people who had a long-acting beta-agonist added or they had theophylline added have no association, but the trend is toward them having worse outcomes. And this sort of 15-percent reduction in mortality is consistent with meta-analyses and clinical trials on the association between steroids and its effect in people with COPD. So it's not a design issue, but it's something that you always have to keep in the back of your mind as to how are we controlling for confounding by indication, how can we make sure that we have adequately controlled for disease severity.

So from this study we see that the addition of a medication is not always associated with worse outcomes, which is what we were really concerned about when we saw the theophylline results. The results depend on the medication that was added. Medication additions may or may not indicate severity of disease, but more than likely they do. And the results are not always going to be unfavorable to the regimen with an additional medication, at least in this case with COPD and respiratory medications.

Risk of Events

- Risk of cardiovascular or cerebrovascular event in those exposed to NSAID
  - CAD = 1.18 [1.11 to 1.27]
  - Non-CAD = 1.14 [1.08 to 1.21]
- Amount of COX-2 selectivity matters
  - Rofecoxib results
    - CAD = 1.01 [0.78 to 1.31]
    - Non-CAD = 1.32 [1.04 to 1.67]

Risk of Mortality



CAD – Dashed line; Non-CAD – Solid line

Now, finally, I'm going to talk about effect modification and effect modification between disease and the risk of, here, nonsteroidals on cardiovascular and cerebrovascular events and ultimately mortality. So we studied nonsteroidals and people with and without preexisting coronary artery disease to look at the event risk. When we look at the risk of cerebrovascular or cardiovascular events in those exposed to nonsteroidals, it's relatively consistent between those with coronary artery disease and those without coronary artery disease; there is an increased risk in having the events. It is really dependent on the amount of COX-2 selectivity in that those people exposed to rofecoxib - Vioxx in this case: If you had coronary artery disease, you really weren't at an increased risk of having cerebrovascular or cardiovascular events in the follow-up period, but if you were at a very low cardiovascular risk, had no prior history, then this drug sort of pushes you over the edge. So there is some effect modification going on here by the presence or absence of cardiovascular disease in people who are exposed to the COX-2's. When we looked at mortality, the difference isn't quite so striking. Here the coronary artery disease group is the dash line and the noncoronary artery disease group is the solid line, and this is the risk of mortality during follow-up. The values that cross 1 are not statistically significantly associated with a decreased risk of mortality, but all of the point estimates show a decreased risk of mortality with a nonsteroidal, something that really hadn't been looked at in detail with these drugs. Most people had focused on the risk of heart attacks and strokes. Well, when you look at all-cause mortality, some of these drugs may actually be protective, and naproxen and ibuprofen are the two that really jump out at you. The difference actually between the cardiovascular disease group and the noncardiovascular disease group sort of bears itself out and is an important difference in whether or not people are deciding to take these agents. This paper is actually going to be published in the fall in the American Journal of Medicine, but it has led us to some additional work, which is to better characterize coronary artery disease risk and to look at other potential risk modifiers in people who are exposed to nonsteroidals. So we are actually drawing information on lipid profiles and liver function tests and renal function tests to get a better grasp of how this effect modification may be playing itself out with respect to exposure of these drugs.

This is the last slide. Medication use, is it a key health behavior? I would say yes, it is a key health behavior. Is it a complicated confounder? Absolutely it is a complicated confounder. The key is to try and tease these two things apart from each other and get some informative information out of here, but ultimately I think it's an interesting and important covariate in all population-based research that we do. With that, I would be happy to answer any questions or take any comments. (Applause.)

QATO: Any questions?

McCLINTOCK: That was really informative because I'm new to this domain. My question is, given the project that we're working on, if we have a constellation of drugs that are being used, can we use them to infer what was the second item there, mainly diagnosis of what it is? Someone could be taking an inhaler and that could be because of asthma; it could be because of GERD, which is

arguably stress or anatomical muscle relaxation, which is very different than asthma, which is, you know, inflammatory or whatever; or it could be something like, you know, atypical Mycobacterium intracellulare, which is a parasite.

LEE: Sometimes. The specificity of medications for diseases varies. Sometimes we know what drugs are used for, and some drugs have very good specificity with respect to that, but lots of times we don't. The drugs for hypertension, for example, ACE inhibitors, we don't know if it's actually for hypertension. We don't know if it's for chronic kidney disease. We don't know what it's for. So inferring disease from drug exposure isn't always a great idea when there is a lack of specificity. There are classes of drugs for which you can make a pretty strong argument that that is what the condition is that is being treated. So other than sometimes, there is no real good answer to your question.

McCLINTOCK: Well, do you have a list of ones we can do and ones that we can't?

LEE: There are lots of algorithms that point you to drugs for identifying diseases, and some are better than others. I'd be happy to sort of point you to those.

McCLINTOCK: Great. Thank you.

JORDAN: Can you use this to ask the question about women who may be pregnant in a study that you've captured and whether you can look at outcome in pregnancy?

LEE: That's an RFP right now. AHRQ just put out an RFP to look at that issue. So there are people who are moving toward using this information to look at the teratogenic effects of drugs within pregnant women. Especially, I think, the focus is really on ACE inhibitors right now, because there's some evidence in the literature that would suggest that they may not be a good idea to take when you're pregnant and does that hold true for ARB's. So yes, these techniques allow you to look at those issues. As long as there's enough of a population to look at and that they're using these medications in the natural setting, then you can answer questions like that.

LINDAU: Thanks. I thought your talk was excellent, high energy, and very impressive. What you presented were analyses where you're trying to look at the effect of the medication.

LEE: Right.

LINDAU: When I look at the variables for which you adjusted, in a population study, looking at health, say, and social factors, your adjusted variables might be the key variables of interest in our analyses, and then we need to control for the medication.

LEE: Exactly.

LINDAU: So what's the take-away message? Say we're looking at a relationship between blood pressure and social connectedness, and we want to control for whether or not the individual is using antihypertensives. What I'm hearing from you is that it's not enough to just control for whether or not they're taking antihypertensives, or is it, and can you comment on that?

LEE: Maybe. First I would argue that you wouldn't need to control for antihypertensive use to look at that relationship, because antihypertensives are known to lower blood pressure, so if you ignore that effect, then you may get a spurious relationship. So I'd say that looking at medication use is a critical covariate in your analysis. How you look at it, it may depend on the question of interest. Are you interested in sort of the effect modification, the co-relationship between medication use and your

covariate of interest, or do you just want to control for it, and maybe it's simply enough to include it in your regression and look at sort of a dose-response within your regression. It doesn't have to be included as -- you're right. My focus is using this as "the" covariate that I look at, but it doesn't have to be that way. It can be like any other regressor in any other analysis that you do, and I would contend that it's an important one in most analyses.

LINDAU: So then a follow-up to that. Obviously a complication is that many individuals who are taking an antihypertensive are taking more than one antihypertensive, or they're taking an antihypertensive for treatment or prevention of, say, diabetes complications and they don't actually have a hypertension problem.

LEE: Yes, yes.

LINDAU: Any advice on how to deal with those things? For example, in our study we collect measures that would be compatible with the Charlson comorbidity score. We know whether the person thinks they have or have been told that they have high blood pressure. So you start with some indicator aside from the medication of what diseases the individual thinks they have, or how do you go about that?

LEE: There are risk indices, like the Charlson, for example, that use medication data. I think they call it Rx-Risk now. It used to be the Chronic Disease Score. But it uses that information to sort of control for these other conditions that an individual has. I would not call them complementary to each other. They sort of would be supplements for each other or substitutes for each other. Being a medication guy, I like that information. I like the information you can get from drug-use data. And using a risk adjuster, like the Chronic Disease Score, I think provides a valuable tool when you're trying to tease apart the actual relationship between your covariates of interest and your outcomes. So I would encourage you to look at those. When you do have the ever/never sort of, the drug is present or it's not, you can use that information. What you can't get at is dose-response and multiple medications. But you can get at the ever/never issue, and ever/never is a lot when we're talking about trying to tease apart drug effects because lots of drug effects are taken out of the equation if you look at ever/never.

LINDAU: Thanks.

QATO: Thank you both for very excellent presentations. (Applause.)

GAVRILOVA: It just happened that this is the second time that I'm talking at the last session of this Biomarkers Workshop. The last time I was talking about the CCBAR website, a topic which I considered to be a little bit boring and it was very hard to keep the audience attentive. Today I'm in a much better position because today we have a very interesting topic: innovative methods in biomarker collection. We have a very impressive team of panelists which I would like to introduce.

INNOVATIVE METHODS

Steps Toward Circadian Psychophysiology

Frank Wilhelm

WILHELM: Thank you. It's a pleasure to be invited to this conference.

My talk will be on psychophysiological measures and how they could be applied outside of a laboratory. This is the kind of paradigm we have been using in NIH-funded research at Stanford University and Harvard where we studied anxiety and anxiety disorders and we wanted to know what kinds of reactions people have to anxiogenic stimuli.

In general, there are three systems that we as scientists can study in humans: Self-report, people say, "I am afraid"; on rating scales, the physiology, like heart racing; and the behavior, people avoid certain situations or run away from them. The area of specific interest, of course, is the physiology domain.

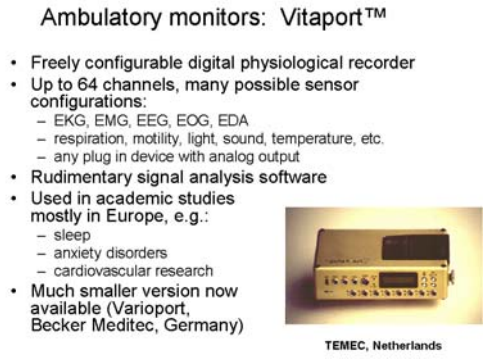
First is the HPA axis, measuring salivary cortisol. I will not talk much about that here. I will more focus on the sympathetic-adrenal-medullary system, which, with many cardiovascular measures we have explored in the laboratory and outside of it - like heart rate, blood pressure, pre-ejection period, pulse transit time, pulse wave amplitudes, T-wave amplitudes and so on, electrodermal activity, another sympathetic measure through different pathways, and stress hormones, like epinephrine or norepinephrine - I will not talk about here. Then an important system, the vagal system, that has not been much explored. Steve Porges, of course, has been one of the primary researchers in that field, and there are only a few cardiovascular measures really that tap into this respiratory sinus arrhythmia phenomenon. And last, the respiratory system, which has proven to be very important in certain anxiety disorders, looking at volume, timing parameters, variability, and that means respiratory regulation instability, and pCO2 as a measure of hyperventilation.

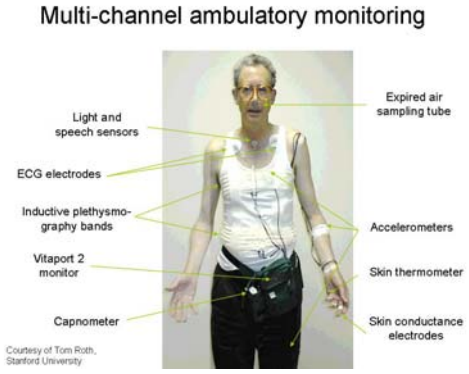
First, I'll show you what we do in the laboratories in trying to simulate real-life anxiety situations. For example, with virtual reality, there is a scenario of a freestanding tower, an open elevator platform subjects with fear of heights stand on, and then under control of the experimenter, they get higher and higher, up to 60 meters. They can look around. This is what the lab scenario looks like. There's the VR head-mounted display, and the sound, of course, matches their experience. When they look around, they can look down and they see what's underneath the elevator. Of course, they get anxious. This is the group selected for high anxiety, a student population, and when they drive up the elevator, they get very highly anxious, 7 on a scale from zero to 10, and when they drive down, it's reduced. The control participants don't get very anxious. Some physiological measures match that variable, like here, electrodermal activity -- I'm sorry, this is in German -- and the non-

anxious group doesn't show much response. This is a sophisticated attempt to bring real life into the laboratory, but what is also a fruitful approach is to bring the laboratory to real life.

Here is an example where Paul Grossman, a frequent collaborator, who is now also at the University of Basel, did a study with several laboratory mental stress tests in the morning. You can see here the repeated heart rate increases. The average of the daily heart rate recording in this record is set to zero, so you see increases of maybe 8 beats per minute during these mild to moderate mental stressors, like speaking in front of an audience, and you see that those repeated stressors habituate quickly. We only plotted data here of minutes when the subject was not physically active. Nevertheless, and surprisingly, when this person left the lab and engaged in normal life, we saw higher heart rates than during the stressors in the lab. And most of all, when the subject engaged in a really exciting task in the evening - watching a World Cup soccer game - you can see incredible heart rate increases. This was not due to jumping up and down, because these data segments were removed. Now, the point I want to make is that we don't actually know much when we study people only in the laboratory. This concept is called "ecological validity."

So we were in search of ecologically valid biomarkers of anxiety and stress and specific markers for certain anxiety disorders, and we utilized a variety of ambulatory recording systems over the past 15 years. Activity watches, everyone knows those; heart rate watches. Vitaport is a very good system, in three model generations now. The University of Amsterdam has a good system, the VU-AMS. Now we are using the LifeShirt system in Switzerland and also had a number of customized sensor devices. For this approach we coined the term "Circadian psychophysiology."





This is the Vitaport device now produced in The Netherlands. I'll show you a few examples of those. I'm sure most people are not familiar with the leaps technology has done in the past ten years. This is freely configurable. That means any kind of sensor can be stuck into those sockets, and the software can be configured to have any amplification or filtering, and even these modules can be stacked up to eight times, so you get 64 channels. We've had two stacks of those in an ambulatory study, but we have also been using this in the laboratory because it's actually a very good system for all of these typical measures people measure in the laboratory. It does not have such a good signal analysis software, so I was actually writing a lot of the software for the analysis myself. This has been used mostly in Europe in sleep, anxiety disorders and cardiovascular research. Now there is a much smaller version. If people look on the Internet, it's called Varioport. It's produced in Germany by the developer also of that system, Klaus Becker.

This is the kind of enthusiasm people can have for this kind of research. You see Professor Walton T. Roth, whom I have worked with at Stanford University, a psychiatrist, and you see the kinds of sensors we used in a clinical study.

In a study I did - my PhD thesis - we put people on airplanes. People who are phobic are flying as part of a repeated exposure treatment. We monitored the progress of exposure and also the acute anxiety people experienced in an airplane. You see here that the self-reported anxiety goes up. The heart rate goes up a lot as well, from 80 to 105 beats per minute. That's much more than people typically see in laboratories with any kind of stressor. We have had phobic participants who had heart rates up to 180 beats per minute during flight. That tells you more than the self-report, clearly. Of course, the RSA, this is respiratory sinus arrhythmia, a parasympathetic measure, goes down in the phobic group.

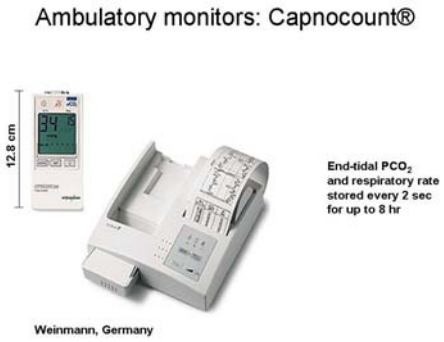
Here you see an individual record of a pre-flight baseline in the clinic, and then the incredible respiratory instability in minute ventilation, which is how much air you breathe in a minute, during the flight. We think this indicates acute bouts of anxiety and attempts to relax. And after the flight in the clinic the post-baseline looks very stable.

You can do with psychophysiological measures what people do with the SCL-90, for example. This kind of psychophysiological profiling hasn't been done much, but would most likely be a fruitful approach to look at profiles or aggregate scores rather than at individual measures. With this you get a profile for each individual patient in their reactivity to the flight situation on all these kinds of items. The upper items are self-reported diary items, and the lower ones are measured physiology. You see how people deviate from the normal score of 50. These are t-scores, so if a person is more than one or two standard deviations above or below -- that is, 10 or 20 points -- it indicates a reliable abnormality.

Now, with this system we also examined the physiology of patients with a diagnosis of panic disorder. This is a record of 30 minutes of breathing while sitting quietly. The figures depict the depth of breathing from breath to breath over 30 minutes in a normal, control person, and then for a typical panic patient. You can see the kind of outliers that are produced in this patient by the respiratory system. So the regulation is quite chaotic and unstable, which relates to a prominent symptom in this anxiety disorder, namely shortness of breath.

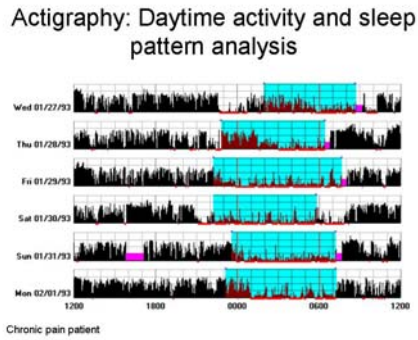
This is a study of exposure treatment of driving phobia that Georg Alpers, a PhD student in our lab, did. We examined how hyperventilation might improve with repeated exposure. Initially there is a strong tendency of the driving phobics to have a drop in pCO2 in the expired air, and as they go back to the lab, it normalizes compared to controls. You can also see that they improve from the first to the second drive considerably.

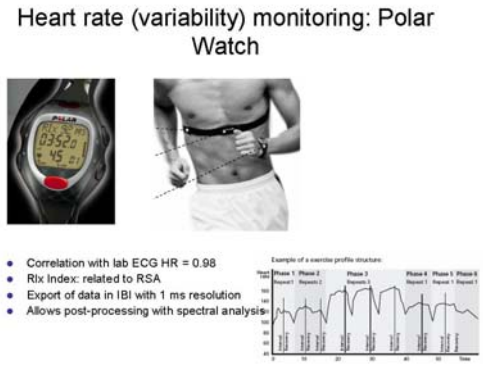
This is the device we've been using to do that. One can download the pCO2 and respiratory frequency data to a PC so it can be used as a stand-alone ambulatory measurement device.



Many people are familiar with these kinds of devices, the Actiwatch or the MicroMini Motionlogger.

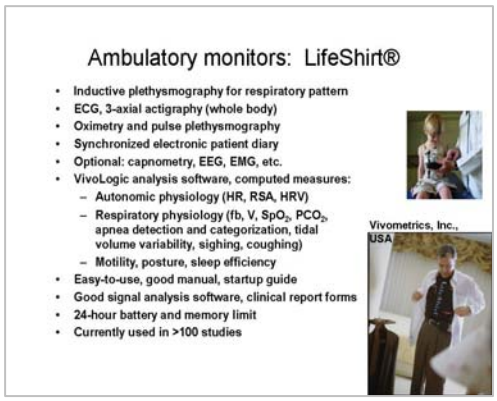
The kinds of analyses people are doing with this are, for example, looking at sleep problems in chronic pain patients. This is one of the simplest but very fruitful ambulatory research approaches that has been used in more than 200 studies now. One device still costs about $900 but if they are used in larger quantities the price should drop considerably.



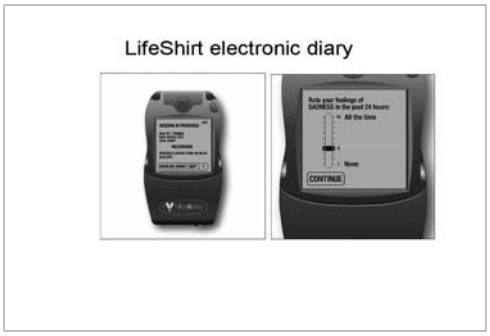


This has not been used much but may be a cheap and effective solution for some questions. The Polar heart rate watch model 810i costs maybe $300 or less. Besides ambulatory heart rate you even get an index of RSA from the device. It records interbeat interval of cardiac activity with a simple sensor band around the chest that is quite accurate in detecting R-spikes in the ECG. It has a wireless connection from the sensor to the watch where the data is recorded. A very effective, low subject burden solution. It allows post-processing of the interbeat interval data. I have written special software to edit and analyze this kind of data to compute reliable spectral and nonlinear system parameters, which you can find on my website. On the lower right you see what kind of laboratory protocols some researchers do with this device.

This is another monitoring system from the Free University of Amsterdam. It is small and also very effective. It can do impedance cardiography to derive some important indices of sympathetic cardiac activation, like pre-ejection period. An alternative system does EDA, electrodermal activity, as a sympathetic measure, and, of course, heart rate and RSA are always measured, as well as motility using a little accelerometer integrated in the device. It comes with very good signal analysis software but it has small memory and that necessitates online R-wave detection. This means that you don't get raw ECG, which would be good for post-processing if there are many measurement artifacts, which is usually the case. For impedance cardiography only ensemble averages across several beats are stored to save memory. Thus, no full disclosure analysis of wave forms is possible. Other devices, like the Vitaport or LifeShirt, have much more memory and do not impose this restriction. That's something one needs to consider in doing this kind of research.



We've been using the LifeShirt system for a few years now. As you can see, it consists of a garment with a lot of sensors woven into it; for example, inductive plethysmography to record respiratory patterns; ECG, a 3-axial actigraphy sensor at the chest for whole body movement, which is better than just measuring arm movement with an activity watch, of course. It is now even available in pediatric sizes. It can also measure oximetry, pulse plethysmography, and a variety of other physiological functions. It comes with a synchronized electronic patient diary, which I'll show you later. It has very advanced analysis software that was really built to make it possible for clinicians to do this kind of study, or even use it as a kind of diagnostic device, for example to screen for sleep apnea. I think there are now more than 100 studies that use the device, few of them published because the device is so new. It has a 24-hour battery capacity and is really easy to use because it has good manuals and support. That's what is lacking for the other multi-channel systems.

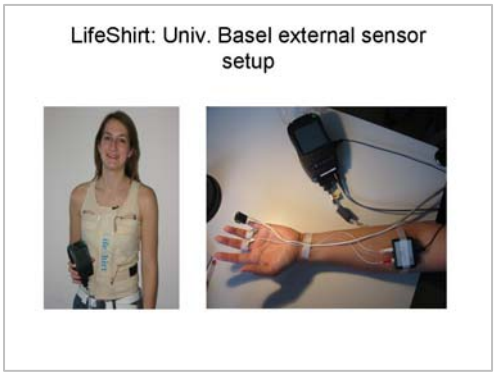


This is the in-built electronic diary and people have sliders to do ratings or answer questions.
In a recent project, we used a separate e-diary because the screen was brighter and we could program it in German more easily for our patients. We have multiple selection options for what symptoms they have or in what context they are, also rating sliders, and one can put in times, how long one stayed in bed and so on.

This is a display of what kind of raw data you get initially, and the amount of data you actually are recording is astounding. Thus, an efficient analysis software is very important, and also ways to deal with artifacts or abnormal readings. One way to deal with these is to use minute-by-minute medians and interquartile ranges. Medians are less susceptible to outliers, so this is something that really helps when looking at the data.

Then you get tables with analyzed values that you can export into Excel. There's a lot of parameters that can be extracted optionally - up to 150 - but for specific studies, depending on your interests, there are specific ones you would typically want to look at.

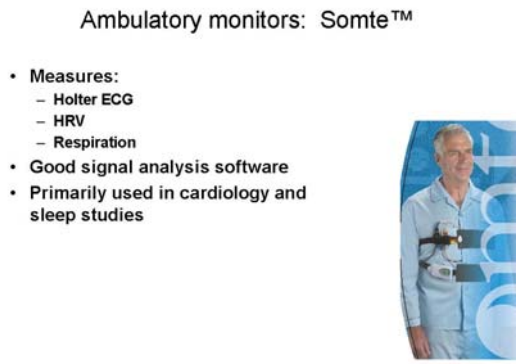
We have also used some additional sensors, for example, for measuring EDA, that is, electrodermal activity, in anxious patients. You see the pulse plethysmography sensor here for measuring pulse wave amplitude and pulse transit time as indices of cardiovascular sympathetic activation. You can see how this is set up. It's a little inconvenient to wear these sensors, but for some questions that can be addressed with them we think the ratio of information gain to subject burden is okay.



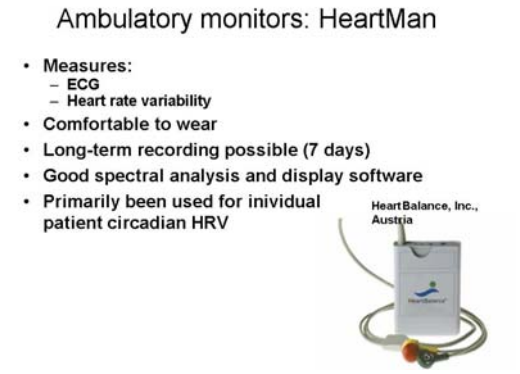
The second signal from the bottom represents electrodermal activity data we're getting during a course of exposure therapy in patients with panic disorder and agoraphobia. You also see all the other channels of respiration, ECG, heart rate and so on. The accelerometry at the bottom is important to quantify how much people are physically active. When people move more, heart rate and electrodermal activity goes up, and when you're interested in stress- or anxiety-related heart rate, you need to take this into account.

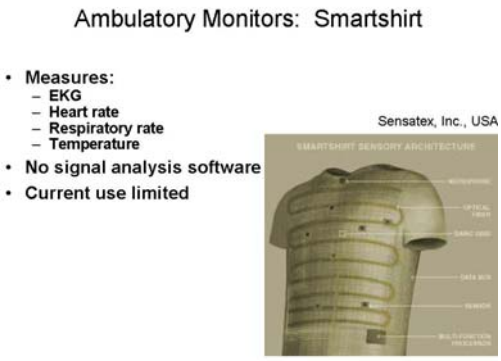
This is an interesting little recording we did where a patient with social anxiety disorder had to give a monthly presentation at his company at a round-table. When you look at the heart rate (HR) signal, you can even see that he got anxious in anticipation. With the speaker before him, he already has an elevated heart rate, and then he's next and his heart rate is at 157 beats per minute almost for the

entire 10-min presentation. He actually said, "Well, it's not such a big problem, but it might not hurt to monitor it." Actually, his amygdala was clearly firing and that is the feedback we gave him, and he was progressing much better in therapy after that.



There is another monitoring system that also measures respiration and ECG. It's a very small system. You'll find all these things in the proceedings, so I'll just jump over them here. Some things you might want to look up on the Internet and see if they might enhance your own research.

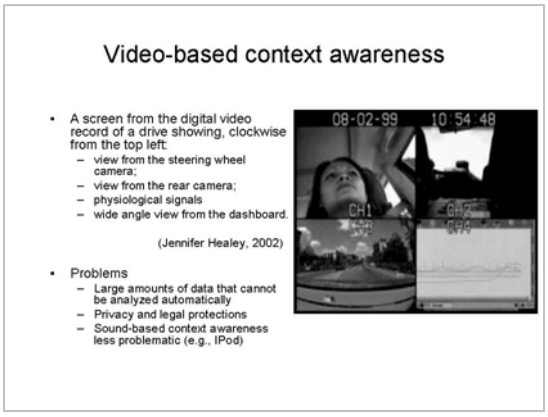




This is a device that measures metabolic activity, and it's maybe a good research tool for a population that is obese.

Here is another one. I'll just jump over that interesting device to monitor autonomic rhythms across the day with very good software for spectral analysis.

There's not really enough time to say what I planned to say. Let me conclude with a statement that the crux of ambulatory monitoring is not that you record these things but that you know what you're doing and that you are planning the design and the protocol of the study carefully. The main problem is physical activity when you're outside of the laboratory. If you're not supervising patients they are walking around at different speeds, going up and down stairs, and that elevates heart rate and affects any of these autonomic and respiratory measures tremendously. Actually, the explained variance for heart rate across the day by physical activity is about 60 percent. Another factor that explains variance is social engagement. When people speak, the heart rate goes up by 5 to 10 beats per minute, and that's not a panic attack or anything. Then there are substance intakes. Several times a day you can get an increase in cardiovascular activity just from the food, caffeine or nicotine intake. There are, of course, sleep/wake effects that you need to track. How do you do that?



Well, you really want to avoid the circadian and random variation in biomarkers you're interested in, and thus you need context awareness: What are people actually doing? With this kind of study design you're getting closer to a laboratory study where you always know what people are doing because you're telling them what to do. In an ambulatory setting, you can obtain context awareness by looking at specific channels. For example, an accelerometer will tell you if a person is physically active; a light sensor will tell you if a person is inside a building our outside. You can look at items in an electronic diary that people are asked to fill out at certain times. When they engage in important context changes, for example when they enter a car and drive, or when they go home and watch TV, you want to know that. In some studies it is very advantageous to use a semi-structured protocol, like in the flying phobia study I told you about. For each step of the exposure treatment we told them in advance what to do, where they needed to go in the airport, when they need to enter the airplane, and so on. To obtain the exact timing of these steps participants were asked to press an event marker. Thus, later in the analysis we were able to track down each situation and thus synchronize data from patients across the same situations, which allowed meaningful Group x Task statistics. This was very effective for making sense of the data and for getting good results.

Of course, if you do not want to impose such a structured protocol on subjects you can do other things. At a very basic level, you can stratify the physiological 24-hour data of each subject by activity levels, or even simpler, just look at the minutes when physical activity was very low. Or you can use more advanced statistical models to overcome confounds, like a method Dr. Roth and I developed for an ambulatory study a few years ago called "additional heart rate." With this method you can look at the heart rate that is in excess of the metabolic demand related to physical activity, which is largely the emotional component in the heart rate. You first need to do a walking calibration for each subject, and use this data later to look at residuals in the ambulatory record. For this you need to obtain minute ventilation data as an index of metabolic activity, since both are related to oxygen consumption. This is explained in an article in *Biological Psychology*.

To come back to the beginning, here you see the recording of a person who was having a pretty flat heart rate across the day, but during the soccer TV game, there was a remarkable increase. When you look at minute ventilation as an index of metabolic activity, the pattern very much matches heart rate. But at the end of the day, during the TV soccer game, the heart rate was clearly in excess of what would be expected based on the minute ventilation. So that's additional heart rate. That's the kind of things you want to look at in population-based studies of stress in daily life.

Now, we did this little study that is not yet published where we examined normal subjects' additional heart rate across the day. You can see an interesting circadian pattern that kind of resembles the cortisol rise and fall across the day, and we are just looking into the kinds of interesting interactions these kinds of data might have.

Thank you. (Applause.)

GAVRILOVA: Thank you, Dr. Wilhelm, for a very important presentation.

We have some time for questions.

WILLIS: One of the interests of a number of people here, including myself, is if one can use some of these kinds of measurements in the context of a survey, and I guess that would encompass questions like what would the cost of the device be, to what extent could the device be used by somebody unsupervised, how do you obtain the data back from the person and so on. Have you had thoughts about those kinds of issues?

WILHELM: What kind of sample size are you thinking of?

WILLIS: Well, just to give an example, a retirement study. It would be close to also the study that Stacy is involved with. We have subjects that we have all over the country. Normally we interview them every two years, but we could interview them on other occasions. We go with a personal interview. We could leave devices in the homes or give people devices and we could go back and collect those devices again. The basic survey collects enormous amounts of data on economic situations and various other kinds of data. So the question would be could one supplement the kind of data that we collected in a survey setting, the survey questionnaires and the like, with data that's collected in this way?

WILHELM: Yes, you certainly could, yes.

WILLIS: Logistical cost issues and that sort of thing.

WILHELM: You could, but it comes at a high cost. If you want to do these kinds of detailed measurements in a sample of, say, 500 subjects, I think no one has ever done that, because the recording might be feasible, but then the analysis and really dealing with that data takes overcoming confounds and so on. It's not that you have one score in the end. It's not that it's that far, except maybe for this apnea scoring during the night with the LifeShirt, that's quite advanced. But you don't want just the mean heart rate across the day. That doesn't tell you much. You're better off just using actigraphs and monitor activity because that explains much of that. So it depends on your specific protocol. You are probably interested in autonomic function, how much sympathetic activity, how much parasympathetic activity there is. Unfortunately RSA is like heart rate. It responds very much to physical activity. It's not only an emotion indicator or a health indicator, an indicator of cardiac mortality and depression, but it's actually very much an indicator of physical activity in the very first place during ambulatory studies. So you also need to look at deconfounding or unmasking this kind of data.

LINDAU: Along those lines, we go into the home to collect data, and we collect all kinds of data, interview data, and then we do things like prick people's finger for blood spots, for example, or send them to the bathroom to give us a vaginal swab specimen, which are stressful exposures. Even if we weren't going to leave equipment at the home but we were able to use some of this physiological measurement during the course of the interview and we would have an event monitor where we could track at what point we pricked the finger, et cetera, I wondered whether that would be feasible. Especially thinking about the LifeShirt, one limitation I could imagine is having the same person wear the same shirt. Is there some of the equipment that you think would be amenable to just using it, say, during a two-hour interview where you could record events and that would be okay to use on multiple respondents?

WILHELM: Well, we usually wash the shirt after one use, and it can be washed many times, but if you want to use the same shirt, maybe not. Maybe you should have three or four shirts with you for

that day. That would make sense medically. And other devices where you have sensors on the fingers and so on, you would have to clean those anyhow. People are doing that kind of study, that they are monitoring the physiological response in the homes of people. That makes a lot of sense.

LINDAU: One of the things we thought about when we designed our study was some measure of stress, and there was a lot of discussion about cortisol. We decided that it was not appropriate for us to use because we really only had one point in the home and we weren't going to leave instructions for people to do, you know, seven samples over the course of the day. So to what degree do you think these physiological measures might be substitutes for or better than measuring something like salivary cortisol as a measure of reactivity or stress response, and can you get useful data doing that just over a two-hour period in the home during an interview interaction?

WILHELM: Yes, you can. And I think cortisol is also problematic. It's not the solution. It's also confounded by many things, and you need a very substantial stressor to see a good, significant response. With heart rate, it's better actually. It's much more sensitive. If you have people sitting and have basically a baseline and then a standardized stress, like finger-pricking or so on, it's a very good measure of stress actually. It has been one of the strongest effect measures in all of our comparison studies with anxious patients. EDA is another good one, but it's not so easy to set up and so on. But, as a minimum, we would use something like the Polar system, and you can take that off, clean that with alcohol, and use it on the next patient. That's really the minimum. I would do it even as an event marker on this watch. You can ask people to press it. We did a study where we did this kind of home monitoring with a Polar watch in PTSD patients while they were remembering their traumatic event and we got very good, interesting data.

LINDAU: Have you published it?

WILHELM: It's in press now, yes.

LINDAU: Oh, great.

WILHELM: I can send that to you.

GAVRILOVA: Any more questions? If there are no questions, the next speaker is Professor David Almeida.

Assessing Life as it is Lived: Daily Stress and Diurnal Rhythms of Salivary Cortisol

David Almeida

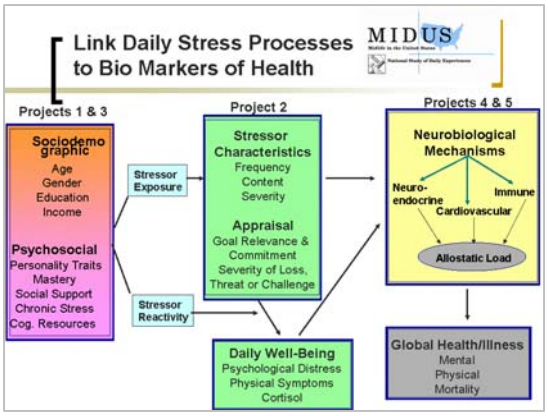
ALMEIDA: As I bring up my PowerPoint, I just want to thank Stacy and Lis for inviting me to this really wonderful workshop. I have to first off admit that I am fairly new to the biomarker game and I've learned an incredible amount already.

Today I want to share with you some of the work I've been doing with the Midlife in the United States study, the MIDUS survey. I direct one of the projects of the MIDUS called the National Study of Daily Experiences, or what I'll call the NSDE. My students refer to the NSDE as the "nasty study" because of NSDE, and I hope my participants don't. Because I'm part of a methods panel, I'm going to dedicate the majority of my talk to describing how we're collecting data in the NSDE. In fact, I'm doing this as well because we're in the middle of data collection. The NSDE is a fairly large telephone diary study that combines information from these telephone interviews with daily collections of saliva that we assay for cortisol. We're doing this in a national sample where we send out packets. I brought props with me that I'll pass out in a second.

Before I get into how we're doing the study, I want to just give you a quick background to the National Study of Daily Experiences. Our overarching goal to this project is to investigate how daily frustrations and irritations sort of get under the skin or how day-to-day living can wear you out. The way that we try to do this, we've developed what we consider a daily health paradigm. So getting away from only concentrating on global characteristics of health, very stable indicators of health, we chart the fluctuations of characteristics of health across short intervals and across daily circumstances. Now, I think that this talk hopefully will dovetail nicely with the previous talk, because even though we're looking at short intervals, they're not as short as the previous talk. We'll be looking at variation from day to day, and I'll try to make the argument that it's important to look at people's health from day to day. We've been doing this primarily relying on self-reports of health, including reports of emotional well-being, such as psychological distress or negative affect and positive affect, as well as self-reports of symptoms that occur on a given day. In this new round of data collection, we've included a biomarker of what I consider daily health, which is salivary cortisol, and I'm going to direct most of my talk to discussing how we're doing this and how well we're doing this.

Now, we think the main benefit of this daily health approach is that it allows us to get a very interesting source of variability, actually multiple sources of variability in health. First is that we can look at between-person differences in health; that is, some people are just less healthy than others, and I think the majority of the talk so far has sort of focused on these sort of individual differences in health, but there's another component of health that might be important, which is within-person variation in health; that is, some days we're just more or less healthy than other days. Some days we have a sinus infection and some days we don't. Some days we have headaches and some days we don't. Even though these five individuals we have here look quite different, on any given day the differences might be reconfigured. The way that we're trying to carry out this paradigm is through

multiple nested studies that comprise the MIDUS study. So my study, and this is a national study of daily experiences, the primary goal or one of the aims of the study is just to chart how variations in daily experiences, and primarily I'm interested in stressful experiences, are connected to daily well-being, both self-reports as well as this biological indicator of cortisol. Here we've taken great pains to assess multiple features of the stressful experience, not just the occurrence of the experience, but the content as well as the severity. We also get a lot of information on the meaning of these experiences. I can give a whole other talk on the measurement of our daily stressors, but I really want to get to our saliva. So the first goal is to link within a person how fluctuations and stressors and stressor characteristics co-occur with fluctuations in daily well-being. On days that you have stressors or certain sorts of stressors, is your daily well-being compromised compared to days when you have fewer stressors?

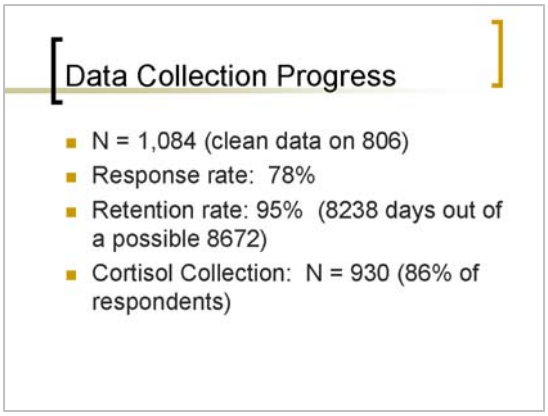


Now, because this is part of a larger program project, we're able to link characteristics about people and their life situations to this daily stress process. So from Projects 1 and 3 of the MIDUS, we have sociodemographic information and psychosocial information - a wonderful panel of psychosocial information, actually, from the MIDUS that we can link to the types of experiences that people are having in their lives - so the types of stressor exposure that we have. For example, we know that younger people report more frequent stressors and they also report more stressors that involve interpersonal tensions compared to older people. In fact, Kira Birditt has done a very nice analysis showing how age is also associated with how individuals in our project react to daily stressors. So in the face of an interpersonal tension, younger people are actually more likely to report psychological distress compared to older people. So we've done a whole series of studies mainly looking at self-reports. We're just now collecting the saliva that will allow us to look at exposure and reactivity to stressors with salivary cortisol and diurnal rhythm of salivary cortisol. The final aim of the MIDUS, and I'll just mention this briefly, is that a subsample of our participants are brought into GCRCs where they're given a full biomedical assessment of their health and a subsample of those folks are given a neurological assessment of how they regulate emotions. So in the end we'll be able to link not only how daily stress processes might map onto a whole range of biomarkers of health but also look at individual and group differences in daily stress processes and how they map onto health. So that's sort of the overall conception of how the NSDE fits within the whole MIDUS framework.

I want to really get to some of the nuts and bolts of how we're carrying this out. Maybe this will answer some of your questions of how we are trying to carry out some biomarker assessment in a large national project. Participants in the NSDE complete eight consecutive daily telephone interviews. I should mention that these participants are a national sample and they agree to answer the interviews every night over eight consecutive nights. During these interviews, we ask a whole range of things that happened to people in the past 24 hours. This is just like a standard survey, but the response frame is always in the past 24 hours, or since the last time we spoke. We get information on how they use their time, including how much time they spent getting emotional and instrumental support; again, self-reports of physical health symptoms and some substance use; some information on their affect in the past 24 hours; how much they got done both at home as well as at work; again, an in-depth assessment of stressors. We also ask about some positive events. We have a checklist of medications that we ask people every day. And on the final day of the interview in sequence, we ask them about how their week was typically for all of these indicators.

To sort of take you through what a typical participant goes through when they enter and complete our study, after completing a fairly long baseline interview from Project 1 of the MIDUS as well as a very lengthy questionnaire and a cognitive assessment, participants are forwarded to our shop. We send out approximately one week before contact a recruitment letter, a summary of previous findings of the study, we send them a check for $25 before we even try to recruit them into the study, and a home saliva collection kit. This is my first prop. In this kit there are instructions about the collection sheet. There are 16 Salivettes in there. There's a UPS biohazard package and instructions on how to get UPS to pick it up, as well as a stress ball. The stress ball is not any sort of experimental manipulation. It's just a token of appreciation that we give to people. If they are very difficult participants, we'll also send them a clock as well that says "MIDUS" on there. I'll pass this out. Feel free to take a look at it and you'll get a sense of what people get. You can see that there are a lot of sort of instructions here, but the first thing that people should see is, right here. It says, "Stop. Do not begin until an interviewer calls you." So they get instructions, but really we'll go over these instructions as well when they get them. I'll just pass this around and you can take a look at it.

Every week we send out between 20 and 30 of these recruitment packets, and we've been doing this now for three years. So every week for the past three years we've been sending this stuff out. Approximately a week after you get the packet, one of our interviewers will call. During this first call, if you answer the phone, we will quickly attempt to recruit you into the study. If you say yes, we do the first day interview. The first day interview takes around 20 to 25 minutes, because part of that interview is an explanation of how to do the saliva collection. The second day of the study, so this is the next day, people start the saliva collection, and then they also will complete the interview that day as well. This happens between days 2 through 5. On days 6 through 7, our interviewers start badgering the participants to mail in their saliva collection kit, and then finally, the last day of the interview, we thank them from the bottom of our hearts and let them know that they might be called again to get on an airplane and go to a research hospital somewhere.

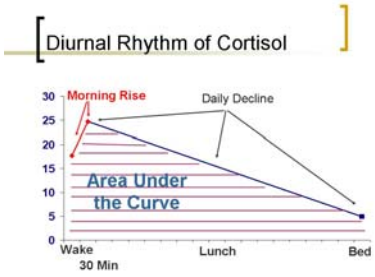


This is our data collection progress so far. Actually this is a little out of date, but when I made the slide, we were at close to 1,100 participants who have completed the protocol. Our response rate is 78 percent. This is 78 percent of people who have completed a very long baseline interview and a very lengthy questionnaire. So the actual participant rate, if you were to multiply out, doing the baseline and the questionnaire, it's probably hovering more around, I would say, between 40 and 45 percent. The good news is that our retention rate is great. We have wonderful interviewers who will not stop calling people until they complete the next day interview. So right now we've completed over 8,200 daily interviews. This is out of a possible 8,600. It's 95 percent of the possible interview days. Once people say yes, they do stay in the study and answer every day. The other piece of good news - and this is the first time I've ever attempted to collect a biomarker and I was very suspicious that people would do it, and when I explain what we ask them to do, I think you'll understand, but, to my delight - 86 percent of the respondents so far have. Actually, it's a bit more: it's 86 percent of the respondents who we have actually received their kits with their saliva samples.

I'll tell you a little bit about our collection. We ask to get 16 samples of saliva from these participants across the 4 days. We have 4 samples on 4 consecutive days. Participants are supposed to provide their first sample upon waking, their second 30 minutes after waking, then before lunch and before bed, and there's a whole set of instructions of what you're not supposed to eat, and brush before you

provide these samples as well. Our target N is at least 2,000 individuals, which would mean, if everyone did it, we would have 32,000 possible samples of saliva. One of the big challenges that I've discovered studying biomarkers in the field is that you have no control over compliance. We can give them written instructions, we have home instructions, but we still can't control what they do. So we have three checks on compliance. During the telephone interview, we ask when people provided the sample. We provide them a home collection sheet in the kit where they write down what time they took their sample. We also ask them what time they got up in the morning and we check whether that's consistent with their wake-up. We also have a smart box that an engineer at Cornell designed for us which automatically records every time you open and close this box. There's a time stamp and a date stamp. So hopefully, when someone says they did the sample, that time and date is going to be one of the several hundred times they can open and close this thing. So we're still working out the smart box, and hopefully it didn't outsmart us. I'll pass this around. You can take a look at it as well. Already it has been opened and closed probably 1,000 times, since it's my sample smart box. So this is how we're doing.

Like I mentioned, 930 people have provided the sample. This is a bit of a different estimate, 89 percent here. So far we have received just over 14,200 samples. We're very happy with our ability to assay these samples for cortisol. People are not missing the samples. Very few missed. Only 434 have missed. So far 237 are out of range. So therefore we have almost 96 percent usable samples of all the total samples.



Comparison of ***Morning Rise*** Cortisol with Wust et al. (2000)

Study	N Persons	N Days	Wake Mean	Wake (SD)	30min Mean	30min (SD)
Wust et al. (2000)	509	2	15.12	(6.3)	22.95	(9.1)
NSDE (Initial Assays)	**930**	**4**	**16.58**	**(11.1)**	**23.37**	**(14.4)**

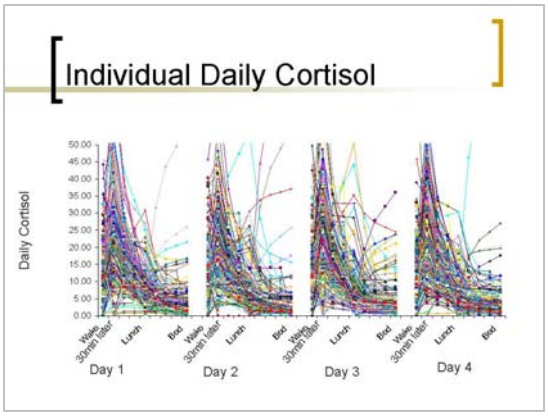
Comparison of ***Daily Decline*** of Cortisol with other Studies

Study	N persons	Sample per Day	Average Slope	SD Slopes
Kirschbaum (1999)	66	25 X 1 days	-.11	.04
Kirschbaum (1994)	20	40 X 2 days	-.11	.04
Cohen (2001)	176	9 X 2 days	-.10	.04
Smyth (1999)	39	3 X 24 days	-.09	.06
NSDE (initial assays)	**930**	3 X 4 days	**-.12**	**.07**

I should mention how this works. I can go into much more detail here, but when people send back in their saliva kits, they send them to our Biocore in Madison. Madison waits until they get enough samples, and then they ship them to Kirschbaum's lab and Kirschbaum does the assay. You all know this, but this is how we parameterize our cortisol. Cortisol follows a distinct circadian rhythm or diurnal rhythm with a morning rise and a daily decline. So on every day we have an estimate of the rise, the decline, as well as the area under the curve.

The first thing that we're doing is checking the validity of our collection with other more controlled studies. This is just a comparison of our initial assays with another published paper that had much more control over collection, and you see that the estimates are very similar. Here is a comparison of our daily decline, our daily slopes with other published studies. Again, it's extremely close. So this certainly made us feel better that we can ask people to do this in the mail and that they will do it well.

Now, what do you get out of collecting so much spit from so many people? I think the real power of it is not just looking at the means but it's going to be looking at the ends of the distribution. If there's some sort of dysregulation that's going on that's going to be implicated in health, it's probably going to be at the ends of the distribution. And having so many people, when we look at, for example, the bottom and top ends of the distribution, we just have a lot more participants to work with, thereby increasing power.



Our challenge, and maybe it's the benefit of doing this, is that with cortisol, now we have even another source of variation. So we have cortisol varies in three ways: It varies within a day, this diurnal pattern; it varies across days, so your diurnal pattern might vary from day to day depending on the types of stress you're experiencing, depending on whether it's a workday or not, depending on whether you're with your kids or not; but also cortisol is going to vary across people, between-person variability. This is what a third of our data look like so far. I almost wept when I saw this. For me it's so beautiful. I've never worked with biological data before, and now I'm sold. I've never had self-report data that looked so nice to me. So here we see the obvious morning rise and then we see the obvious daily decline, but what's hidden in here as well is a large degree of day-to-day variability; that is, your rise and your decline is not consistent every single day. And there is a lot of between-person variability. Some people rise faster than others or their decline is flatter than others. I'm not going to go through this, but this is how we model it. Using multilevel modeling, we can model the within-day variability, that's L1; the across days variability, that's Level 2; and then the between-person variability. Again, I'm running out of time, but we can also add predictors at all three levels, actually. So, for example, at Level 1 we've added control variables, such as: "Did you take medications that day? What time did you get out of bed that day?" those sorts of things, when they took their medications and when they got out of bed. Level 2, we have a lot of information about the daily context of people's lives, daily stressors, as well as all the other things that we're getting in the telephone interview, and then finally at Level 3 is when we include information from the other projects of the MIDUS, so information about sociodemographic and psychosocial characteristics that's going to feed into this. Oh, there's even more. So preliminary findings, again, where we see quite a bit of between-person variation and within-person variation. I've run out of time. It's similar for the rise and the decline.

Let me just tell you about what we're doing right now with the cortisol. First we're looking at between-person predictors in this diurnal rhythm of cortisol. So we've done analyses looking at education, age, gender, personality, and chronic stress, showing that not only are these variables related to the diurnal pattern, the overall level of this pattern of cortisol, but also we're finding the variation from day to day in cortisol is associated with these things. We're looking at within-person daily predictors. Why is it that on some days your pattern is different than other days? We are focusing again mainly right now on the context and appraisal of daily stressors, but we're also looking at sleep. We're finding out that if you wake up earlier than you typically do, that's going to certainly affect your pattern of cortisol, but if you are a late riser and you get up earlier, then that seems to affect your diurnal pattern the most. Finally, we're looking at between-person differences in within-person covariation. For example, we're looking at sociodemographic differences in stressor reactivity. So people who have less education, on days that they experience stressors, their patterns in cortisol are very different than people with higher levels of education. Then what we haven't done yet but we will do is that we will take, again, these individual differences in daily stress processes and link them to the other biomarkers of health in the MIDUS project. And then, finally, we have looked

at this in another study, in our sort of pilot study in Wisconsin, showing that social connections interact with daily decline in predicting interleukin-6.

So I will stop there and just leave you with this conclusion from Kafka who wanted to find out as much as he could from his then-lover, Felice. Kafka was very dissatisfied with the letters that Felice was providing him, so he tried to convince Felice just to give him a diary every day. So he said, "Don't make an effort. Just keep a little diary for me." I feel like that's what I'm asking my participants as well. If biomarkers were around when Kafka was around, maybe he'd ask for her to spit as well. Thank you.

(Applause.)

GAVRILOVA: Thank you for this very interesting presentation. There may be some questions. Any questions?

LINDAU: Thank you. That was a really interesting talk. Emma Adam was here last year. I don't know if you know Emma, but she gave us more of the basic science of the cortisol. The amount of effort that it took to collect the samples is tremendous, it sounds like, and you're still doing that. There are possibilities I know of, but have you considered or will you be getting any other measures from the salivary specimens aside from cortisol?

ALMEIDA: No, unfortunately not, mainly because it's something that we hadn't planned, and with a study this size, you know, just budgeting, we decided not to do it, and with this number, and it kills me to say it, but we're not storing the saliva as well. In two other studies that I'm doing now, we are getting other measures from saliva. We're getting DHEAS and alpha-amylase.

GAVRILOVA: You mentioned that you are going to link to other biomarkers from the same study. Which biomarkers do you plan to use?

ALMEIDA: Well, right now we're very interested in linking daily stress processes with immune functioning, so mainly IL-6 right now. We will have an opportunity to look at neurophysiological functioning. Owen Temple in the back will be working with me on this as well, looking at frontal asymmetry and how that's associated with reactivity to stressors as well. I should also mention the other thing that we're doing. We have a cardiovascular challenge test that's part of the MIDUS Biocore and we'll certainly link our daily stress processes to this laboratory-based assessment of stress reactivity.

GAVRILOVA: Thank you. Our next speaker is Donald Ingram.

Biomarkers of Aging: Marking Time or Moving Forward?

Donald Ingram

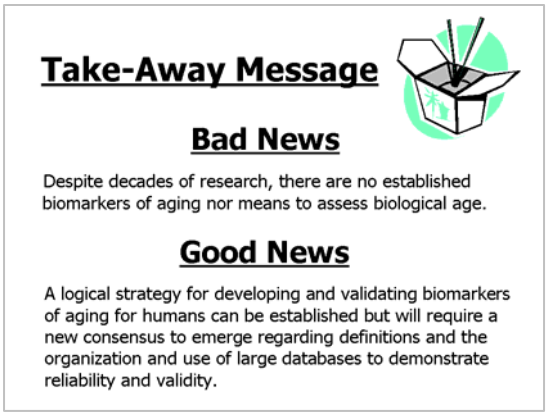
INGRAM: It's my pleasure to be here. After the talk, I think you'll understand why I'm really happy to be here, because while the talk is not necessarily consistent with the title of "Innovative" because I've been at it a long time, the issue is marking time or moving forward with this concept. I feel like I'm in a friendly audience -- I hope I am -- with most of you willing to understand some of the concepts that kind of draw a blank stare from biologists. They just don't get it. And they're not multivariate until recently. I mean the genomics revolution created that. But for a long time they just haven't quite understood some of the concepts that I'm going to present that I hope you'll easily grasp.

Now, when we're talking about biomarkers of aging, and I've heard the term "biomarkers" used in very many contexts today, I'm going to refer to it in very specific ways that I hope I can communicate to you, but it's very difficult. I mean I have colleagues who think it's just impossible. I mean you're talking about too complicated, too complex of a set of biological processes. I don't refer to aging as a process. It's a set of processes that occurs stochastically. So how can you ever really study it? I mean, from the surface we recognize it, it certainly has a face validity, but we don't really understand it that well. Yet we know, and you've already seen today, about the inter-individual variability that occurs. It appears to be that this may be very different, biological age, but, you know, is it really just on the surface, is it superficial, or is there something more in-depth that we can measure - in ways I'm going to talk about - and have it be a meaningful and useful exercise.

That refers to the challenge which has always been in the forefront: Can we stop aging? The answer is no, but the Laboratory of Experimental Gerontology and others in biogerontological research are hoping that we can slow it down a bit. But how would that be measured? You know, here it is, the magic elixir, and we're going to do our clinical study here. How can we go about arriving at a conclusion that aging had been retarded to some degree, and by what degree? So this area of biomarkers of aging has been at the core of this and sort of the question has fallen on its back, as you see here. And yet it doesn't stop the entrepreneurs of the world from claiming that various formulas they have are anti-aging -- don't have a clue what that means -- you know, that you can take hormones that will slow aging, that you can go to clinics and have your aging evaluated through various tests and then have these supplements recommended to you. And really, as you know, there's no conceptual basis to these claims and there's no empirical evidence to support them.

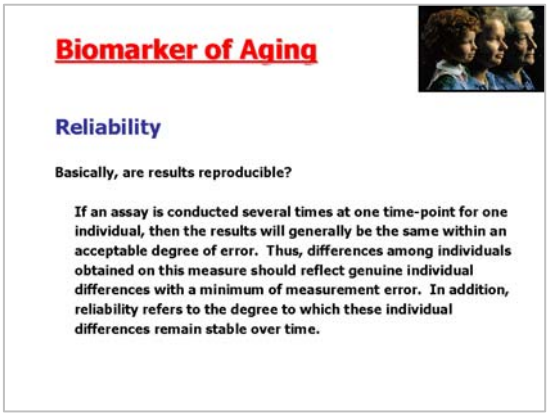
I just want to put things in perspective from Hollingsworth, who had the unique challenge of assessing the damage, biological damage done to thousands of individuals following the Hiroshima and Nagasaki atomic bomb explosion. In the laboratory it did appear that exposure to radiation did accelerate aging, and in other ways than just increasing cancer incidence. So he had the task of looking at the survivors of this event and determining whether aging had been accelerated. So he had that view of it, and I'll give you a summary view at the end of the talk.

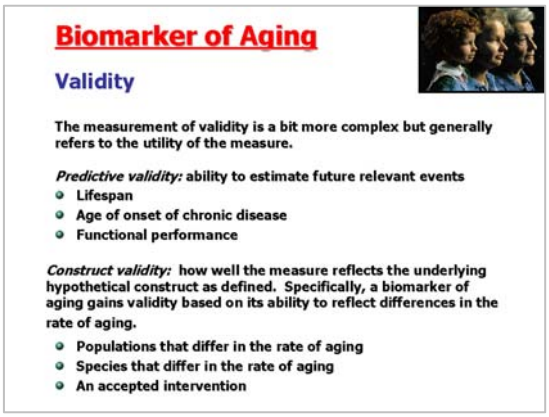
Really biomarkers of aging, many of you in the room don't realize it, but it is the ugly duckling of biogerontology. It just has not come across as a very pretty concept. There have been initiatives that cost a lot of money by our institute that were done in animals that didn't come to any fruition and it has just lost favor, and I'll try and give you a sense of how that has evolved. So the take-away message today is going to be bad news. There are no established biomarkers of aging or means to assess biological age. But the good news, and I hope to have this friendly audience give me some encouragement, there is a logical strategy for developing validated biomarkers of aging for humans, it can be established, and I hope a new consensus will emerge regarding definitions and organization and use of large databases to demonstrate reliability and validity. So, you know, I feel like this is a friendly audience.



Now, this is my definition, and I think it's one based upon a utilitarian concept: A variable design to measure the rate of aging that can predict potential life span, health span, or function -- I like all three. Combined would be great, but any of the three is pretty good -- of an individual or a group of healthy individuals, and therefore be useful to assess the effectiveness of an intervention. This is the utility I see. Now, don't confuse that. It's related to the issue of biological age, which first appeared in the literature also called functional age or physiological age, and that's more of an index that's composed of individual biomarkers of aging that one would claim is a more accurate index of aging than the mere chronological age, and these related concepts have had the same checkered history.

I don't need to explain to most in this audience what I mean by reliability, but biologists sort of don't quite understand. They kind of confuse the terms, reliability and validity. You know what it means. It has got to measure something reliably. Then validity, as many in the audience will recognize, is a process by which one accepts proof of whether there's utility in the measure, and that can be predictive validity, which I've already mentioned to you. These would be very useful instruments if they could have this kind of predictive validity. But the construct validity is a little bit more difficult. And you'll see as we go along that we haven't come up with any consensus that there are populations that differ in the rate of aging. So this is why I think this group is so important. There are species that differ in the rate of aging, and I'll show you that a strategy will provide some construct validity if we move forward in that fashion. And there is no accepted intervention that alters aging rate.





So this is sort of the simple view of what a biomarker of aging is, and this is something that correlates with chronological age, you know, and if we were to take

that as the definition, I mean there's thousands, tens of thousands of things, I mean the color of our hair, the wrinkling of our skin. These are all correlated with chronological age. What do they mean? What purpose do you want to use them for? This concept, though, of biological age was borne out by looking at a number of what were thought to be pretty good function measures, let's just say heart rate or blood pressure, and finding individuals who are biologically younger, you know, over here, or biologically older, and this was the basis for a lot of multiple regression studies. They simply took the residuals of these individual differences and classed them together in a multiple regression equation. Again, you can find tens of thousands of these, and some may be more important. I'll show you. This one is still important. I think this is probably a reliable and valid biomarker of aging, forced expiratory volume. Here's blood pressure, hemoglobin, serum albumin. You know, they're correlated with chronological age, but what utility would they have in terms of prediction of acceptable criterion or the assessment of an intervention?

So there have been a lot of these studies and there have been books written on this and just have not achieved consensus in the field. These multiple regressions, you know, would come up with this regression equation that would predict biological age, and, you know, maybe it would come up with that and maybe it would come up with that. And these were reported back in the '70s and some turning into the '80s, you know, that would attempt to look at the estimated biological age as a function of chronological age. And if you think about it, and my colleague, Paul Kosted (phonetic), claimed in a very elegant expose of this approach that, you know, if you were perfect, if your equation was perfect, it would be perfectly useless, that you couldn't replace, that you have a perfect correlation with chronological age. So this is kind of a silly approach is what he looked at it as. Moving away from that, there are others using our database that the Baltimore Longitudinal Study of Aging published some years ago, again, looking at: Well, okay, so biological age is not a good concept. So maybe just profiles and maybe individuals differ on these. I'll come back to this later. Okay. So that's sort of the concept of biological age, and out of that grew the concept of a biomarker of aging which we began to think about, you know, because if we had an intervention, could we alter this trajectory of change. I forgot to add that most biologists still look at aging in this very strictly linear view. I mean quadratics and that kind of thing confuse them, but they look at linear correlations and have not gone to more complex models.

All right. If this were the intervention and retarded the rate of aging, based upon what? How would we prove it? Well, just to let you know, the field of biogerontology sort of just said, you know what, we're just frustrated with this whole issue and we went to a different set of criteria. At least for animal models, for animal studies of interventions, you know, it's got to impact mortality, it's got to impact morbidity, and it's got to impact function. So we moved away from that. That was based upon the following observations. You know, if aging is so complex and made up of many different complex processes, why is just taking food away from an animal such a robust intervention, and this is the calorie-restriction paradigm, looking at survival of animals, long-lived hybrid animals that have this characteristic when they're fed about 5 percent and these animals are fed about 40 percent, and you see the remarkable increase in the median and the maximum life span that biogerontologists claim has really affected the rate of aging. In addition to that is the counting of diseases, in this case typical tumors that these mice get, and the age of onset in the controls, and the incidence of these tumors is higher compared to the calorie-restricted animals. This was profound. I mean my first time to see an animal with my own eyes, it's like, wow! I mean, these are all genetically identical, same age, same environment, and I bet you could pick out the old mice there, the controls versus the calorie-restricted animals. So aging, although it can be complex, at least this intervention can have very, very robust effects on the rate of aging as measured by a number of parameters, including mortality and disease and function. Thousands of studies have demonstrated this.

What we didn't know was its application to longer-lived species, specifically humans, and so the National Institute on Aging, the study that I've been with since its inception, started in 1987 a study

of 30-percent calorie restriction in the rhesus monkey, and Dario has done an outstanding job of just telling you about that species. And just looking at the phenotype of aging, it's just remarkable, for everything from cataracts to the skin, the spine. I'm a psychologist, so I do study behavior, and the decline in locomotor activity, the decline in simple learning ability, it's all there, and even the decline in food intake, by the way, which becomes a problem, and the types of diseases they get - diabetes and cancer and heart problems - are all very reminiscent of human aging. This here is a 39-year-old that died at the age of 41 and was in fact the longest living rhesus monkey in captivity.

Now, just some hats off to Dario again. Our animal is a laboratory-reared animal, and there are some benefits of looking in the field, but it's sort of in a different direction, I think, than what he was expecting. We were stumped some years ago by this study done on a related species, but out in the field he was chain-smoking, drinking martinis, eating junk food, and was addicted to sex. Here he is at a strip bar. Now, this monkey lived to be 52 years old, which would be the longest-lived monkey of that species. (Laughter.) So that would give you sort of the life history of these animals. But the bottom line is that the rate of aging in the monkeys, if you just take this simple metric, is about three times faster in the rhesus monkey than in humans, and that becomes important for reasons I'll tell you.

We started this study not with the intention of looking at life span and mortality and late life function but with a lot of enthusiasm and promise about molecular biomarkers of aging. That was drilling down to the most elemental components of aging, the processes that drove aging and its manifestation across species. This happened to do with the integrity of chromosomes and whether they could keep dividing or not and DNA damage that would keep the integrity of the entire cell, the glycation of proteins, the oxidation of proteins. These became sort of the promise of aging research that we'd be able to use and to evaluate interventions. In our hands, in a monkey study we could get correlations across chronological age, but note the variability, and this correlation is not much to write home about to begin with. This variability just did not turn out to be reliable over time. The assays are tricky, they did not turn out to be reliable over time, and in general we just abandoned them in favor of looking at the more rigid criterion of does it increase life span, does it reduce disease, and does it maintain function for our intervention, but that's a long-term study in rhesus monkeys. Can you imagine doing that in humans? So we're stuck with the challenge of developing biomarkers of aging for humans. We have written about the strategy. I dare say, and I haven't checked lately, but I'd be surprised if there were 100 hits on any of these papers. We have got a lot more hits on this one which was published in *Science*, and I'll show you a slide from it, but it outlined what we thought was a reasonable criterion -- Dario, again, has already shown this slide -- that many people haven't even approached in this very simple statistical screen, which is: Okay, this is the first step, but you need to look at longitudinal change, because cross-sectional change does not always duplicate longitudinal change. And that change needs to be in the same direction, I might add, and in the same magnitude. It needs to be stable over time. I'll explain more about this as a validity measure a few slides away. So a lot of people looking at biomarkers of aging hadn't even taken that simple approach, and this group can appreciate why cross-sectional versus longitudinal change is important; why you need to look at stability of the measures over time.

So here is the summary, then, of a lot of work that came from a colleague, Eitaro Nakamura, and I using those criteria. So we're looking, then, from a screen of about 30 simple blood chemistry measures and hematology measures, and this is sort of the coefficient of cross-sectional differences, longitudinal change, and the stability of these three measures. So we proposed that as a way of a first pass in screening those variables, and then I'll show you what we did with that. So those are just three. These variables fell out, cross-sectional and longitudinal, stable over time, and then we submitted that to a principal component analysis that, through a little bit of moving of the equation, yielded this equation to measure a biological age score. Okay? So that's the percent lymphocytes, the serum concentration of albumin, creatinine, alkaline phosphatase, and calcium. So you can come up

with a biological age score and you can plot it then for each individual monkey across time and you can get this kind of measurement. And you'll see this was only about seven years into the study. The control group, experimental group - there's some divergence for adult monkeys and old monkeys, and then a standardized score, but that was not significant. You know, we were excited about the method, but I could not generate a lot of enthusiasm by my other colleagues.

All right. Let's talk about why, because we needed to validate that approach, and I talked about that as a possible validation criterion. Here's a measure of DHEAS, which this slide is not going to quite convey, but just take my word for it. Major cross-sectional change with age; longitudinal change in the same direction; stability of this great variability over time. Individuals high in this adrenal steroid maintained that level and low individuals maintained their low level, and we've published extensively on this. Again, this metric of 3 to 1, and you see chimpanzees have a 2:1 ratio, and, like the other species, have higher ratios. And here is then that human change is a percent of maximum of DHEA in males and females. Note that there's no difference in that slope, and there's really, as you know, no consensus that men and women age at different rates. The same for rhesus monkeys, but measuring this slope to that slope, remember the metric should be about 3 to 1 and it's about 2 to 1. But here is what I've been trying to push, but, again, without consensus. We had a lot of data on a primate aging database, and Dario did show that, and I sort of was the catalyst for getting this started, but it has gone in directions of not quite the way I want it to go when it's moved out of my hands.

What the purpose was in my mind was the following, and we've already published on this. Look at all the data points we can get from these many different places to look at a candidate biomarker of aging. This is serum albumin and albumin/globulin ratio, and what we want to do is get confidence in that slope, what it is, and we can do mixed-effects modelings to account for all these differences. This is work done in collaboration with David Allison at the University of Alabama, Birmingham. And then, and this is hypothetical, if we had all these comparative data and we had the standardized slope of change, then the validity logic would follow like this. If this is a valid measure of the rate of aging, then the rate of change should be proportional to the species' differences in life span. It should be greater in shorter-lived species and less in longer-lived species. So this was the logic. And these are real data on albumin. This data is in the iPAD. Two species of macaques and these are some apes, and we measured the age rate of change in albumin, and that would be, I believe, a way of going about validating a biomarker of aging. It makes sense to me logically that it would apply. This is just to show you that albumin in humans has met validity in other fashions. This is the slope of change across age. These are geriatric populations where most of this research has been done in sick and in well populations. These are well, but it still declined. So it's not just morbidity is the point about albumin. Most physicians just think about it as being a measure of poor nutrition or some other thing going on, but this shows that there's change geriatrically in well patients. This thing was a survival study done in Japan looking at individuals in the lowest quartile of albumin and their survival compared to those in higher. So it has been able to predict longevity at least in geriatric populations. This is DHEA again and the BLSA that we looked at, and this was the maintenance of DHEA over time, not the absolute levels, but the maintenance over time, the rate of change, and we published this in that *Science* paper; that healthy men who maintained their DHEA did have a higher survival advantage.

The last part of the talk is real data that is unpublished, again, from our collaborator, Eitaro Nakamura, who is at Kyoto University. These are data that Eitaro has gathered on healthy men during healthy visits at the Kyoto Hospital. You look at some typical measures: Systolic blood pressure; hematocrit, which I don't think we see in the U.S. population; albumin, again, in Japan still shows this great scatter, which is reliable scatter, across age; and here's the one that really is always reliable and shows a change, the FEV. So looking at a principal component analysis, you see that they are related and explain a good portion of the variance. So they look good, they have construct

validity in that sense, but how do we establish validity? Are there human populations that differ in the rate of aging that can be used to validate biomarkers of aging? I've said, well, there's not. We have these possibilities, but I said no.

This is still controversial. This might be a good chance right here. That's an emerging consensus on this. This could be a good chance. This is still controversial. But many gerontologists believe that diabetes is a form of accelerated aging. So what I'm going to show you is that. That's all we have right now. These are diabetic subjects and healthy subjects, using that formula derived by principal component analysis. These are the individual slopes across seven years, and you see the average slope is here, and the average slope of the diabetics - and you can see they're a little bit steeper - is there. Can we conclude from this that diabetics are aging faster? We're using this as sort of a validation form, so the argument is kind of circular. But this is the challenge where we are, in terms of validating biomarkers of aging.

This is the study I showed you before. Barkan and Norris looked at those who had survived who were biologically older versus those -- I'm sorry, these are deceased up here, and those who survived are here. Again, look what's falling out here - blood pressure, albumin, globulin - as differentiating between those two populations. So here is a validation process that we can have.

All right. Here's the last perspective, Hollingsworth. The bottom line is, he says the question then and now is: Do the physiological young live longer than the physiological older, and, if so, what parameters of age predict this validity? He's saying there are massive data banks that exist, and these data would be here. So here is the conclusion. I believe a consensus can be reached on how to define biomarkers of aging and to validate them for human studies, including the statistical methods applied.

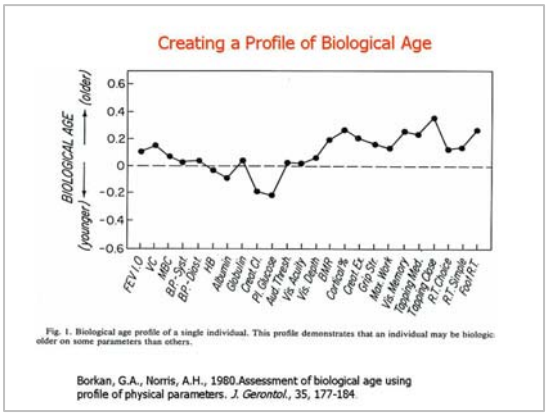


Fig. 1. Biological age profile of a single individual. This profile demonstrates that an individual may be biologic older on some parameters than others.

Borkan, G.A., Norris, A.H., 1980.Assessment of biological age using profile of physical parameters. *J. Gerontol.*, 35, 177-184.

Beginning about a year and a half ago, in conjunction with the Alliance for Aging Research, we convened a working group designed to move toward a consensus. The working group has produced a proposal. We're currently seeking funding and there has been some discussion with the organizers of this workshop regarding that initiative. Increased utilization, expansion of databases, and tissue banks that would offer opportunities for validating biomarkers of aging will emerge. Acceptance by regulatory agencies of a method for defining and validating biomarkers of aging. At our workshop there was FDA representation. They're dying for a consensus to emerge from the field. They're looking for leadership from the field to emerge. Why? Because companies are approaching them and saying, "We have this product. What claims can we make? How do we approach you with that?" There will be emergence of interventions that show retardation of aging in preclinical studies. Our institute sponsors a huge preclinical program now in mice, evaluating such putative interventions, to look at life span, disease, and function.

Regarding more clinical trials directed toward aging interventions that would incorporate reliable and valid biomarkers of aging as their endpoints, this is the big one: generation of research funds to support this effort. We've been walking this proposal around insurance companies thinking they would be very, very interested. They haven't been. So this meeting today is very important and our representation from our extramural staff is very, very important, and I hope we develop that kind of

a momentum. These are a number of collaborators over the years. I've already mentioned several of the important ones. Let's hope that happens. Thanks very much. (Applause.)

GAVRILOVA: Thank you.

CARTER: I'm way out of my field here, so I'm about to say things I hope you don't type down. As a biologist, listening to all this is very painful for me because I see it as kind of a shotgun or a fishing approach. I understand why it's done this way, but in the last 20 or 30 years there's a considerable amount of information. I'm not critiquing your paper. Don't misunderstand me. I'm reacting to the whole idea of a biomarker and how you might go about selecting something that makes sense. The first thing is that some principles come up, and I may be wasting your time, but the body is hierarchically arranged. It isn't a random process. Health is arranged in a hierarchical fashion. It's arranged that way through an evolutionary process. We know actually something about this. We know that the nervous system is sort of sitting at the top of the story. So I think one of the things that needs to come into this story is an awareness -- maybe all of you know this and I just am picking up the wrong cues here as I listen, but there are certain parts of the body that are sort of health-oriented, the neurological systems, the parasympathetic system, parts of the brain stem that are there for growth and restoration, and others that are there for the process of defending the body against an insult, an illness, whatever. So my guess is that if you wanted to find the most sensitive indicators of health, you'd look on the healthy side of that physiological system, and we know some of those. We know that cardiac vagal tone is a pretty good indicator, and it is allowed to be there by the evolution of the autonomic nervous system, certain areas of the brain that are allowed to be online when you're safe and healthy. Other systems, like cort, that are brought in when there's a challenge, they give you different kinds of information. So, again, since I don't know this field, maybe there's a lot written about this that I'm just not aware of, but it seems to me there's some logical ways to go about determining what would be based on physiological principles, and especially putting the nervous system into the story a bit more. Again, as someone who studies the brain, it just seems to me unusual that that's not where you would go first, because that's so important. We know if you take off the brain, things stop working. Death is the outcome of that. But we also know in the brain, you know, which parts are necessary for good functioning. The last thing I want to say is that prairie voles, with all that oxytocin, live two and a half to three times longer than rats, mice, and much longer than hamsters. I mean I gave some to Rockefeller University. I don't even keep them that long. They had them seven years. I don't know if you realize how unusual that is. I forget, and Debbie would know, the hamster's life span is like a year and a half, two years maybe.

AUDIENCE MEMBER: Maybe two at the most. A rat is three.

CARTER: A rat is three. So, now, is there any relevance there? I think so. And what I couldn't tell because you went too quickly over the social mammals, but some of those primates you're showing us are highly social and some are not, and the behavior of an animal, the sort of lifestyle, the social system tells you also a lot. You have to look in a more general way. It tells you a lot about the physiological systems that sort of dominate in that particular species. So there is a lot of really interesting stuff out there. Of course, if I had nothing else, if I could only measure one thing, I'd probably measure oxytocin, but that's a bias of mine. That's not necessarily the answer. I certainly would think that the RSA, respiratory sinus arrhythmia, vagal activity has to be one of the major biomarkers. Frank didn't give you all the background there. I mean he was telling you tools for measuring it. It's not easy to measure, but it's extremely sensitive, because virtually every organ, the immune system, the heart, everything is innervated by both parasympathetic and sympathetic activity, but the parasympathetic is on the healthy side, especially the RSA component of the parasympathetic activity.

INGRAM: If I can respond, you know, I beat up on biologists in my introductory comments, but it's a different approach. I play both sides of the game, just so you know. Biology is a bottom-up. You just stated it in terms of a hypothesized mechanism. And as a quick aside, I did a search on oxytocin and positive neurons in the paraventricular nucleus. They don't change with age. So you would have started with a good hypothesis that didn't turn out to be a good biomarker of aging. Stereological analysis.

CARTER: Swab, s-w-a-b, right?

INGRAM: Okay. I mean let's just say I'm wrong. Let's just say you did that and you found it didn't change. That's a bottom-up, and the people in the room, I think most of them are top-down. You know, they take a big population view and then try to make some sense out of it. Either approach is okay, and I think you'll agree that biology now has been pulled from the bottom to top-up in terms of the genomic revolution. So you have these huge databases. I know where you're coming from. I was raised in the same environment. So fishing is now a legitimate exercise, because the bottom line to it is that these measures have to have reliability. You mentioned your oxytocin assay. It may be great, but it's not going to be robust in terms of, like, oh, we didn't catch it at 9:30 a.m. in the morning. In terms of its utility over time, how good is it going to be in terms of population studies. I think that's the difference.

CARTER: It's not the perfect tool. Don't misunderstand me.

INGRAM: No, no. I'm trying to give you a sense of why this alternative approach. That's what I'm trying to give you a sense of. It doesn't negate hypothesis testing. You know, this would be a good mechanism to look at, just like, you know, DNA damage on the surface of it. We say, like, "Oh, that's incredible. Oh, that's very important." It hasn't proven to be incredible at all unless you do have a huge mutation in your DNA and have some defect; but, for normal aging, it hasn't proven to be that worthwhile.

CARTER: All I wanted to point out, the biomarker approach is not one that I would normally use, and the reason is that it's very difficult in a dynamic biological system for one variable to show you very much. At the same time, if you're looking, I would still look over on the positive side of the equation, not necessarily just on the negative, and maybe you need both. You know, as I say, this is way out of my field, but what we're trying to do, I think, is to index systems that are then keeping the body healthy. Is that the basic idea?

INGRAM: Yeah, but I'm going to provide another bias too. You know, Dr. Gavrilova mentioned my training in psychology, and I have sat through so many lectures looking at this gene or that gene or this nick in the DNA. I don't really care. How is that organism functioning? If aging is so complex with all of these different processes going on independently, they all converge on function: How well do you move around, how well do you learn, how well can you see, can you hear, can you move, can you react? These are the important meaningful terms. It all comes back to that, and I used to have to sit through that, but now I think I'm in the right place.

McCLINTOCK: In your final remarks there, I think you addressed my question, but let me ask it anyway, but, as a prelude to that, let me say I totally agree with you, Sue, about how to distinguish or how can we get a measure of a system that's causal versus a biomarker, which to me means it's just an indicator or a spurious correlation, but still a fabulous indicator. So I have two questions, one of which is, if we use the analogy of a clock, when we say "biomarker," are we going for the hands on the clock, which are a really good indicator of the rate of aging, or are we going after the gears inside that are really driving the system? Both are going to be biological measures, but when we say "biomarker," which are we going after, the indicators I call the hands or are we going after the gears?

Then the question is, what's the phenomenon? You, but everybody in our group as well, keep dancing back and forth between disease and something called aging. Let me just reveal my intellectual background, which is I was hired by Bernice Neugarten at the now Department of Comparative Human Development, and she pounded into me that aging is not the same as disease. Sometimes aging is getting better. So with that, I would say not only are we going after the hands, are we going after the mechanism, the gears, but when we say "aging," what are we taking about and how is it different from a disease process?

INGRAM: It was pounded into me by Nathan Shock, and that's too long of an answer and I don't think it's the focus of our discussion. I would like to address the first part. We're going after the hands. If we were here at a conference on aggression and we were all interested in coming up with a new instrument to measure aggression in a paper and pencil test, we would go through the exercise of, you know, what items on that paper and pencil test would give us some indication of whether a person was going to be aggressive or not, and we would argue over those different items and we'd come up with a 100-item test. We make sure it's reliable. Everybody takes the test and we come back in a week or so and we take the test again and we show it's reliable. Our answers to those questions are stable over time. The next thing we might do is we might go out to the closest prison and we might then measure the paper and pencil responses of that group of people on death row and we say, wow, those people are really aggressive and provide us, therefore, some construct validity to our instrument. Biomarkers of aging are not geared to get towards the fundamental mechanisms at all. The guesses about what they may be can provide you with some insight into what are good candidates, but it's a measurement instrument, it is a substitute for the underlying process, and you only have to demonstrate that it's reliable and valid and that's it.

McCLINTOCK: I was very struck by how lung volume -- I can't remember exactly what it was called, but how, as you pointed out, the lung volume seemed to be a really good predictor of some things, and you (agree) as well.

INGRAM: Yes. You would agree that's pretty vital.

McCLINTOCK: I mean that was pretty stunning there. Is that a measure of smaller volume from, again, microbes and bacteria or is that a measure of plasticity/elasticity of the tissue, and is it a good indicator, because without enough oxygen you get oxidative stress and it kills you? So how do you interpret that? That's my question.

INGRAM: Those are all good questions, but I'm standing up here as a, you know, psychometrician or a gerontometrician. I don't really care. The point is, is it reliable and can I predict something with it? That's really the bottom line.

WILHELM: How much do you take behavioral predictors into the equation?

INGRAM: I just said that's my strong bias. I think that the ability to move and the ability to react in terms of measurement of reaction time is probably one of the most reliable biomarkers of aging. I'm using that term loosely. We have not proven its reliability and validity in that context. I definitely believe it.

WILHELM: So physical activity, for example, diet, wouldn't that predict something more than these functional markers of aging? Behavior might change at any point. You can make behavioral changes even over the age of 60 and you can extend your life.

INGRAM: Sure. If you do an intervention, like, say, you give some exercise and training and movement, you can change that. Right. So you now have an intervention that may alter what we

have not agreed to is a good biomarker of aging but is a candidate biomarker of aging. Things change for reasons. You say, well, exercise can change the rate of aging. I mean that's still a very debatable concept, but that's where we are. We've not agreed on a biomarker of aging, you know, how to validate it.

WILHELM: How about stress response, physical stress or mental stress response, do you have any of those markers in there? Systemic vascular resistance I found in one study, in a cross-sectional study --

INGRAM: Yeah, and heart rate variability might be another one. No, we don't have those. Just so you know, when we went down this road, biomarkers of aging, and when the field sort of halted and we were advised by our scientific advisory committee: "No, you don't have to do that. We don't really care. What we care about now is whether your monkeys live longer, are going to have less disease and function better." I mean that works. We can all agree on the consensus that the intervention that did that, you know, that's it, it did it. But biomarkers are substitutes for the ability to predict any or all of those, and so we gave up that line of research -- well, we damped it down.

LINDAU: I was going to make a comment about the discussion. I've been at a bunch of meetings in the last few months that have interdisciplinary groups of people at all levels, meaning kind of the most microscopic to the most macro level, but who are interested in this interface between the social and the biological. My perspective is that we need both the people who are looking at the gears and who are looking at the hands and we have to feed off of and talk to each other, which I think is a part that we haven't been doing. I don't think that doing what you're doing is a threat to looking at oxytocin mechanisms. I also think that most of us would accept that there is not a universal answer. So you may have come up with a series of markers or hands on a clock that predict time very well and somebody else could have constructed a different clock with different kinds of hands that also predicts time very well, and I don't hear you saying that those would be mutually exclusive.

INGRAM: It's all empirical. You know, prove it. That's the challenge.

LINDAU: Right, and by making those observations, it begs those of us who do the work at the more microscopic level or on the physiology side or on the psychological side to understand what the components are of the gears and how they interact with one another.

INGRAM: Sure. So what I'm saying, you know, the real challenge is a consensus on how to validate. We're taking a poll. Who believes this is a good validation process, diabetics versus nondiabetics, the data I presented you? Who believes that's a good validation process? (No hands.) Nobody? All right. So we don't have a consensus on that. We have no consensus. You have to go one way or the other. Who believes it's not a good validation process? (A show of hands.)

AUDIENCE MEMBER: Validation for what?

INGRAM: For a biomarker of aging. (A show of hands.) So most people are sort of on the fence. Okay. Men versus women. How many of you believe that men age at, say, a faster rate than women?

LINDAU: I think the problem with that question is that it may be very culturally and socioeconomically specific, and so it's hard for me to get my head around generic men versus generic women.

HALL: And additionally, men and women on a number of certain macromarkers age at different rates at different times.

INGRAM: That's correct, because men have equal life expectancy at older ages. I think the consensus in the field is that men and women do not age at different rates. Would you agree?

McCLINTOCK: Well, what about gonadal aging?

AUDIENCE MEMBER: Do you mean in a life span do they not age differently?

INGRAM: Now we're chopping it up in terms of, you know, like middle age or old age and sort of get into that sense. It's a little more difficult.

McCLINTOCK: But the answer is obviously going to be system-specific, right?

INGRAM: But I'm not interested in that.

McCLINTOCK: Yeah, but we are.

INGRAM: I'm interested in disease and function and death. If this stuff I gave to you made your blood pressure go lower and your heart rate go lower and lowered your cholesterol, we would not call it an intervention in aging. We'd call it a cardiovascular intervention. That's what it would be. Okay? So, yes. I mean you wanted to complicate the discussion. You're right, systems age at different rates in different individuals, it's true, and if that's the case, then, you know, this conversation about a biological age or biomarkers of aging, as some of my colleagues have argued, it's just that you can't do it. It's too complex. That may be how we arrived at this conclusion. I don't think so. I tried to show you that calorie restriction just has incredible effects on aging. I mean there's no denying that those genetically-identical mice aged at different rates. There's no denying it.

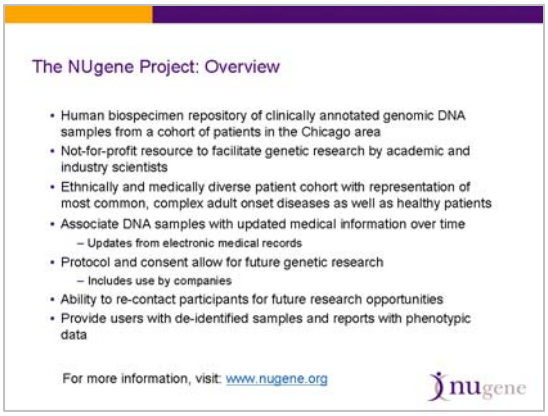
GAVRILOVA: Thank you, Dr. Ingram.

Minimally Invasive Methods to Collect and Extract DNA for Population Research

Wendy Wolf

WOLF: Thank you so much to the organizers for the invitation to come and speak with you this morning, and thanks to all of you for getting up on a Friday morning and coming to hear this presentation. Also thanks to Tom for that lovely introduction.
So what I'm going to speak about today is mostly about methods that are used to collect and extract DNA for population-based research. I'm going to touch only a little bit about the NUgene Project; but if anyone has questions either about the project or anything I discuss today, please feel free to let me know during the presentation, and we'll try to address those as we go along.

Ever since the completion of the human genome project there's been a lot of interest in trying to understand the genetic basis behind health and disease. And, really it's this representation here that we need to go and address in the next, say, decade or 20 years or so, and that is: How does human genetic variation predispose certain individuals to disease or affect their health status? One of the projects that's been initiated since the completion of the human genome project is the HapMap Project, and that's essentially a large-scale project to try to tease out the variation between individuals across the human race. The goal of these sorts of studies is to create basic knowledge that's useful in generating health assessments, improving diagnostics to detect diseases earlier and improved selection of therapies based on genetic information.



The NUgene Project is essentially an initiative at Northwestern University to create a repository of genetic specimens from a small subset of the patient population that's seen at Northwestern-affiliated hospitals. It's basically designed as a nonprofit resource to facilitate genetic research into the causes of common disease. We have both an ethnically and medically diverse population that's served through the Northwestern hospital community, and our population has a number of different common diseases represented within our group as well as normal individuals that seem healthy at least at this point. As part of the initiative we associate DNA samples with medical information, the bulk of which is actually pulled from electronic medical record sources, which requires that we adequately address privacy and confidentiality protections as part of this study. The project is designed to facilitate downstream future genetic research, so it's a pretty broad consent process that our participants go through. Users then apply to use the resource for a particular project and this requires both IRB approval as well as approval by a sample access committee that reviews each request.

As you all know, the causes of different diseases follows this broad spectrum. There are gene-gene interactions that are responsible for many common diseases, but most common diseases fall within this middle range where there's a mix of genetic interactions and environmental interactions that make an individual susceptible to a given disease. While the NUgene Project really doesn't address the environmental component so much, some of the downstream techniques that we're going to talk about today can be used to address those sorts of issues.

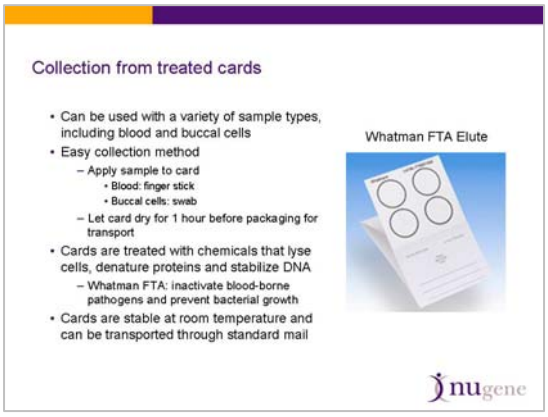
Here is a list of very commonly used collection techniques to collect human biospecimens for downstream DNA isolation. The two major sources are white blood cells, which are typically collected from blood samples, and buccal cells, which are epithelial cells found within your mouth, basically the inner lining of your cheek.

The major advantage with white blood cells is that it produces a fairly clean population of human genomic DNA for isolation. Buccal cells, however, have the advantage in that these methods are very easily administered so a participant can actually do these collection methods in their own homes. It's noninvasive as opposed to finger sticks or blood draws, and they're relatively inexpensive techniques when compared to blood draws. One of the major down sides of buccal cells is that there's a pretty high level of bacterial DNA contamination that you would have with the actual DNA specimen itself. There is a combination of both human DNA and bacterial DNA in these particular samples. And then these are some of the collection methods, swabs or brushes - which we're going to go through in more detail - saliva, mouthwash; and there are also methods to incorporate treated cards into basically any approach by applying these cells to a solid substrate for transport.

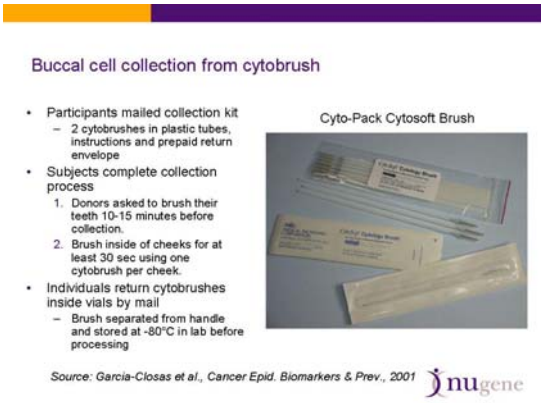
As part of the NUgene Project where we're banking large quantities of DNA in archival form for future use, we chose to use blood draws to isolate large quantities of DNA. And in this particular schema we have a blood draw that's done in the context of a clinical setting, so this particular project is well suited to this type of analysis - you have the patient that is participating in the study through a clinical setting and is in theory pretty amenable to having their blood drawn for research purposes. Those samples are then extracted through an automated system and then stored in archival form in minus-80 freezers for downstream use. Even though we collect the same volume of blood on each individual participant, our yields vary quite a bit using this technology. Our average yield is about 250 micrograms, which is enough for thousands and thousands of typical PCR-based assays. However, the range of yield that you will notice from any given participant is quite dramatic. We have samples where we essentially isolate no DNA, and then we have those samples where we isolate upwards of 600 micrograms; the yield covers a very broad spectrum. As far as noninvasive techniques are concerned, you'd want to consider these methods to improve donor compliance with your particular study, as well as facilitate the collection of large numbers of samples, simplify collection and transport, reduce the transportation and storage rates; and a big consideration for many studies is minimizing exposure both to the participant as well as the study personnel to blood-borne pathogens. In addition some studies that also collect blood samples, either through blood draws or finger sticks, also use noninvasive means of collecting DNA as a backup source.

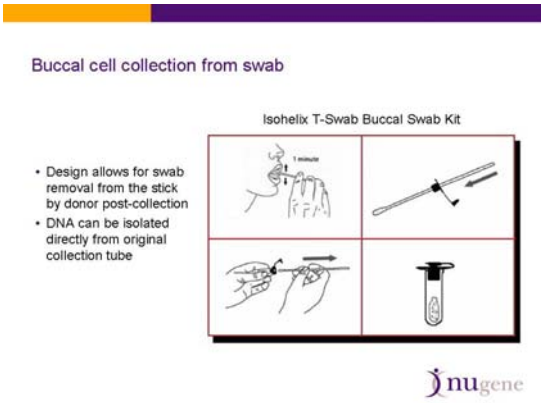
Among the considerations one would want to think about when designing a study to collect specimens for downstream DNA isolation would be ease of use of the actual collection process itself, as well as ease of DNA extraction and storage considerations for the lab personnel. Finally DNA yield, purity and quality information is also important as these are not the same between different collection measures; transportation regulations; and finally cost, which is a significant issue. So if we just focus for a minute on factors that affect DNA yield, these are some of the ones that have been reported in the literature: age, which is mostly an issue related to protocol compliance. You can imagine that it might be more difficult for some individuals to comply with instructions for collection; for example, donor status, diseases or normal. In the case of blood draws it's white blood cell count that actually dictates the potential yield from that individual sample; for sample volume,

there's an optimal volume for most collection methods; storage temperature post-collection, is a big issue; and interestingly enough, whether or not the collection method is done supervised or unsupervised is another consideration. This was a small study that was conducted by the Coriell Institute where they found that the yield of DNA was actually higher from individuals that gave a mouthwash sample in an unsupervised setting with no research personnel present as opposed to a supervised setting. They hypothesize that maybe the mouthwashing was a little bit more vigorous when participants weren't being observed. There's also the consideration of time lag between collection and extraction.



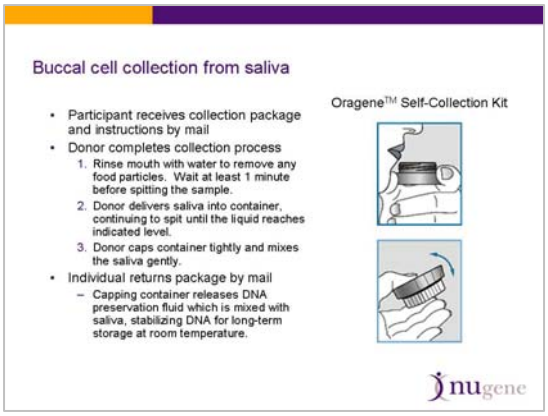
I'm going to launch now into a couple of different methods that are used for collection and highlight some of the different tools that can help you collect this material for downstream genetic DNA isolation and downstream analysis. Collection from treated cards - these are essentially filter-based cards which have different chemicals on them, and they can be used with a variety of different sample types. It's a very easy collection method. In the case of blood samples it's most often used with a finger stick. The person can just apply their, basically, drop of blood to the circle on the filter and that sample then dries and is then ready for downstream storage and transport through mail. In the case of buccal swabs you would swab your inside cheek of your mouth and then apply that brush to the circle area on this filter. In some cases there are specialized filters that can help an individual detect whether or not cells have been applied to the circles by different colorimetric changes. And the nice feature about some of these types of cards is that they actually help facilitate the DNA extraction process by lysing cells upon contact with this filter, breaking down proteins and preventing bacterial growth, which is useful in terms minimizing issues that might occur if mail got delayed. The cards are stable at room temperature and can be transported through standard mail using prepackaged mailing materials.





This slides focuses on buccal cell collection. This is a cytobrush, which has nylon bristles at the end, and in this particular example I'm highlighting a protocol from this particular study. In this example participants were mailed a collection kit with two swabs, and they were instructed to brush their teeth 10 to 15 minutes before collection. In theory this would remove a lot of the bacterial cell contamination within the mouth prior to collection. Then individuals were instructed to use that swab against their cheek for about 30 seconds, one swab per cheek, and requested to return these brushes inside plastic vials through the mail. Once that package reaches the lab, the brush is separated from the handle by the lab technicians and stored at minus-80 prior to processing. So a

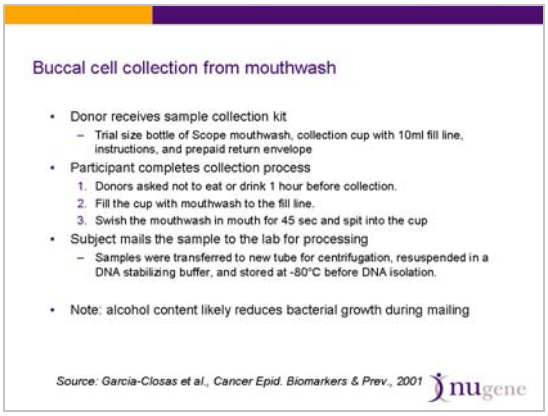
slight variation on that, an example that tools can actually make a big difference in your protocol, is the use of this Isohelix T-swab kit. The big difference here is that it's a swab as opposed to a bristle, but it also has a special cap that can be used by the donor to seal the sample in a small plastic vial and ship a much smaller specimen through the mail. In addition, the swab portion can be detached from the stick again by the participant reducing mailing cost and reducing downstream processing that would be necessary in the lab. This is also very nice because the DNA can actually be isolated in this tube once you decide to do so. There's also the opportunity to collect buccal cells through saliva samples, and this is one particular tool from a company that came out fairly recently called Oragene.



The product has probably been on the market about two years now and it's basically a vial that's about half the size of a contact lens case. In this case, the participant again receives the collection kit through the mail. They rinse their mouth with water first to remove any food particles, and they are instructed to wait about a minute or so before spitting their samples. In this particular container, they're asked to spit two mils of saliva, which doesn't seem like a lot; but when you're actually spitting, it's quite a challenge. As part of the instructions, they recommend if you're having troubling doing this, to add a little bit of sugar to your tongue to help make you salivate a little more.

PRETI: You're collecting non-stimulated saliva, you're not giving the patient anything to chew?

WOLF: That's correct. And in this particular case, your donor delivers the saliva into the container up to the proper volume, they cap the collection vessel, and this actually releases DNA preservative agents into the solution. So that's a good mechanism of stabilizing the cells and preventing downstream bacterial cell growth during shipping or storage in the lab.



Moving on to buccal cell collection from mouthwash samples, again in this case a donor would receive a collection kit that would include a small trial size vial of mouthwash, such as Scope, which seems to be the preferred vendor in the literature, which is great for the company producing Scope. They have a collection cup with a ten-mil fill line and instructions on how to complete the collection process. This particular protocol was pulled from an article that was collecting data on whether or not mouthwash samples should be used as opposed to cytobrush or cytoswab samples. And in this instance they were asked not to eat or drink for an hour prior to the collection. They filled the cup with mouthwash to the fill line and then swished the mouthwash around their mouth for approximately a minute. The subject then expectorates the mouthwash back into the collection cup and mails the sample through the mail. Once that sample arrives at the lab it's stored at minus-80 degrees before any DNA isolation occurs. One of the things to note in these different protocols is that it's often difficult to compare research study results because there are variations in terms of things like what donors do prior to the collection in terms of reducing contamination for the samples, as well as the time that people use for either cytobrush collection or saliva or mouthwash. The nice feature about the

mouthwash protocol, in particular, is that the alcohol content in the mouthwash reduces the likelihood of bacterial cell growth. So if you were to draw a comparison between just two methods, the cytobrush method and the mouthwash method, here are some of the advantages of each approach. In the case of the cytobrush you have easy assembly for large-scale mailings, much easier than you would for the mouthwash approach; light-weight postage, obviously a little brush is very light weight; and the processing is very cost effective. However, mouthwash has some unique advantages as well. On average you get much higher DNA yields from a mouthwash sample. It's higher quality DNA, higher molecular weight DNA, which means it can be used in more robust downstream assays; and I'll go into that in a little bit more detail about this in a few minutes. In addition an important consideration is that it has much higher percentages of human DNA as opposed to samples collected with cytobrushes. Although the percentages vary quite widely, the mouthwash sample would contain roughly 90 percent human DNA, whereas the cytobrush samples can sometimes be only about 10 percent human DNA in the sample.

So far I've presented a lot of different information about different collection techniques, but this is a grid that compares the pros and cons of each different approach. A couple highlights here: while the median DNA-yield is useful information, I would say it's actually more important in most cases to consider the range of DNA-yield that you're likely to get from a given collection method, and this is because if you have different assays designed in mind to do on your resulting DNA samples, you want to make sure that the bulk of your participants are giving you enough DNA for those assays. The range is really critical in determining that because it's very rare that you will get, say, 30 micrograms per mil from every participant for a blood sample. Another important consideration is the molecular weight of the DNA which is an indicator of quality. For most instances here you have high molecular weight DNA in the case of mouthwash samples, the Oragene saliva sample, and then samples drawn from blood. These types of collection methods can be much more useful for some of the robust and newer DNA analysis technologies that are coming down the pipeline. That includes things like whole genome amplification which can be used to propagate DNA and amplify it, so you have many copies for downstream studies.

Specimen storage at room temperature and how long you can actually do that without being detrimental to your sample, is another consideration, mostly for mailing but also a consideration in terms of what sort of tools you'll need in the lab to store the samples prior to doing bulk processing towards the end of your study. And in the case of mouthwash it's weeks; buccal swabs it's very low - days. If your sample gets caught up in the mail, that could be a problem for you. Then there are some techniques that have combined purification and collection schemes. In the case of blood spot samples, you can actually punch the blood spot from the filter paper, put it in a tube and directly do PCR-based assays on it, which is a very nice feature. I should note a caveat to most of this grid, except for this particular line, is that a lot of this information is pulled from a commercial product brochure. However, I think that most of it is pretty aligned with what I've read, and then, of course, the pros of those techniques that are noninvasive in nature.

Does anyone have any questions? Yes.

LINDAU: I do, regarding the DNA yield from blood spots. What's required? Is that an amount based on, say, having four full spots, or is that a punch; and then what's the minimum yield one needs in order to do the PCR amplification?

WOLF: Sure. The first question is how -- whether that yield is representative of basically having a full card on a participant or not. It depends on the card because different cards can hold different volumes, but in this case the median yield is representative of a full card for an individual and so is the range; and I imagine that this very low end has to do mostly with donor compliance and actually getting the sample applied, which would probably be – hopefully, you won't get too many samples

that fall in that range. As far as typical amounts of DNA that you would need to use for a PCR-based assay, it's normally around 10 nanograms for most applications, and most PCR-based assays are pretty resilient, so you can use maybe lower-quality or lower-purity DNA samples and still get pretty robust information from those assays.

LINDAU: Can I ask one more question? Is there any evidence that freezing prolongs the life of a blood spot as opposed to keeping it at room temperature?

WOLF: None that I've seen.

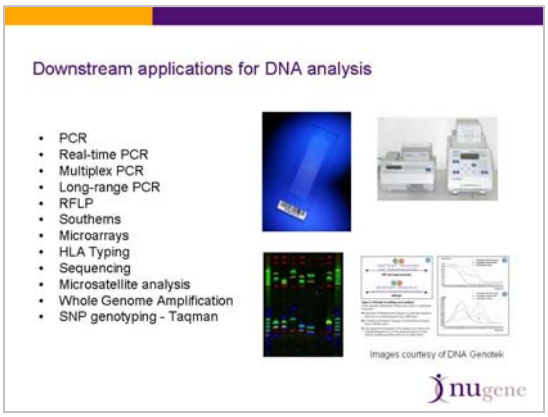
LINDAU: And I had one other comment. I still think, as you made the point, that doing mouthwash might feel embarrassing; that even though we're not actually traversing the skin barrier to get those things, that there's some level of invasiveness. I certainly agree that I would feel that blood spots are more invasive than a mouthwash, but maybe not all subjects would necessarily feel that way.

WOLF: Right. That's actually a very interesting point. In a couple of the articles that I read where they were performing studies to compare different methods, there were examples of participants who declined to do certain methods because they felt uncomfortable doing so, and that was most noted with the mouthwash example.

So transportation regulations, are a big consideration. You have your samples collected by the participants. Now they have to get back to your central lab or core lab for processing and storage, and one of the things to be aware of is regulations through the International Air Transport Association, or IATA. These are actually broadly applicable regulations despite the fact they have air transport in their name because most mailings these days have at least one leg that involves some level of air transportation, and the regulations in this area are only getting stricter. Changes that occurred last year indicated certain parameters for shipping specimens with minimal likelihood of pathogens, and in these particular cases - mostly buccal cell samples but also some diagnostic specimens that contain blood - they have pretty clearly outlined guidelines for how you must package those specimens and label those. Basically you need a leak-proof primary and secondary packaging and an outer packaging that's labeled as an exempt human specimen. Examples of packaging solutions that different companies have used as part of their product offerings, in the case of FTA cards where the pathogens are inactivated on contact with the filter paper, you can actually ship the cards with a desiccant package. One of the down sides of FTA cards is they're extremely sensitive to moisture so you need to have a desiccant package in there to pull away any moisture that the cards might be exposed to, as well as a pouch to protect the card from exposure to liquids and then a mailer which gets labeled. In the case of Oragene saliva-collection, the packaging would include the capped vial as a primary leak-proof container, a secondary container, which would be a biohazard envelope - and one of the new provisions as part of this is that you need to have an absorbent material in this package in case the primary package leaks, and that material has to be able to contain that full volume that's contained within the primary container as well as an outer mailing envelope.

At this point you've gone through all these hurdles. You've used the right collection techniques for your samples. Participants are mailing the packages back in. Your central lab is collecting them. One of the things you need to consider is what sort of DNA isolation method you're going to use. General methods include use of organic reagents such as phenol and chloroform, which gives you very pure DNA, but it's pretty laborious, and most people don't enjoy working with those compounds. Modified salting-out protocols have been used for many years, tend to give a little bit dirtier samples than the phenol/chloroform method, but they're typically much faster. And, of course, there are commercially available kits to help you. In this particular case some of the common vendors are Qiagen and Gentra; and I highly recommend if you have a large study, that you go the

commercial kit route, just in terms of standardizing your isolation process after you've gone through all this effort collecting information and specimens from your participants. I've run across, in my literature searches, a few tips to improve yield from buccal cell samples. In the case of Proteinase K, one of the big considerations you have when you have a sample is whether you are properly lysing all the cells to make the DNA within the cells available for capture. This, in effect, maximizes the potential DNA that you can pull down, isolate and store. And that's pretty routinely used in most buccal cell extraction methods. However, a little bit more unclear as to whether or not this is useful is the addition of glycogen to the DNA extraction procedure. Glycogen is a co-precipitant such that when DNA is bound to it, it becomes a little heavier and easier to extract. So it's most useful in cases where you might expect to have a low DNA yield from your sample, such as in the case of cytobrushes.



So, really, the end goal of all this effort is to be able to analyze DNA through a variety of downstream mechanisms depending on your research area of interest and technology that will support that. This is a list of commonly used techniques to analyze your resulting DNA samples. In the case of the buccal cell samples where you isolated lower molecular weight DNA, there are some of these protocols that might be a little bit difficult to conduct with those samples. One example would be long-range PCR, so if you're actually trying to amplify up a fragment that's large in size, maybe on the order of 5 KB or so, that can be a problem because your DNA samples tend to be a little bit more highly fragmented across the genome than some of the other DNA collection techniques. Another consideration, too, is that some of these assays use quite a bit of DNA in each sample. So Southern blots, for example, some individuals are still using Southern blot methods to detect polymorphisms between different individuals. In the case of cytobrush sample, you're probably not going to have enough to detect genomic DNA from a single cytobrush or even a couple. And then as I mentioned before, some of the other techniques that are a little bit more sensitive, such as whole genome amplification to basically make copies of your existing DNA sample require a little bit higher quality DNA. Does anyone have questions about any of these techniques or approaches and how they relate to different collection methods? Yes.

NIELSEN: One question. With the [unintelligible] can you do the whole genome amplification from blood spots? You're talking about the limitations in that particular method.

WOLF: You know what, I can't say for certain whether you can or not. I think you can, but I'm not certain about that. Does anyone else?

AUDIENCE MEMBER: I'm pretty sure you can.

McDADE: The limitation would probably be quantity micrograms of DNA which is lower, a lower yield of blood spots, but if you -- you can't concentrate the amount of DNA you get in a blood spot or you could use multiple drops of blood, so you can elevate it for some of that.

WOLF: It's a fantastic technology. It's still a little bit early stage, but one of the considerations for using downstream whole genome-amplified DNA for downstream genotyping is how robust it is and representative it is compared to your original sample so that there might be some of your original sample that you would want to save for some of your more sensitive assays, such as sequencing, for

example, and use your whole genome-amplified DNA for some of the less robust techniques for analysis.

WANG: Can you talk about establishing permanent cell lives even though it's really expensive.

WOLF: That technology exists. Again, it's very expensive, and the -- was it the technical details you were asking about?

WANG: I guess I'm thinking all of these techniques have resources, how much material is available. I know some larger scale projects are vested. Creating the cell lines so they don't have to necessarily deal with the triaging aspects of who will get samples and who won't.

WOLF: That's a great point. So for those studies that do have access to a fair amount of financial backing and technically savvy staff members, that certainly is an option to create immortalized cell lines for downstream propagation of genomic DNA on a given individual. The process is pretty lengthy and very, very expensive, in the order of at least thousands of dollars to tens of thousands of dollars per sample.

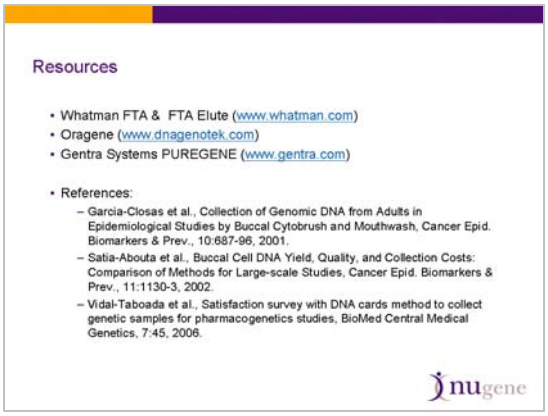
Okay. I'm going to move on to the final part of my presentation, and that was to give you a little bit of information about what this all costs. Here is an example of data pulled from a couple of different commercial entities that provide either extraction kits or the collection resources for biospecimen collection. So this top panel is information from Gentra systems. The middle panel is information from DNA Genotech, which makes the Oragene saliva-base collection product, and the bottom section is information from Whatman on some of the treated cards that they offer. If you're talking about large-scale projects, it almost always makes sense to do your purchases in bulk. So, for a blood sample, you're looking at about a little over a dollar per mil in terms of just the extraction cost for isolating the DNA. In the case of this particular buccal cell kit, which actually includes the nylon bristles, you're looking at on the order of a little bit less than a dollar per sample per brush for both the bristle brush and the extraction kit itself. Mouthwash - you can do a sample for roughly $1.50 each at a bulk rate, and that's just the extraction. That doesn't include the trial kit of Scope that you would need to send your participants or any of the mailing costs. That's just the extraction. For the Oragene collection apparatus, it's still pretty expensive. It's about 12.50 just for that little vial and the downstream reagent that you would use to isolate the DNA. And then for Whatman cards, there are different ranges of cards with different attributes. This indicating card is one of the ones with this colorimetric change when buccal cell samples are applied to the circle, and those are the costs outlined here. This mostly just covers costs of the extraction and some of the tools that you need to facilitate the collection. There are a couple of studies out that there have looked at overall costs for collecting specimens for large scale studies, and this was one that was written up in 2002 comparing buccal cell collection costs, both cytobrush method and mouthwash method as a fully loaded example of the cost for study with 75,000 mailings. In this particular example you can see that with a response rate of about half, it's about ten bucks more expensive more for the mouthwash method than the cytobrush, and a lot of this comes from the postage differences, so the weight of postage, and the processing and handling steps because, for the mouthwash examples, if you remember, once they come into the lab, you actually have to spin down the cells and then store them prior to isolation. It requires a little bit more manual manipulation than the cytobrush method in the lab. This is obviously information based on 2002 dollars, so that's another consideration in terms of scaling up for your studies as well.

And with that I'll just wrap up with resources. Some of the different tools that I featured are listed here. Among these particular websites, Gentra Systems actually has really nice information available on the web in terms of different protocols and technical support guidelines that cover considerations

from the collection step through the extraction step; and these are the full references that were mentioned earlier through my talk. Yes.

HAUSER: This is extremely helpful. I was wondering if I could have a copy of your slides?

WOLF: Sure.



McDADE: Will they be included in the proceedings? They will be in the proceedings.

HAUSER: I want to know if we can get them sooner.

McDADE: Talk to Wendy.

WOLF: Sure, talk to me afterwards. I'd be happy to give you a copy. Are there any questions about anything I presented or other things you'd like me to comment on?

MILLS: That was an extremely helpful presentation, very informative; but I wonder if you can shift your comments to issues of the ethics of gathering DNA. You talked about it from an investigator's point of view; but what about the ethics? You collect this information for subsequent analysis, and what happens if you encounter some individual who has some predisposition to some serious deadly disease? Can you talk a little bit about that, please?

WOLF: Sure, sure, I can. I can tell you a little bit about personal experience with grappling with that issue through some of the discussions we've had around the NUgene Project. In this particular instance when the study was originally submitted to the IRB at Northwestern, the panelists decided that we needed to include a provision during the consent process that allowed people the option of receiving back their results in the event that they were medically significant in nature. And so we actually rolled out the study with that provision. We were a little bit hesitant about certain aspects of the study; but basically after a year's worth of experience and conducting a small DOE funded ELSI (Ethical, Legal, and Social Implications of the Human Genome Project Program) study, we gained greater insight into how participants viewed this provision. I'd say about 20 percent of our participants from our first year were participants in this downstream ELSI study. And what we found from that experience where NUgene participants were interviewed after participating in the ELSI study was that a third of them expected to receive results back, a third of them hoped to, and a third of those individuals did not expect to ever receive results back. And this actually did not match with the percentages of individuals that necessarily selected that option to be recontacted in the case of medically significant results. I think that one of the things we struggled with is, can you put that provision in consent? And no matter how much you talk about it, can you adequately prepare the participant for how likely that's going to be or with what time frame you might actually have that knowledge and whether or not they'll actually want to receive it at that time? Those were some of our considerations, and they were just highlighted by this particular ELSI experience. Yes.

WANG: If I could just add, I'm actually program director for Ethical, Legal, and Social Implications Program. I have to say that there are ongoing efforts now specifically looking at some of the issues that you talked about, not only on how -- is there an obligation on researchers to report back unexpected findings, and they are actually using a lot of the experiences in neuroscience and the whole issues of body and whole body imaging as a way to insert sort of that dialog - but part of the issue I think, and I'm glad you brought it up, and I'm glad you got funding - we weren't particularly

helpful for you. And part of the issue is that the genome institute has been struggling with it, how often do you actually need to go back and ask people, do you still want the results, because there's a camp of people that say you just need to do it in terms of the initial consenting process, but then the question comes how do you know they're informed? And how do you allow for people to change their minds? And a lot of the expenses that is being argued about is, do you need to go to counselors or other professionals? And that's just expensive. So I think that some people are now weighing not to return any results and do the blanket because they are research; but it's really something up in the air, and I know they have some working groups looking at that policy.

HALL: It seems to me there are medically and ethically significant differences between the kinds of results that you might find and that I would assume is not going to be very obvious to the participants. There would be a huge difference, for instance, from someone entering into this and thinking, well, maybe they'll find that I've got 30 percent greater than average chance of developing type II diabetes in my 50's versus someone downstream deciding to look for, say, prevalence of Tay-Sachs carriers in non-Ashkenazim population in Chicago, which would have huge impact or incidence of Huntington's disease or something; two examples that have had a lot of debate about disclosure and whether people want to be informed or not. Could I hear you speak to that?

WANG: Well, the examples you gave, the specific tests would not be included on these large sort of genome-wide tests, sort of scans; but I think the other piece is that there are -- even though you have the genetic information, sometimes people don't want to know. And I'm not sure there are always safeguards in for the participant to be protected from the investigators because the investigators feel that they have an obligation to tell them when they, in fact, are choosing not to know; and the very lively example now is the APOE.

HALL: I guess part of my question is that participants might want to know some kinds of results, and they might not want to know other kinds of results, or they might need a lot more thinking about whether they would want to know other kinds of results. Do you have any mechanisms in place for dealing with that?

WANG: I know that some people are doing the consent forms where they have to check off, which is sort of a nightmare for people trying to keep track, but I think that's one way of doing it. But the other piece, and I think this is where social and behavioral scientists can actually contribute, is to really start doing more research about more risk assessment and decision making, specifically around what do people understand about genomic data and what does it really mean, and really start developing algorithms to help convey this sort of information.

WOLF: I think those are all great points, and I'd like to add that I think it definitely depends on what type of genetic research you're doing and whether it involves largely monogenic disorders or more common complex disorders where it's a little bit more difficult to tease out the relative importance of a given variant or a polymorphism in terms of that person's risk assessment for disease down the line. And I also agree that it's definitely the case that we need to give participants the option, even if they've said in the beginning that they want to receive results, to give them that option again if we do have to recontact them as to whether or not they want to receive results at that time. It is one of those issues where people's thinking on that evolves over time just based on their family experiences, experiences with friends and exposure to different media.

McDADE: Thank you. This is certainly a challenging area particularly since a number of studies are banking samples and haven't actually decided what things they're going to analyze yet. Okay, one more question.

LIPTON: I'd like to get people's more recent experience. As long ago as about eight years ago we were cautioned by some of the IRBs that we were not to be disclosing genetic results to our participants because of the risk of their being denied health insurance coverage for a particular condition if they themselves knew that they were genetically at higher risk for that condition. And so we just said, okay fine. We're not going to worry about it, and we have been telling people we're not giving them any results. Is that still the case for most people?

WOLF: Actually, is somebody in here with genetic counseling experience?

WANG: Sorry, I keep talking through your presentation. There is legislation right now, and I think part of the difficulty is that there's a lot of anecdotal stories about it but no empirical evidence. I think if you talk to most genetic counselors, they would actually say the same thing; but I guess the other piece to keep in mind, since we're so focused on individual consent and protections, I brought the issue yesterday at the lunch about the huge DNA databases going out in public, that in fact there are a lot of people who are donating their DNA for these large studies where there's broad access, and so we may be promising them that on an individual basis, but they may not really understand the magnitude of where their data is going to be going, so they may be losing it anyway.

McDADE: Thank you. I think we need to move on. Much work to be done in this area obviously. So, Wendy, thanks for a very informative presentation.

WOLF: Thank you very much. (Applause.)

THE PROBLEMS AND PROMISES OF BIOMEASURES IN STUDIES OF MARGINALIZED COMMUNITIES: AFRICAN AMERICAN YOUTH AND NIHILISM AS A CASE STUDY

Cathy Cohen, David Williams, and Joel Frader

Moderator: Thomas Fisher

LINDAU: Welcome back from the break. I'm going to give a brief introduction to Dr. Thomas Fisher who has agreed to moderate this very exciting session. I think the more technical orientation to genetics was a very good segway to this talk, especially as we started to open questions around the ethics of collecting such information.

This is a unique session for the biomeasures conference, in that it is an opportunity to use a new research idea as a case study for combined thought and discussion. Imagine that you were part of the team that was going to start such a study. What kinds of considerations need to come into play and do the brakes need to be put on? So we really want this to be an engaging and interactive session.

So in introducing and welcoming my colleague, Dr. Thomas Fisher, I'd like to say that he is an emergency medicine physician who received his doctor's degree at the University of Chicago and continued on in fellowship as part of the Robert Wood Johnson Clinical Scholars Program. I believe he is remaining at the University of Chicago, where he will be a faculty member in the Department of Emergency Medicine. He's been doing fascinating, important research, and is very passionate about how to provide better health care to impoverished communities of color, who have points of access to health care that oftentimes don't extend beyond the emergency medicine setting. His work has involved the integration of methods used by individuals in the business school and outside the formal discipline of biomedicine, and I think he's a wonderful individual to bring together a very interdisciplinary panel and very interdisciplinary audience for this session. With that I welcome you, Thomas, and turn it over to you. (Applause.)

FISHER: That was really kind and flattering. Thank you, Stacy. We have a really star-studded panel here and a very interesting topic to tackle. I'd like to thank Stacy for coordinating this rich interdisciplinary conference and having the vision to begin with a topic of this nature in the discussion of biomarkers. So to begin this discussion, I'd like to give a little format, a little road map of where we're going today. I'd like to begin by introducing the topic, nihilism and measuring biomarkers among African-American youth, and then I'll introduce our discussants. Then, I'd like to open it up for discussion amongst the group. I expect it to generate a lot of interest, and I'd like to allow plenty of time for discussion.

I'd like to introduce Dr. Cathy Cohen. She is a professor in political science at the University of Chicago and the former director of the Center for the Study of Race, Politics and Culture. She was the recent recipient of the Robert Wood Johnson Investigator's Award in Health Policy and sits on the National Advisory Committee for the scholars in health policy fellowship. Her major areas of interest are American politics, African-American politics, women in politics, lesbian-gay politics and social movements. In April 2005, Dr. Cohen sponsored the first Feminism and Hip-hop Conference. She is the author of the Boundaries of Blackness, AIDS and the Breakdown of Black Politics and co-

editor with Kathleen Jones and Joan Tronto of Women Transforming Politics: An Alternative Reader. Her articles have appeared in numerous publications such as the American Science Review, GLQ and the Du Bois Review. In addition to her academic work, Dr. Cohen has remained active politically working in organizations and planning such events as Black AIDS Mobilization, the Audre Lorde Project, Kitchen Table: Women of Color Press, Black Nations/Queer Nations?: Lesbian and Gay Sexuality in the African Diaspora and the Black Radical Congress. Today she's going to present her work on this case study surrounding black youth and nihilism. (Applause.)

COHEN: Okay. First of all, thanks for showing up for the non-PowerPoint session. I actually want to begin by thanking Stacy for the opportunity to participate in this important discussion of biomarkers and African-American youth. I am indebted to Stacy for this opportunity not only because she extended the invitation to attend and contribute to this year's workshop, but more importantly because she's been an important interlocutory in my engagement with the topic of nihilism among African-American youth. Stacy, ever the pioneer, encouraged me to think about the possible biological effects of nihilism in this population, and so here I am today, a political scientist at the fourth annual Chicago Biomarker workshop raising what I hope will be three interesting questions that will generate lively and contentious debate and discussion among this group.

Now, the first question, and I'm just going to list the three questions and then go into them, the first is: Is nihilism a central force in the cognitive and emotional lives of some African-American youth? The second: If nihilism does work to influence the behaviors, attitudes and aspirations of young African-Americans, might it also impact their bodies and health? Third and finally: If there is a biological effect of nihilism on parts of this population, is it even something that we should pursue with new research that includes biomarkers?

Now, before I take up the task of trying to answer these questions or at least engage these questions a bit, let me say what I mean by nihilism. I want to also again acknowledge that as a political scientist interested in the relationship between social structure and individual and community health, my precision with regard to medical terminology may not be optimal, at the very least. I think I took biology and chemistry in college. I can't remember back that far truthfully.

So on the topic of what we mean by nihilism. Cornel West in *Race Matters* discusses the significance of nihilism in black communities suggesting that nihilism must be a concept that researchers concerned with African-American youth across a broad spectrum of disciplines, including those concerned with health and medicine, that these researchers must make central this concept to their analyses. He clarifies early in his discussion of the term that he's not utilizing the concept of nihilism associated with philosophers such as Nietzsche. Instead, West writes that the nihilism he is exploring and concerned with should be understood as, and I quote: "The lived experience of coping with the life of horrifying meaninglessness, hopelessness and most important lovelessness," end quote. He argues that, again, quote: "The most basic issue now facing black America is the nihilistic threat to its very existence. It is primarily a question of speaking to the profound sense of psychological depression, personal worthlessness and social despair so widespread in black America." Now, of course Cornel West is not the only one writing about nihilism, despair and alienation in black communities these days, and I will say often, with no data to support their assertions. Recently there's been a spate of articles and books that highlight the dire condition, in particular of black men in the United States. On March 20th, 2006, a *New York's Times* article begins with the sentence, and I quote: "Black men in the United States face a far more dire situation than is portrayed by common employment and educational statistics, a flurry of new scholarly studies warn. And it has worsened in recent years even as an economic boom and a welfare overhaul brought gains to black women and other groups," end quote. Later in the article, the author writes that the client, and these are clients from the Center for Fathers, Families and Work Force Development: "That the clients readily admit to their own bad choices, but say they also fight a pervasive sense of hopelessness," end quote.

So it is that kind of stark reality of poverty, imprisonment, disease, violence and other life-threatening conditions that West and the clients of the Center are referring to when he writes, and they talk of, an eclipse of hope in black communities. Now, I'm sure that many in this room are familiar with the statistics that undergird this claim of nihilism and despair? Unfortunately, I actually don't have time to recount in detail such statistics, but they highlight the disproportionate numbers of young African-Americans who are the victims of homicide, increasing rates of incarceration, HIV and AIDS, poverty, unemployment and a lack of educational attainment, just to mention a few of the topics. These broad statistics suggest an experience of loss of opportunities, depleted resources and daily racist encounters with individuals and the State that make the possibility of nihilism all too real for far too many young African-American people. And while there is really much research to draw upon when thinking about the emotional disposition of young African-Americans, it has been new data that I've acquired that suggest that we might need to take another look at the emotional lives, strategic behaviors and health of young African-Americans through the lens of nihilism.

For example, data I gathered from three small focus groups conducted in Chicago in 2004 with African-American young adults, ages 18 to 21, suggest that a substantial number of these young people believe there is little hope of progress during their lives and so their goal has become one of survival. Specifically they are engaged in a political strategy of survival that I have called a politics of invisibility. Some of the participants, especially those most vulnerably positioned, indicated they are constantly engaged in the strategy of invisibility - making themselves invisible to authority figures like the police, teachers, welfare caseworkers, correction officers and others that they believe are out to get them. Now, of course, one of the most troubling parts of that revelation, especially for a political scientist, is that a strategy of invisibility is largely a politically disempowering strategy, for if one is invisible to government authorities, then your needs and concerns are unlikely to be heard. Furthermore, through a politics of invisibility, young people lose any power to hold entities accountable.

However, in the realm of health one must wonder what does it do to a person's world view and physical health to live a life, at least partially, of invisibility? How might such a survival strategy manifest itself in physical consequences, such as depression, stress, or hypertension? What does the self-imposed and society-supported goal of invisibility- at least to authority figures including medical authorities - do to their bodies and health? Correspondingly, data from a national survey of young people ages 15 to 25 that I mounted through NORC in the fall of 2005 indicates that African-American young people do in fact believe themselves to be under attack and thus suffer from greater political and social alienation, and marginalization than other young people. For example, African-American young people are more likely to believe that they are not treated as equal citizens in this country, they're more likely to believe that immigrants are treated better than blacks, and that leaders care very little about people like them. These same young people -- and we're talking about a general respondent pool of about 1,600 respondents and 650 African-American respondents - express attitudes of societal alienation and marginalization that should be noted. African-American respondents were more likely to believe that it was more difficult for black youth to get ahead, that black youth receive a poorer education than other young people, that the police discriminate much more against black youth than white youth; that people judge them by what they can buy and what they can own, that they never get their fair share, and that racism would never be eliminated in their lifetime.

Now, one of the most troubling findings was the degree to which African-American youth seemed to have internalized some of the demonizing directed their way. Nearly 80 percent of African-American respondents agreed with the statement that young black people have the wrong morals about things like sex and work; that was 80 percent in agreement compared to 32 percent of whites and 34 percent of Latino youth agreeing with this statement. Now, again while there is not time to

systematically review the findings I've mentioned here, I do believe they provide, support, at least for the idea that at least some African-American youth may be dealing with feelings of nihilism, fatalism, and despair that can impact their social, political and, yes, health behaviors. Given this possibility, the second question remains, which is: Do these nihilistic feelings impact their bodies for physical health?

Now, previous research has repeatedly found that factors such as low socioeconomic status and perceived experiences with racism impact the health and health behaviors of many African-Americans. David Williams, whom you will hear from very soon, in his article race, "Socioeconomic Status and Health: The Added Effects of Racism and Discrimination," writes that, quote: "A growing body of research also suggests that in addition to its effect on health indirectly through socioeconomic position, exposure to racism, and discrimination can also more directly adversely affect health," end quote. He goes on to note that, quote: "Some research also suggests that the subjective experience of discrimination may be an important type of stress that can adversely affect health." Finally, later in the article he asks, I think, pretty poignant questions. He states: "What does it mean for a child to grow up in a society where he or she is viewed as being inferior and where those messages are routinely communicated in multiple ways? A small body of research suggests that the prevalence of negative stereotypes and cultural images of stigmatized groups can adversely affect health status. These studies have found that internalized racism is positively related to psychological distress, depressive symptoms, substance use, and chronic physical health problems." Now, in addition to Williams' work, other researchers have found that psychological and physiological stressors among African-American youth lead to negative somatic health conditions, such as exaggerated blood pressure reaction marked by increased cortisol release. Did I sound like a doctor there? The results of such studies suggest that, like depression, we might expect nihilistic feelings to impact the health behaviors and the physiological health of African-American youth. Moreover, like clinical depression, we should expect that nihilism will impact both the body and brain leading to consequences such as mood swings, the inability to experience pleasure, cognitive abnormalities, high blood pressure, and dysfunction in the cardiovascular and immune systems among African-American youth.

Now, of course one of the related questions that researchers will need to take up is whether nihilism as I've detailed it here is really just another articulation of depression. Specifically while a genuine concern about nihilism may be warranted, it may be that the bodily manifestation of nihilism is exactly the same of what is observed from depression particularly in this population. Now, this question can only be resolved by implementing studies that allow us to measure both nihilism and depression through survey questions and biomarkers among the same respondent pool to see if differences in these concepts can be discerned. Now, finally I think it's time to address our third question, which maybe is the most interesting question which will lead to a larger discussion, and that's, namely, should we, or basically you as researchers, turn your attention to trying to identify biomarkers of nihilism among African youth? Thus, our case study.

In closing I want to review ever so briefly two concerns that others have voiced with regard to such a research agenda. One is the issue of reduction of science. Professor Troy Duster in his presidential address to the American Sociological Association warns of the increasing tendency among funders and government agencies to look at only data inside the body in search of exclusively genetic and biological explanations for social issues. Now, I have to say I'm in agreement with Dr. Duster that we make a grave mistake if we focus our attention exclusively on biological and genetic matters when trying to explain, for example, the lives and behaviors of young black adults and young children. Clearly in the case of nihilism, the major interventions must be public in nature, changing the live condition of young black people. I think we all might agree that in our best vision of a transformed society it will be a long-term project to fully address the more desperate living condition of some black people.

So the question again is, why pursue a biological research agenda, when a social and political intervention is really what is needed? Now, one possible answer to this question might be that the demonstration that nihilism can impact one's bodily health might actually provide greater impetus to the public and to the government to address the social issues giving rise to nihilism. Another answer might be that while we're waiting for needed social and political interventions, there may actually exist medical interventions that can mitigate the consequences of nihilism on the health of young African-Americans. Through trying to identify the existence of biological markers of nihilism and its physiological consequences, we might, in fact, or you might, in fact, pinpoint the appropriate interventions to change any negative trajectory of nihilism on the health of young African-Americans. The bottom line, it seems, is that biomarkers of nihilism, if pursued, must focus on the biological affects on, and not the biological origins of, nihilism among African-American youth. But again it seems to me that this is a question we might return to during the discussion. Second and finally, there's the issue of pathologizing African-American youth, dare I say, again. Arguably, more than any other subgroup of Americans, African-American youth embody, let's say, the challenges of American politics in this post civil rights period. However, in contrast to the centrality of African-American youth to the politics and policies of the country, their perspectives and voices have generally been absent not only from public debates, but also from research on this increasingly disadvantaged population. Ironically, as their presence or perceived threats grows in the public mind through kind of media framing them as criminals and sexual deviants, we have actually less systematic information on the political, sexual and cultural ideas, actions, and choices of this group than we possibly did 30 years ago. So I worry that a focus on biomarkers of nihilism will continue this trend silencing the young people at the center of such research as we seek biomarkers that will be used by less generous individuals and groups and governments to once again stigmatize and demonize this group, this time blaming their seemingly bad behaviors on genetic predisposition and biological nihilism that must be treated with medication and not social policies. Thus, I'm of the mind that we should tread cautiously, if at all, with the research agenda focusing on biomarkers of nihilism.

Responses to Cohen

Joel Frader

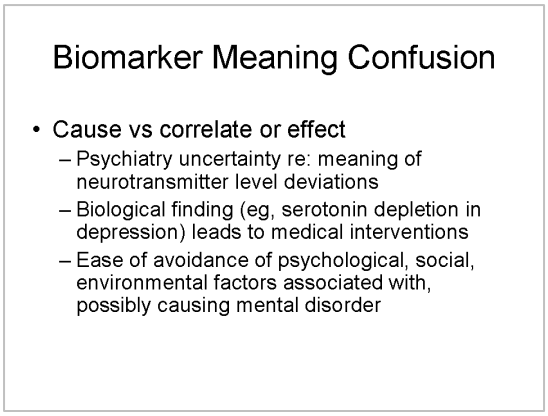
FISHER: Let's continue. I'd like to introduce Joel Frader. He is a professor of pediatrics and a professor of medical humanities and bioethics at Northwestern University. In addition, he serves as chief of the division of general academic pediatrics at Children's Memorial Hospital. His research interests include end-of-life care for children, especially in the intensive care settings, innovation and research in pediatrics as they pertain to surgery and early clinical trials, bioethical aspects of transplantation regarding living donors, and allocation of health care resources for services to children. Dr. Frader served on the Committee on Bioethics of the American Academy of Pediatrics which he chaired for four years, and the Committee on Ethics of the American College of OB/GYN. He was elected to two terms on the Board of the Society for Health and Human Values, including one as the president. He also spent considerable time serving on various federal study sections, data and safety monitoring committees and advisory panels, including those pertaining to ethical aspects of research on human subjects. Dr. Frader has been elected a Fellow of the Hastings Center, a member of the American Pediatric Society and a member of the Chicago Institute of Medicine. In addition he's a clinician whose activities involve supervision of the care of acutely ill hospitalized children and palliative/hospice care for children. (Applause.)

FRADER: I agree with Dr. Cohen on virtually everything she said. I am even more concerned and skeptical than she is and that comes from my admittedly somewhat idiosyncratic view as a clinician and amateur sociologist.

I am really worried about misunderstanding and misuse of biological information about people, particularly youth, whether we call their condition depression or whether we call it nihilism. There are some very important historical reasons to be concerned about this. We, as a public, often mishandle information about biology. The best example during my life in medicine occurred when I was a medical student and a resident. It had to do with the promise of newborn screening tests to save the lives of many young children identified as having sickle cell *disease* prior to the clinical presentation of their disorder, by providing oral antibiotics to prevent overwhelming infection. Indeed, over the last 30 or 40 years that strategy has been extraordinarily successful. However, when screening tests first became available, many more infants, identified as carriers or having sickle cell *trait,* were widely misunderstood to have the disease. Large numbers of these individuals were inappropriately excluded from a wide range of employment, insurance, and other possibilities. It took decades to correct the misconception and turn things around. As a result, many African-Americans understandably worry about the misuse of biological information about members of their community.

We saw similar misuse of biological information in the early days of the HIV epidemic. I remember many, many reports, both in academic sources and in newspapers, on the radio and on television about discrimination against people who were HIV positive in health care settings, in obtaining housing, in access to schools or school programs, etc. Some unfortunate enough to need hospitalization found themselves with dirty rooms, because housekeeping personnel were unwilling to go into the room to clean it. Those confined to beds found themselves unable to get food

because hospital dietary workers would leave trays at the door, because they were afraid to enter the rooms and somehow acquire the disease. Some surgeons, including a very well-known orthopedic surgeon in California, insisted upon wearing cumbersome protective "spacesuits" each time they had to operate on an HIV-positive patient. While surgeons do have a greater likelihood of acquiring HIV in the operating room than the housekeeping or dietary personnel, their behavior constituted an exaggerated response from individuals who should have known better. One might remember that shortly before AIDS appeared, doctors and nurses had experienced very great health risks from hepatitis B infection. Before the availability of hepatitis B vaccine, many health professionals, especially those coming in contact with patient blood in the operating room or in dialysis units, got very sick with high, long-term morbidity and mortality. Yet, those professionals did not take extraordinary measures to isolate themselves from acquiring infection. What happened after the discovery of AIDS was a focus was on the social characteristics of people who had HIV infection, not the biological phenomena. Again, the problem had to do with misunderstanding and misuse of biological information and the identification of biological information with groups of individuals whether or not they were appropriately labeled.



Thus, it seems to me we face a serious problem: pervasive confusion and abuse of biological information, even within medicine. Biomarkers indicative of nihilism pose the same sort of threat. One of the difficulties with modern psychiatry's focus on neurotransmitters has been distraction of the medical mental health community from the suffering of the patients; rather than attend to the *experiences* of patients in their social and psychological environments. In the mental health settings I'm familiar with in pediatrics, once a diagnosis is made, child psychiatrists seem concerned only with the biological, i.e., with writing prescriptions. We see little interest in psychotherapy, providing support, or addressing social and environmental conditions that affect many children with mental health problems, particularly poor children. This raises the same risk that Dr. Cohen referred to in her remarks.

It seems very easy, once we have identified a biological correlate, not even a cause of a condition, to focus entirely on the biology and forget about nearly everything else. We see this all the time in our clinical practice. We see a disproportionate amount of obesity, hypertension, insulin resistance, and other health problems among African-American and Latino adolescents. We all know about troublingly high rates of teen pregnancy. When we struggle to provide treatment and find ourselves frustrated because we do not feel we are making headway, we sometimes ask our adolescent patients why they do not eat better, exercise, take their medication, and so on. Not infrequently we get back some variation on the response, "Well, why should I bother." We hear this especially frequently from the males. They say, "I'm not going to be around long enough to worry about treating my hypertension or losing weight," or something like that. When we talk with many of our adolescent pregnant patients or adolescent mothers bringing themselves and their babies for treatment, they tell us that the most important thing to them is the intimate relationship with their baby, intimacy that they believe is unavailable to them in any other context. Hopelessness is really an important part of the experience of these patients and has a tremendous negative impact on what we are able to provide in the way of medical help. One has to wonder how identification of biomarkers of nihilism, depression, loneliness, or whatever will help the impoverished lives of these youth.

So, I agree entirely with Doctor Cohen's caution. I think we need to emphasize the social and political risk associated with excessive medicalization and distraction or diversion of resources away

from the correlative causes of nihilism or depression, whichever word one prefers. I worry that our fascination with biology means we will not focus enough attention on poverty, on poor schools, on health care disparities. I think we need to take a great deal of care in pursuing this path to biological and, I would add, genetic determinism. Thank you. (Applause.)

David Williams

FISHER: So our final commentator/discussant is David Williams. He joins us from the University of Michigan, where he's the Harold W. Cruise professor of sociology, professor of epidemiology at the School of Public Health and the director of the South Africa Initiatives Office with the Center for Afro-American and African Studies. His research revolves around socioeconomic status, the experience of discrimination or racism, and resulting health effects. He's examined the extent to which psychosocial factors, like stress, racism, social support, and religious involvement are linked to health and social status and can explain socioeconomic and racial variations in health. He's incorporated the MacArthur Network on socioeconomic status and health psychosocial measures, such as MacStatus and reactive responding with biomeasures, such as the measurement of cortisol in two research projects, the National Phone Survey, and the Ypsilanti Study. In addition, he has been involved in the collaborative research of the Network. The Network has undertaken with the Project on Human Development in Chicago Neighborhood. He is a member of the Emerging Science Committee of CARDIA, serving as a link between the CARDIA study and the MacArthur Network investigators. Dr. Williams also serves on the Scientific Directions committee of the Jackson Heart Study, a large NHLBI-funded prospective study, and the Data and Safety Monitoring Board, Multi-Ethnic Study of Atherosclerosis, another NHLBI study. He's a remarkable, active researcher, who serves as PI or co-PI on a variety of projects, for example, psychosocial factors in mental health. His work has received substantial funding from the National Cancer Institute, the Robert Wood Johnson Foundation and the National institute of Mental Health. He's quite a scholar on the social determinants and racial disparities in health. We welcome his comments. (Applause.)

WILLIAMS: Thanks, Tom, for the kind words of introduction. Thanks, Cathy, for a very provocative paper. Since I was commenting on Cathy's paper, I just will share some thoughts.

I think Cathy has done a good job by pulling together in one place a lot of the talks and statistics on the challenges facing African-American youth, with African-American men in particular. She asks three questions. I want to begin with question No. 3. Question No. 3 was: Is there a biological effect of nihilism on parts of the population, and is this something that we should pursue with new research that includes biomarkers? My answer is, yes. I think that social environment affects health. It affects health in multiple ways. The standard used in health policy circles coming out of the Surgeon General's report in 1979, but has been repeated in research by McGinness and others, is that about 50 percent of variations in adult health is accounted for by social and behavioral factors -- much more important than genetics and than medical care. So that, clearly, if the social environment is such a large force impacting health, and that social environment affects health through specific underlying physiological processes, then how does the environment get under the skin? So, for any social factor we presume to have a health effect, I think it is a reasonable scientific question to identify the specific mechanisms and processes by which the larger environment gets under the skin to impact health. But I think as we do this, we need to have specific hypotheses about exactly what we're going after. We need to know why we're asking the question and what exactly we're trying to accomplish. Specifically, we need to think of it not as biomarkers in the abstract, but instead, do you think that health is being compromised through specific impacts on immune function? Is it triggering specific inflammatory processes? So I think we need to have specific hypotheses based on specific research on where we should wear the leverage points to begin to investigate this and pursue along those lines. I do agree with the words of caution that have been expressed, and my first word of caution that I'll come back to and elaborate more on when I discuss questions 1 and 2 is, we need to first of all clearly define and operationalize what nihilism is; and I'll come back to that.

There are a couple claims in Cathy's paper I want to talk about. She suggests that demonstrating the effects of nihilism on health will provide new impetus for addressing the social causes that give rise to nihilism in the first place. Maybe I'm just getting too cynical. I don't think so. I think we have very good examples of where we are now documenting underlying biological effects of social environment, and it hasn't changed anything. I'll give you one example in the area of research on discrimination and health. I developed a scale to capture everyday discrimination. It has been used in a large trial study, a multi-site study of women in the United States. This study has documented that African-American women followed over five years who report high levels of everyday discrimination have more rapid development of coronary calcification over a five-year period of time and have the development of intermediate thickness, that's the measurement of atherosclerosis as developed in the carotid artery. You can actually document physiological effects by adjusting for other risk factors for heart disease on fundamental biological processes. And we have someone like Sally Satel who, in commenting on the findings, says that the finding doesn't say anything because it just shows that these African-American women were probably "psychologically high-strung", and because they were psychologically high-strung, they perceived discrimination that didn't exist, and it's being psychologically high-strung that leads to these cardiovascular effects that have been seen. So the notion that simply documenting an effect with underlying physiological processes will lead to a new impetus to address the problems is what I'm not sure about.

Secondly, there was a claim that identifying biological markers can help us pinpoint appropriate interventions. I'm also very skeptical of that suggestion as well. I think it's only true if you are thinking narrowly of some pharmacological interventions, but I don't think we need to understand fundamental underlying biological processes in order to change fundamental forces that are social in origin. Probably the best example I can think of is early in the history of public health which is the example of the Snow and Broad Street Pump. Contaminated water was the cause of a deadly cholera epidemic in London in 1854. At that point, Snow did not understand what the specific biological mechanisms were in the water. He did he not know what were the processes by which they got into the body, but recognized that if he took bold action and removed the handle from the pump so the community didn't have access to the contaminated water, he could stop the epidemic. I think if nihilism is in fact a problem that is primarily socially caused, the solutions to that problem lie not in identifying the biological mechanisms, but in really taking bold action to get it to social cause in the first place.

At the same time I'm saying, as a scientist, to understand the ways in which nihilism really operates, I think understanding the biology would be important, but I don't think we should confuse the two different reasons. I agree with the point about there's real danger of stigmatization of populations that happen historically, and there's no reason to think it wouldn't happen again. But I also want to make the case that research that is done appropriately and that is appropriately qualified in the area of this era of genomics is exactly important and necessary. One of the things that we clearly understand is that genetics and biology are not static. They're dynamic, and they are responsive to the particular environments in which individuals and groups function. One of the striking findings by sociologists looking at studying race in the United States and studying the impact of residential segregation clearly documents that blacks and whites, for example, in this country, live in very different residential environments. Work by Ralph Sampson, who was at University of Chicago, now at Harvard, and William Julius Wilson, documented in the 171 larger cities, there isn't even one city where whites live in ecological equality compared to blacks, that the worst urban context in which whites reside is better than the average context of black communities. So the distinctiveness of these residential environments is quite striking. And if, in fact, biology operates in interaction with environmental triggers, what needs to happen is to study not only gene-gene interactions but study gene-environmental interactions. When you can look at a group living under very different environmental conditions, I think it is really important; it's a good place to begin to seek to understand how the environment and the biology come together to produce particular patterns of disease. There is some

suggestive evidence of patterns like that that might exist. For example, there have been a number of studies of ambulatory blood pressure that have been done among normal-tensive individuals and found there's no racial difference in blood pressure levels. These are normal-tensive individuals, so no racial differences in blood pressure levels between blacks and whites appear during the day. But all of these studies have found that whites have a larger nocturnal decline in blood pressure than African-Americans do, so that African-Americans maintain high levels of blood pressure even while they're sleeping. What does that reflect? One hypothesis is that it reflects living under very different environmental circumstances so that there is the need to maintain a heightened sense of awareness and a heightened sense of vigilance, and they can never truly relax, even while they're sleeping. And what are the long-term biological consequences and consequences for potentially hypertension for this? We don't really know. So I'm saying I think we need to do more about looking at the biology and environment, but we need to do it carefully.

Now, let me comment on the first two questions. And the first comment I have is to what extent does the concept of nihilism really bias something that is new? I'm really being as provocative as possible when I say, do we really need the concept of the nihilism and do we really need to go out to develop new measures of it, given that we already have other concepts with empirically validated measures of things like hopelessness? We've researched now documenting that hopelessness affects health, independent of depression, for example. We have measures of pessimism. We have measures of negative expectations. We have measures of anomie and alienation. So I am not convinced that we need a new concept of nihilism.

The other point I think Cathy heavily based her claim is to the need for looking at nihilism on Cornel West's work. He talks about the profound sense of psychological depression and personal worthlessness and social despair that's widespread in black America. I think we do have to ask, is Cornel West correct? I think there's reason to believe he is not correct, or maybe he's telling only a part of the story. For example, national data has shown for the last 20 years that blacks either have higher levels of self-esteem than whites do or there's no racial difference in self-esteem. Data for the last 100 years in the United States has shown that blacks have lower rates of suicide than whites do. Now black rates of suicide have been increasing in recent years, especially among the youth, so there are some changes but still, even today, the rates of suicide are lower. The best data we have on depression, rates of depression in the general population, whites have higher rates of major depression than blacks do, and that has been shown in the ECA studies. That has been shown most recently in the National Comorbidity Study and National Study Replication and the new national study we've just completed. Rates, even among the youth of drug use and alcohol abuse, is higher among whites than it is among blacks, although, intriguingly, in the early adulthood rates of blacks really pick up, at a point in which they enter the labor market of going to school. Homicide. We talk about homicide in the black community. If you look at homicide rates over the last 50 years, the rates of homicide today for African-Americans are just about what they were 50 years ago. There is a narrowing of the racial gap in homicide today, primarily driven by the fact that rates of homicide have increased among whites. So the general media perception is that we have these dramatically high rates, yet they've been fairly stable for the last 50 years. I'm not saying they aren't high and that it isn't a problem, but I'm saying it's not new. We are looking at the persistence of trends that have some underlying causes that have been there for a while. So I am not sure if Cornel West is really right, and I am not sure whether the patterns that are being identified are of hopelessness within the population: those things are real, and depression within the population is real. I'm not sure to what extent it's distinctive for African-American youth.

For example, an FDA report indicated that among children who received an antidepressant medication between 1998 and 2002, that the most rapid use of this was among preschool children in the United States. The number of prescriptions for antidepressants was growing 10 percent among children and youth, so that there is a problem of depression. There's other data suggesting that the

rates of depression are dramatically increasing in the U.S. in general. There's other data indicating that the age of onset of depression is also becoming younger and younger in the United States. So I think there's a larger societal phenomenon here, and one of the things if we really want to understand what is distinctive among African-American youth, we need to study not only African-American youth, but we also need to study more broadly, in the population, so that we can see to what extent the patterns are distinct.

I also think the point that Cathy Cohen makes of the pervasive sense among African-American youth, of the persistence of discrimination and the limits of their chances and so on, that's real. That is certainly problematic and symptomatic of some of the larger structural conditions that they face. I'm not sure that's evidence of nihilism. That's evidence they are in touch with reality because that is reality. I think many of my undergraduate students in University of Michigan think America had a racial problem back in the '60s, and that it was solved by Martin Luther King and others. But the work of Deba Patra, et al. and others, showed the persistence of discrimination in American society. For example, this large audit study done in Milwaukee, Wisconsin, used the standard approach of having black and white testers with identical resumes apply for 350 entry level jobs. But there's a wrinkle in this study, in that one of the peer, both one of the black men and one of the white males, claimed that they had a felony conviction for cocaine possession. This study found that not only were there blacks and whites with identical resumes, and not only were blacks less likely to get a job than whites, but that a black male whose record was clean was disadvantaged with an identical resume to a white male who claimed to have a felony conviction for cocaine possession. So it's a dramatic example of the persistence of discrimination in American society, and the black youth are in touch with that reality.

I think one of the challenges in understanding the black population more globally and its youth is to identify not only the risk factors and the problems, which I think are important and paint a very negative picture. You really also need to think of what are the adaptive resources. What are they bringing to bear, that in the light of all of these negatives, that they're still doing that well. And I think we can look back historically, Morris Rosenberg's work in the 1960s is a good example. Prior to his work, there was widespread belief among American psychologists that black kids had to have poor self-esteem. In fact, some of that was used in the 1954 Supreme Court case, and everyone knew that growing up in this segregated racist society, the psychological well-being of blacks would be tremendously damaged. And Rosenberg's work showed, no, they actually don't have poorer self-esteem, in part because of what they were using for the frame of reference. They were not using the white American as a frame of reference, but were using more proximal, immediate context as a frame of reference. He showed this was not something unique to blacks, but he showed it was true for Protestants, for Catholics, and for Jews. It really depended on the specific environment which they were in. For example, Jewish kids growing up in heavily non-Jewish neighborhoods tended to do more poorly in terms of self-esteem than Jewish kids growing up in heavily Jewish neighborhoods. So he showed there were some fundamental psychological processes that operated.

So in summary, I think that the challenges of the black population that are raised here are really important. I am not certain that we need a new construct. I think that we need new research, but we have to do it very carefully. We really need to study the biology not in the abstract, but a biology in interaction with a social environment. And I think what your paper probably points to more clearly is the need for interventions to address the underlying problems that black youth face. Thank you. (Applause.)

FISHER: That's quite a way to start our morning. I'd like to start the discussion with a couple ground rules. We've got about 25 minutes, and in order to get as many people as possible, if you have a comment, I'd like you to raise your name card, and I'll write you down. And if you have something that just has to be said as a follow-up, raise both hands, and I'll get you in.

But I think to begin, there are some very interesting points raised about how the life experience of African-Americans is very, very different from that of the majority community, and that may impact health and health care in a number of different ways. These lived experiences are very different both within the social determinants of health care, the differential access to the goods and services of our society, or increased exposure to negative experiences and environmental hazards. But these things also progress into the health care environment, via access, and then once in health care, differential care by physicians and extenders. That being said, this notion of using biomarkers to measure the impact of these social determinants on health care is intriguing, and I think everybody has raised the concern that this could be potentially used negatively. I'd like to begin with the position that maybe that's not true. It seems like in general, our public adds increased weight to things that are genetic or biological. That is real and tangible, whereas just the description of different experiences and different health care may have some spurious or very difficult link to transition. Whereas if we can measure that African-Americans who have these opinions, also have these differences in health care, maybe this is an opportunity to mobilize resources surrounding these social determinants in health and health care. In contrast, there's also the tension that was raised that maybe this also leads to the decision that these people are just defective, and they're ill, and this is no longer our problem. It's, in fact, genetic. It's not societal at all. I wonder if anybody would like to address that tension that may exist?

McDADE: I think those are great points, and thanks to the panel for really provocative thoughts. I think it's on that point, it's very important to be very clear in our language, when we talk about biology versus genes, and what we're measuring, because the causal arrows can be totally different; so a lot of times, when we talk about biological determinism, when in fact what we mean is genetic determinism. A lot of times when we talk about measuring biology, people immediately assume genes when in fact we may not be doing that at all. So if you are measuring genes, then of course there is an obvious risk of locating the cause of some phenomena, like Troy Duster does in his very nice address, it's bankrupt with respect to intervention because there's nothing you can do, and worse case scenario, you end up blaming the victim. But when you measure the biological impact of social phenomena, which most of us as social scientists are fundamentally interested in doing, and you keep laser precise focus on those social context, then this becomes a very powerful analytic and policy tool for the reasons you're suggesting, because biology gains attention, and it has direct links to health outcomes that people care about. So I agree with the sentiment of a number of people here that measuring biology in this context and modeling it in sophisticated ways specifically in relation to social and economic and political phenomena has real promise.

FISHER: Doctor Wang.

WANG: Even though I'm a geneticist, I think I really experience the panel's perspective. I guess I'd like to hear more about the issue of genomic research and how it influences social behavioral research, because I agree with you that we need to be clear in our language and what we're actually measuring. But it's within the context of genomic research, where there is a huge effort in trying to understand genetic variation and linking it to population migration. There are groups of people who believe that there are unique ancestral genetic informative markers. So I'm wanting to hear more dialogue, given that there's a huge camp of people who have this very seductive technology behind them and ways we can articulate the perspectives here which I hear people that are valuing within that context.

COHEN: Actually I don't think I'm the person to answer that question. I do, however, wanted to just say a couple of points if that's okay.

FISHER: Be my guest.

COHEN: I'm going against the rules I'm sure, to clarify some of the things or address some of the things, in particular that David, who I've known for a very long time and expected him to find all the weaknesses in this paper, which was supposed to be provocative and not necessarily comprehensive. I am not in any way wed to the idea of nihilism as a concept. I think it's an important jumping off point, because of the significance of Cornel West. That has gained resonance within social science communities that I think we have to address. Now, having said that, I think the observation that pessimism and hopelessness has an impact independent of depression, also speaks to the need of trying to understand what Joel also talked about. What's happening in the lives in particular of young African-Americans? Much of the data that we have are for actually older adults in terms of thinking of African-Americans at least 18 or over. For example, the data that I'm refering is 15 to 25, so we can begin to get a clear sense of even earlier attitudes towards kind of the fatalism, or the opportunities that are provided them.

But that said, David, I understood you when you said we need to understand biomarkers and the biology, not because we think it will provide interventions, or it will convince government entities to support the type of interventions that we need, both pharmacological or social interventions but, rather, we need to understand biomarkers and biology, because we're good scientists. And I think I hear different things. On the one hand, I hear people saying, you know what, if we do this right and if we pay attention to the social and political interventions and constructs, we actually can come up with interventions. But I thought I heard you saying we're probably not going to come up with interventions. First of all, can you clarify?

WILLIAMS: Yes, you heard me right. I think that research on underlying biological processes can come up with interventions. But I think if the problem is nihilism, then the interventions we come up with once we understand the fundamental biological processes, would have to be in my mind some kind of pharmacologic interventions. I guess you could develop therapeutic counseling strategies to deal with it as well, but again, to me, those are very downstream interventions. What we need are more upstream interventions to get at the underlying problem itself.

COHEN: So why pursue biomarkers?

WILLIAMS: I think that as a scientist, I want to understand if I am claiming that a social environment affects health, I think we need to trace the pathways by which this happens, so it's more documenting that it is in fact consequential. In theory, if that could be used in the service of getting the country, getting political leaders, policy makers to pay more attention to it, that's good. I am saying that from my view of what is happening in the United States, we are having more and more evidence in multiple domains of how the social environment affects health. I haven't seen a shift in policy based on the new biological evidence, so while I'm saying it's a good thing to do and I endorse it, more from the pure sense of understanding the science, I really want to understand how the social environment has health consequences. I think to my mind it's naive to think that simply if we demonstrate the relationship, even if you can do it with the best scientific evidence you have, that suddenly there will be a tidal shift in terms of the policy making discussions and the ideological debates around these issues. We're dealing with fairly intractable issues in society, and most policy made in this country has nothing to do with science. Sometimes it's great if you can have a confluence of scientific information at the same point that there's a policy question, and science sometimes informs policy making. But as a person sitting beside the Beltway, they don't always say, well, what is the science here and go base it on the science.

FISHER: Mills.

MILLS: I'd also like to thank Cathy and the panelists for a really exciting and stimulating conversation. I think we've really gotten at the heart of interdisciplinary conversation about health and health outcomes. I wanted to say that, David, I was glad that you had raised a question of whether or not we needed another concept. As a sociologist, you brought up anomie, alienation and others, but I do wonder, for example, in terms of the importance of the biomarkers, while I do appreciate the admonition for care and caution, whether it would also be important given that many of the scales and measures that we do have have been normed by white male populations it might be important to have some sense of the biological and physiological responses to stressors that are normed in these specific populations.

The other thing that was really most interesting to me is, Cathy, while you raised your three overarching questions, the real interesting question for me is you raised the issue in what ways does the strategy of invisibility affect the health outcomes of African-Americans, and particularly African-American males. And so I thought that was a very interesting question. Whether we call it anomie or nihilism or whatever label, the issue of the strategy of invisibility and what leads to the necessity of having that strategy.

And then my final comment is I think the discussants and the topic really speaks to the issue of the importance of having ecological models where we have the ability to study, in concert with biology and health behaviors, social environmental things. So really that's what I'm taking from your conversation, the importance of integrating the biology with the social. Again, wonderful, stimulating conversation. I thank you all for that.

WILLIAMS: Can we go back to the question that was raised by Doctor Wang? I mean that's a really important question. It would be useful to have some conversation about it, I think.

FISHER: The question was revolving around the significance of genetics in defining race and, therefore, defining downstream effects of race. Does that sum it up, Doctor?

WANG: No.

FISHER: Be my guest.

WANG: Let me say, I don't know if I can repeat exactly what I said before. I guess I'm wanting to hear more about there's a whole different community that is invested in the technology of identifying ancestral informative markers, and basically that there's specific pieces of DNA that you can identify population groups, which some people refer to as race. That's a whole other discussion in itself. I would love to hear more discussion about what does the use of biomarkers mean within that context, because that other context has a lot of money, and it has a very seductive sort of message. I think there's a lot of private companies right now, where people go to try to find their sense of self, which I absolutely appreciate. They're going to these companies saying I can give you some DNA, and you will tell me who I am, which may or may not be consistent with how they experience who they are socially. And I think this is where the rub is going to come with looking at biomarkers in terms of people's identities that rubs with the biology women.

WILLIAMS: I think part of the context in which this is emerging is what is the role of genetics, not in terms of an accounting for individual differences in disease, but in group differences in disease. So we have these dramatic racial disparities in health in the United States, and to what extent is genetics a central underlying determinant of these differences. And I think some have looked at the human genome project and the work from physical anthropologists for the last 50 years and said there is genetic variation within populations, but it doesn't map onto a racial category. And there's new research saying no. If you really look at enough points along the human genome, you can tell

continental origin, continent of origin, and so that race is still valid as a way of thinking and for use in medicine and for use in health research and for many people, although it doesn't necessarily lead to the other. But some persons see that as validating the need to pursue genetic explanations as fundamental to racial disparities in health.

FRADER: I guess as a clinician my hope is that we can use genetic information to bypass identification with particular groups and go directly to the use of genetic information to assist individuals. So the ability to figure out whether somebody is going to respond in a particular way to a drug that I want to use, when a patient is in the intensive care unit, regardless of what group they may belong to, is much more important to me than some sort of statistical relationship between the patient's skin color or country of origin or whatever. So my hope is that we can forget about the population at some point and go directly to individual uses of genetic information.

FISHER: Can I move on to a couple questions? Dr. Meltzer.

MELTZER: I loved the panel. I thought it was fascinating, and I didn't realize when I came how closely related it was to some work I did a number of years ago. I want to share not just some things that I did, but some things that others have done that's closely related, and it's really at the interface of nihilism and biology.

My Ph.D. is in economics, and I wrote my dissertation on the effects of mortality on incentives to invest in education in developing countries. The idea being if you know your life is going to be short, it doesn't pay to invest; and it's essentially nihilism. In that setting, it's very difficult to test, because the things that cause you not to live very long also are related in a social context to a bunch of other things that aren't necessarily conducive to economic development and so on. One of the big challenges that arises is how do you get sort of exogenous variation in life expectancy, that allows you to test these arrows that Professor Williams was talking about. One of my students, a Ph.D. in the economics department at the University of Chicago, is Avi Stoler. Stoler wrote a dissertation about a year ago that looked at exactly this, using genetic information, and he looked at the change in behavior of people with Huntington disease or a family history of Huntington disease. As we got more information, they were allowed genetic testing. What he discovered was that as people discovered more information about their genetic risk, they became much more nihilistic if in fact that information was negative. And I mention this, because I think it shows on the one hand the power of biological data and information to illuminate these arrows, but at the same time also this sort of policy tension where there are winners and losers as this information is realized. We need sort of the social structures to think about whether this is a good thing and how we should best respond to it. Anyway, I just wanted to share that, because I think it's a very concrete example of a lot of things people have been talking about.

FISHER: Let's get some more comments. Harris.

HARRIS: I was going to go back to the question that we were talking about before. There's a couple of things that I'd like to say. One thing that bothers me about this discussion is the way that people conceptualize genetic influences, particularly genetic influences for the kinds of things that we talk about in behavioral and social sciences, where there's very complex traits, where many genes, many environmental factors are involved. There are gene-gene interactions. There are gene-environment interactions. There are gene-environment correlations where people might be selecting the environments that they're in. You can't really think about genes as being deterministic at all, and I think most people doing research in this area acknowledge this, that genes are not deterministic. I would argue that no gene is expressed outside of the context of an environment, and so the expression of those genes will vary depending on environments and depending on a number of other factors. So the level of analysis that you're talking about where people are looking at populations and

migration I think that the interaction there with behavioral and social research has to do with questions like how does the demographic history, how does the social history of a population through hundreds and thousands of years affect the gene pool in that population, affect which genes survive and why, because they've been under some kind of pressure. When it comes to how this relates to race, I think that is the reason why it's useful. I'm not sure what the best terms are to think about, if race is the right word and people make a distinction between biological constructs and social constructs of race. But the reasons why it's useful in genetic terms is that it helps to understand a little bit more about gene-gene interactions, so that if you're looking at the effects of particular genes, you want to know how the other background genetic variation that you haven't identified is maybe modifying the effects of the genes that you're looking at. This kind of mapping populations to where they've come from, gives you a little bit more information about this background genetic variation, that could be modifying effects that you're looking at.

WANG: I absolutely agree, and I wish on some level, people had that some level of sophistication when thinking about this. I think about living examples that we have day to day, so we think about the drug BiDil that recently came out that was approved by FDA. So, for example, the labeling of the drug says that it is for self-identified African-American people, and so my question was, well, if I say I'm African-American, would you include me in the study? And they actually said yes. So for me it's the whole understanding. So they are using a social and more psychological form of identity and population identifier and imposing it on to a biological process that in the data they provided they didn't actually test.

So, I agree with you is that if we're really going to start talking about biological markers, to make sure we're very clear on all the conclusions that we draw.

HALL: I'd like to go back to the question of nihilism as a construct and bring a perspective from psychiatry, which would suggest looking at nihilism not necessarily as a manifestation or correlate of major depression, but rather dysthymia, which might be a useful concept to think with, because it's gone through, in some ways, some of the mirrors of the discussions that have been going on.

Dysthymia, or minor depression, was typically dismissed as being purely characterological. These were dismissed as neurotic, high-strung individuals who couldn't be dealt with by medical professionals, and they were more or less dismissed until it was demonstrated through a lot of family studies, longitudinal studies, and therapeutic studies that actually they respond quite well to antidepressants. And if you look at the impact over the lifespan of these low grade depressions, not being incapacitated where they can't get out of bed, where they can't eat, where they can't maintain relationships, but more the sort of not really being able to enjoy life, being discouraged from taking initiatives, being discouraged from focusing on prosocial activities, just this constant nagging there's something wrong, which you would expect in someone living in a chronically stressful environment, actually has an economic and quality of life and health impact at least as high as major depression and matches up with some of my own work that I'm working through on issues of this mild depressive symptoms in Central Europe. There again, is a relatively disadvantaged population and where again you also see a lot of social comparison issues coming in, that as you might see in the United States, if African-Americans are becoming more mainstreamed or if their reference group is changing, you'd see this more just as if they were referring to themselves.

FISHER: So we are getting to the last five minutes, and we have three more comments, so I'd really like for us to get them in because this is really becoming a rich discussion, so if we can keep them brief; Next is Stacy.

LINDAU: So I actually am glad I followed you, Cage, because I think it picks up on this. I've been thinking about -- I actually think of the concept of nihilism as I understand it from conversations

with Cathy as very distinct from a concept like depression. And not to make silly use of language, but it seems to me there's a difference between depression which I understand largely as a biochemical consequence of brain neurotransmitters and oppression which seems to be much more related to the concept of nihilism. And I don't want to, in any way, be misunderstood in making a connection between this discussion about African-American youth and the conversation we had yesterday morning which was led by Dario Maestripieri who talked about his research with non-human primates. The relevance of this conversation to that one was Dario described settings -- he's very interested in doing population-based research with non-human primates where monkeys live freely and can be observed in a relatively free environment versus monkeys who are kept in cages and research facilities, often times socially isolated, removed, separated from their mothers, et cetera, and showed some very convincing evidence about the differences in biology you see or the kind of different answers you get to your questions about pathways if you ask the same question in a population of monkeys oppressed and living in cages versus living freely. And so in my mind I'm thinking about what can be learned in thinking about this or is there such a notion as a biology of oppression. I think the comments you made, Cage, are very, very interesting; but to me somehow dysthymia seems -- not to insult your work but seems trivial to what I hear you describing, Cathy, of the experiences of African-American youth. So those are my thoughts. I don't know if anyone else who plans to make a comment wants to respond to them.

HALL: I would just want to say that the issue of dysthymia is not at all trivialized in the situation at all and not to imply some sort of genetic determinism at all either, but rather to bring out, or enhance, the realization that social oppression and social stressors can become locked in in cognitive and physiological pathways or circuits that are very difficult to break out of without some sort of somatic intervention, particularly if they're continuing to raise the stressors on a daily basis. We often think that that locking-in pattern can only happen when you get major depression, where you can't get out of bed, but there's a lot of evidence that you get locked in lower, more chronic, insidious levels as well.

LINDAU: I don't misunderstand you.

FISHER: So we have two more. Lipton.

LIPTON: I just wanted to thank you all for bringing up these questions.

I wanted to bring up another concept, which I think again is more valuable than the issue of trying to define populations. I agree with Dr. Frader and Dr. Williams about the dangers of this, and to bring it back down to individuals, because now the problems I think of genetic studies is they promise to be able to identify individuals who may be more or less susceptible to the social influences that they experience. So there are some polymorphisms that have been identified that, for example, can render individuals more susceptible, if they happen to be in a highly discriminatory environment, where these polymorphisms have absolutely no effect in a less discriminatory environment. We all know the old grandmother's tales about how stress is good for you, that if you can get through this, then you can get through anything and so on and so forth. So I think exploring those issues in individuals; personally, it just seems like a more fruitful way to go, rather than to worry any more about defining continents of origin of our grandparents or great, great, great-grandparents and looking at populations that way.

FISHER: McClintock.

McCLINTOCK: Thank you. I have a question for the panel that I hope you can answer in one minute. You were talking about careful research, and I wanted to know, because I'm embarking on research in this area, what you mean by that. Secondly, what's bothered me about this whole

discussion is the implicit categorization of African-American. There is, I know, huge variation within the African-American population, community, whatever. I assume its borders are fuzzy, and in terms of careful and powerful research, would you recommend that we should invest in comparing variation within, rather than comparing between categories?

FISHER: If I could jump in, I wonder if we could make that rhetorical for the moment and address that at the end and have one more comment, because I know we're about to get into the next session and then wrap after that comment. So the panel, they'll address you personally.

HAUSER: I would just like to reemphasize the importance of the social structural determinants of the life chances of young African-Americans. What worries me about the notion of nihilism insofar as I like it as a structural concept, but I don't like it as a psychological construct. If you want to know why, go to your local used book store and buy a copy of Aldous Huxley's Brave New World. The somatization of perception of negative life chances frightens me more than anything else I've heard in this discussion.

FISHER: I'd like to thank everybody for engaging in this really interesting discussion, which benefited from the fact that we all come from different backgrounds, perspectives, and trainings. I think some of the take-homes are" we've clearly raised that we have very different experiences based on our skin color, and that means different things to different people. There were some topics that were brought up that we will likely continue in discussion during our break.

WILLIAMS: I have a few quick responses to your comment. I think it's a great one. Even the term African-American itself is problematic. If you ask most persons of African ancestry in the United States, African-American, although it's not the preferred term of academics, it's not the preferred term of most black people in America. They prefer the term black to African-American, just even terminology is problematic. I just want to give you an illustration of how dynamic the complexity is. With colleagues in the University of Michigan, we've just completed the largest study of black mental health in the United States that includes blacks of Caribbean ancestry. And what we have found, for example, just in terms of terminology for some Caribbean blacks, their preferred term is African-American. And for most African-Americans their preferred term is black. So it really becomes difficult even as we think about writing about these groups, there's no terms that we can use which is truly valid of what people's perceptions are and so on.

The other point is just to give you one illustration of a finding we have from this study so far, it's in press in the *American Journal of Public Health*, January of next year, but is that there is dramatic heterogeneity within the black population even in terms of health risk. So the finding I'm going to give you is we have found for rates of psychiatric disorders that black Caribbean women have lower rates of psychiatric disorders than African-American women, but black Caribbean men have higher rates of psychiatric disorders than African-American men and that the most at-risk group within our study is third generation Caribbean immigrants who are the single most disadvantaged group on virtually every indicator of health we look at. So it's just to illustrate there is dramatic variation within the population, and all of these terms we use have their limitations.

COHEN: I would agree with David. I think we understand that we have to use kind of multiple concepts and multiple terms to allow people to identify themselves. Let me say, of course, absolutely there's variation in African-American or black communities, but what I'm talking about here is not nihilism across those communities but nihilism as it exists among those who are most vulnerable or most marginalized, in particular among African-American youth who are impoverished. I'm talking about those who are impoverished, those who at a 50 percent rate do not finish high school, African males who a third of them will be in jail, on probation, on parole, that there is a subset of young black people who face what I would consider to be kind of chronic despair that we might, in fact,

want to pay attention to. In terms of careful research, I guess the thing I've learned from just chatting with Stacy over coffee is the need for collaborative work across disciplines and across perspectives, and I guess that's what I would be advocating here in terms of thinking about careful research.; It's dialogue. It's individuals, who think about not only the genetic connections between groups we might call race, but those of us in the social sciences who think about the social construction of race and how in fact it varies from historical period to historical period, and group to group. So to me, that is kind of careful research in the same way that I think David spoke, which is being very clear about the hypotheses of the mechanisms through which we believe these social conditions impact the body. I can't do that, but I can do that in conversation with Stacy.

LINDAU: Okay. Thank you. That was wonderful. (Applause.)

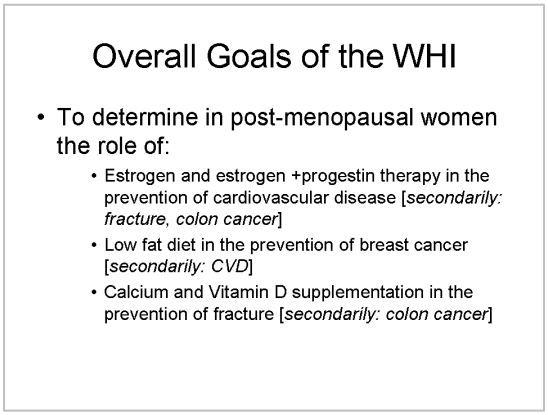
LINDAU: We are going to enter into our last scientific part of the session, and I am very pleased to thank and introduce Dr. Robert Wallace who is not only a participant in the conference here but is the chairman of the Advisory Board for the National Social Life, Health and Aging Project; and whose job it was today to introduce Richard Suzman and who now is pinch hitting for him. Dr. Wallace is a Professor of Epidemiology and Internal Medicine at the University of Iowa. He's co-director of the Center on Aging with the Carver College of Medicine; and he is an active researcher. He is heavily involved with Health and Retirement Study, and he's been very involved with the Women's Health Initiative Study among many, many others. He's widely published, and I know many of you know him already. So he has agreed to share with us a new perspective in relation to the other talks that we've had which is a large volunteer study, the Women's Health Initiative Study, and the biological measures that were collected in the context of that study as a possible source of data use and experimentation by individuals like those of us in the room. So with that, Dr. Wallace, I welcome you and thank you. (Applause.)

FROM POPULATION TO PUBLICATION: PIPELINE ISSUES IN ACCOMPLISHING THE NIH ROADMAP MISSION FOR INTEGRATED HEALTH RESEARCH

Biomarkers in a Large, Multi-Center, Multidisciplinary Cohort: Lessons from The Women's Health Initiative

Robert Wallace

WALLACE: Good morning. I'm going to continue with the thread of the presentations earlier this morning on biomarkers, and talk a little bit about the Women's Health Initiative (WHI) experience relevant to biomarkers, both because of the nature and the size of the study. I will limit my remarks and hopefully engage you in some discussion of the issues. I view the topic this morning as much about the social and behavioral aspects of investigators dealing with large data sets and a large volume of blood and DNA as it is with the scientific studies that have been conducted and the potential for future studies.



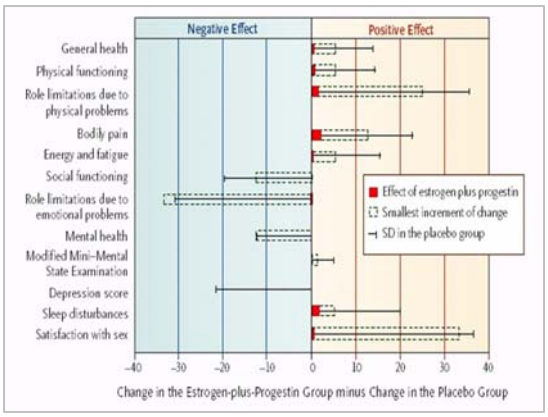
I appreciate that you may have varying levels of knowledge about the WHI. It is basically a set of three large preventive clinical trials that basically have three randomized interventions among post-menopausal women: 1) estrogen and estrogen plus progestin therapy; 2) a low fat diet trial; and 3) and the third was calcium and vitamin D supplementation. Additionally, there was a large observation study made of potential trial participants who either couldn't or wouldn't participate in the randomized interventions.

WHI was a long time in coming and was supported early on by the women's caucus in the Congress, and came to fruition in about 1992, with the strong support of Bernadine Healy, who was then director of NIH. The basic belief at the time was that adult women were not being studied sufficiently, relative to men, on the issues of chronic disease prevention, in general, and cardiovascular prevention, in particular. The slide shows the goals and outcomes of WHI. The hormone interventions were primarily an attempts to prevent cardiovascular disease, the low-fat diet to prevent breast cancer, and calcium and Vitamin D to prevent fractures. Those of you who know the outcomes of these trials, which have been published, know that when you do the research, they don't always come out the way it was thought.

I just want to give you some sense of the scope of the study and the sample size and just enough about the study design to follow my discussion. Basically, women were recruited first into the

hormone trial and into the dietary modification trial. Those that either felt they couldn't enter once having gone through a screening process or who were turned down because they had certain health or other characteristics were put in a cohort which we call the Observational Study. The randomized, calcium/Vitamin D study was superimposed on the other two trials; there was no additional recruitment of subjects. I'm not going to say too much about this any further except to tell you the total sample size after recruitment was about 161,000. WHI was performed at 40 sites in the United States and probably is the grandmother of all clinical trials as far as I can tell.

I want to just make a point that the cultural history of estrogen therapy goes back a very long way. It goes back to the '40s. For those of you that are interested in women's health, this is a fascinating historical journey. This slide shows an advertisement from the American Journal of Obstetrics and Gynecology from 1968. In a sense it's a sort of "polypill," containing two different drug classes—estrogen and a tranquillizer. But it's more than that. I think it represented a view of the attitude of the medical profession toward menopause and toward women and the nature of menopause.



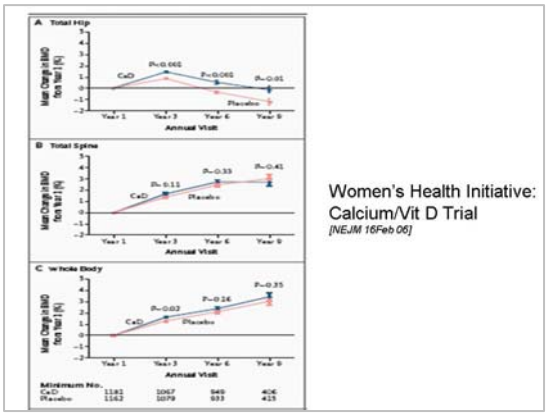
The next slide is a chart from the *New England Journal*, exploring the effect of estrogen plus progestin on measures of quality of life in that WHI trial. The findings are not important here; I want you to observe the scales and other measures on the left, and to let it serve as an invitation to participate with WHI investigators on issues related to social and behavioral factors in women's health. I believe this to be a very rich dataset on many issues related to the sociology and psychology of aging, and over time these data will be available to all for investigation. There are many geriatric and gerontological measures, including a standardized physical performance test. There are cognitive tests and dementia evaluations on subpopulations as well as a detailed hospital utilization experience. I hope you will participate with me and others on this project.

WHI is also a rich source of potential and actual biomarker studies. We collected blood at baseline and later as well, and separated out into cells, and some of which had its DNA extracted. The rest of it was saved as buffy coat, plasma and serum, and this has formed a large repository that is available, most efficiently through WHI investigators. The next slide shows examples of some of the studies of biomarkers that have been performed in WHI, including work in inflammation, nutrition, the clotting cascade, lipids and lipoproteins, measures of vascular function and endocrine measures of obesity. Many female sex hormones have been measured. In addition, many genetic studies have been performed. I hope this will show you something about the capacities of this study.

The WHI family is large and I hope will get larger as the specimens are exploited. Next I would like to discuss some of the general issues regarding the use of biomarkers in large studies. First, let me describe some of the burdens. With a large number of specimens, you have the problem of tracking the specimens--where they are, how much has been extracted and what's left. As you might guess, all of this requires a computerized database. Somebody has to pay for the storage and retrieval of these specimens, and in this case it is the National Heart, Lung and Blood Institute. There are complex logistics of obtaining stored specimens. Can they all be identified, and can they be retrieved in a timely manner? How do you obtain remnant specimens and send them back for storage or to another investigator? And all of those things need to be worked out.

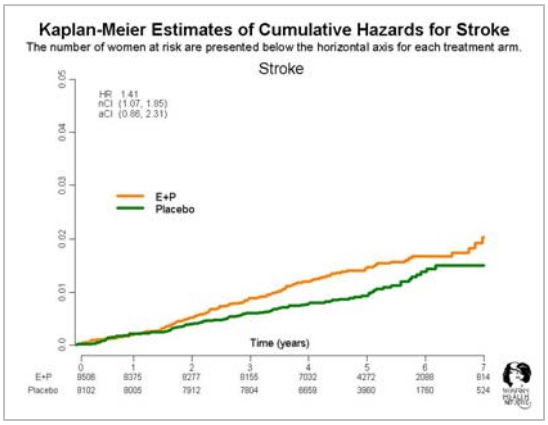
One needs to worry about unapproved uses of distributed specimens. If you give an investigator specimens to do a certain number of tasks that have been scientifically endorsed, what if they want to do that next gene or something appears in a journal the following day, and they want to try it, is this allowable? You must maintain rules of engagement about how the specimens are used and about returning them when the assays are completed. Another issue, on the bottom of the slide, is whether at some point the data from the laboratory determinations should be released to the public. This is a data-sharing issue, and as you know, all large NIH-funded projects require a data-sharing plan. WHI has received requests for data from persons outside the study, and the usual issues of maintaining confidentiality and allowing investigators to get their original studies published pertain.

As shown in the next slide, a very important issue is prioritizing the specimens for distribution among study applicants. First, we need criteria as to who is a suitable investigator to apply for the specimens. Right now that requires among other things affiliation with an existing WHI investigator. Of course, the scientific promise of the study is important. Some relevance to WHI themes may be an issue as well, as is the question of the nature of specimen oversight once they're released. Typically you have to be at an institution where there's an IRB and institutional responsibility for accepting specimens and doing the right thing. The funding issue is very important as well, because there is a delay between receiving approval from WHI and receipt of funding, if it comes at all. The question is whether we should hold those specimens for someone while the funding is being determined. The specimen process would be simplified if the applicants had funding already in hand.



Despite the large amount of work and the committee process to make priorities and dole out the specimens as the system allows, there are additional issues of study priorities. There may be special priorities that come from NIH and their consultative processes.

There may be some inquiries from industry or the private research community that place a call on specimens. There may also be some priorities dictated by historical and ethical concerns. An example, while it wasn't fully anticipated, the WHI hormone trials caused some serious adverse vascular effects, and the calcium/ Vitamin D trial caused an increase in renal stones. So there were ethical as well as scientific dimensions to use some of the specimens for elucidating the mechanisms of the adverse effects.



As an example of an adverse effect, the next slide shows the stroke outcome from the estrogen/ progestin intervention. The trial was stopped at about five years because of this and other adverse effects, and then the difference between the curves started to narrow a little bit. Some specimens were devoted to trying to explain the findings shown here, as well as others like it.

Another use of the specimens has been to furnish them to private companies, and we've had several requests from such organizations. We were approached by Perlegen Sciences, a

pharmacogenomic company that wanted to look for genes and SNPs that might be related to the breast cancer and vascular outcomes we found in WHI. After substantial scientific consultation and a vote by all of the site investigators, we decided to give them specimens.

An Industrial Affiliation: Genome-wide Scan - I

Approached by biotechnology industry to use WHI to validate techniques:

Stage I – 1000 cases and 1000 controls in the cohort study for each of 3 conditions:
-Coronary heart disease
-Stroke
-Breast cancer
Pools of 125 participants
Perlegen's 360K tag-SNPs is completed.

Genome-wide Scan - II

Stage II
-For each disease 9000 SNPs selected empirically
-1000 SNPs in candidate genes added.
-Specialized 10K chips for each disease have been designed.
-Individual-level testing of 10K SNPs

Genome-wide Scan - III

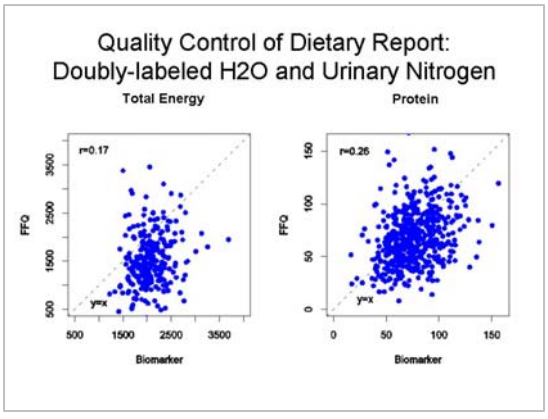
Stage III
-Individual testing of 335 CHD, 258 stroke, and 349 breast cancer cases and matched controls in hormone trial)

[Recent inquiry: Using WHI for Epigenome studies]

There were some special ethical issues as well, related to maintaining the anonymity of specimen donors and the fact that we never consented the population originally to share the specimens with a private company. All of this has been worked out.

They are using high-throughput techniques, including chip technology, and are funding this work themselves. This march of science and technology continues. Industrial biotechnology affiliation is something that I'm hoping will work out for mutual benefit, although there are many issues to consider and adjudicate. In the end, however, it will open up an investigative capacity that otherwise we wouldn't have.

In addition to collaborative partnering with other academics and companies, some of the best collaborative activities may be within NIH, and one we are prosing to partner with is "CGEMS" from the National Cancer Institute. The website is shown on this slide. This is a collaborative study exploring genetic markers for cancer. They want to do about 500,000 SNPs across the genome and follow them up with replicate studies. I think that the number of SNPs is not important here, but only that some very important things can come from this. The point is that collaborations among the studies is very important, and you should look for them if you have population studies with appropriate biomarkers. They are out there and there are many collaboratives that are looking for partners. For genetics, a sample of 1,000-3,000 people isn't enough. but it can be pooled with other studies to good end.



Another use of the biomarkers has been to validate questionnaire responses. An example might be for dietary questionnaires. They are always a little bit suspect and while they "work," they can have lots of problems. Perhaps one of the hardest things to measure quantitatively by diet questionnaire is total energy intake, which is so central to studies of obesity and many other conditions. WHI is conducting a doubly-labeled water study to obtain an objective biomarker of energy intake. We were also interested in protein intake, in verifying protein intake, and in this case we are comparing dietary

reports with urinary creatine levels, again for purposes of validation. The point is that biomarkers can also be used in certain instances to validate the other kinds of information. There have been calls while we were doing this study worrying if we could get it done.

In summary, the WHI has several strengths with respect to studying biomarkers, which are: are its size, its broad age range among elderly women, a large representation of minority participants, attention to the detail of questionnaire information on a wide range of health and behavioral topics, detailed information on health outcomes, the availability of blood for biomarker studies and a large number of investigators from many institutions who can bring ideas from many disciplines. The necessary rigor and documentation that comes with conducting clinical trials yields a high level of data-cleansing and accuracy.

However, there are several potential weaknesses as well. It's a volunteer cohort, as Stacy noted; only women are represented, and there are difficult consent issues. We have a large observational cohort, but it's a very selected cohort and at least at baseline not very representative of the general population. There are many logistical difficulties in conducting the study for many reasons, and there are 42 institutions, 42 institutional IRBs and many investigators who must weigh in on all the issues. The specimens are already dwindling, particularly from participants who developed heart disease and cancer because that's where there has been the most scientific interest in this study. For other types of participants, there is still a good supply of blood specimens for new ideas related to biomarkers.

There's so much more that could be said about all of this, and so these are the people. As have some of the others, I might show you a picture of my lab, but I don't have enough time to show you all of the investigators and project staff and the 161,000 participants who gave so much of their time and lives to this study. Thank you. (Applause.)

LINDAU: I think we can take five minutes for questions, if there are any.

NIELSEN: Just a quick comment to bring in again a bit of that discussion from the National Academy's workshop last week. Richard Suzman and several other people highlighted the fact that in these rich databases - perhaps I think someone threw out the number 17 percent of the data has been analyzed to its full potential, and then someone had a much more pessimistic estimate of around 7 percent - and I think that you're presenting how the opportunities exist for getting involved in this study is really important for this group to be aware of, and it also exists for many of the other large-scale studies where data-sharing is part of the agreement when NIH in particular is funding this kind of research. So I would encourage anyone who is considering launching into this field of biomarker investigation to really take advantage of what's already been done out there and is in some ways underexploited and ready to be utilized. Thanks for putting that in there.

WALLACE: Thank you. That is the message.

MILLS: I want to ask, you mentioned in data these approximately 17 percent are minority, and in terms of health-disparities research can you talk a little bit about the categories, classifications of minorities of the sample in which you have adequate Latino/Hispanic representation - within group comparisons, Puerto Ricans, Central Americans, so forth so on - or is it just general?

WALLACE: I'm sorry I didn't bring the slides and hadn't anticipated that question, but I can tell you there are many, a few thousand African-Americans, a few thousand Hispanics, and a broad range of others racial and ethnic groups. One of the study sites emphasized Native Americans, and there are a substantial number of Native American women as well, a number of Asians, and so there is an ample opportunity to conduct disparities research. Please note, however, that our participants are volunteers, and not population referent. If you're looking for geographic representation, it certainly

would never be that, and I would never claim it. On the other hand, Stacy and I and others discussed earlier whether these participants become more like a general population over time. That would be my own speculation, although I could not prove it. Over time it might in fact be possible to explore this cohort of women who are pretty healthy at the beginning, but then start to get their diseases at different rates. These differential rates can then be related to the wealth of risk factor information in place.

WOLF: So you had mentioned a little bit about the Perlegen studies where they were proposing to do very extensive genotyping on roughly 2,000 individuals. Can you comment on your experience in reconsenting those individuals and how they reacted, whether they sort of viewed this as a possibility based on their original consent, and go into a little bit more detail about that?

WALLACE: Just to be clear, in the end we didn't re-consent them for these genetic studies because we didn't think the Perlegen activities were different from the general nature of the genetic studies which we had explained in the original consent. And we don't do it for a lot of the other candidate gene and SNPs and other associational studies that are being performed. More recently, however, NIH, with good reason started to get worried that as this brave new world of determinations comes about, that we were behind in our consent, and it was somewhat outdated by modern standards. So we went back to re-consent the survivors to reconfirm participation in the long term follow-up and for future, yet unspecified biomarker studies. The consent rate for subsequent genetic studies was close to 80 percent, which was lower than the consent rate for the extension study as a whole. The extension study is a death and disease outcome surveillance activity to determine the long term effects of our interventions. So some women didn't want any more genetics studies, and there's probably a study lurking there about women who had originally agreed to it but now refused.

WALLACE: Again. Thank you. (Applause).

CLOSING

Stacy Tessler Lindau

LINDAU: Okay. Well, thank you all. I first and foremost want to thank the participants, especially those of you who had the stamina to stick it out for a day and-a-half and through a really broad variety of presentations, speakers and ideas. This is the fourth year, as you know, that we've done the workshop. Some of you have even stuck it out for four years, you've been here every year; and some of you are new this year.

And that's the next group I'd like to thank. The National Institute on Aging funded the conference this year; and hopefully if there's money in the budget next year, we will be funded again. This is a collaborative grant in conjunction with Lis Nielsen who was a tremendous collaborator and resource for us. Georgeanne Patmios also gave us a good deal of advice as we were going down the path of planning for the workshop. So I want to thank you two, and I've enjoyed the collaboration very much. I thanked Alicia Frasier last night and yesterday. I can't thank her enough. She joined the Biomarker Core this year and just dove right in and made the workshop happen at a level even better than the year before; and every year I think we increase in our performance and our accomplishments. And so thank you, Alicia, and also to the rest of the Biomarker Core staff, which includes Natalia and Precious, who has been an ad hoc member of our staff; and Amelia, who also stuck it out with us and helped with the microphones. I also, of course, want to acknowledge the Center on Aging, which is led by Linda Waite and Diane Lauderdale and is housed at NORC. And the individuals at NORC who have been very, very supportive, including Kathleen Parks and Adelle Hinojosa, who helped with a lot of your travel details. If I have forgotten to thank anyone, my sincere apologies, and I do appreciate everybody who came, participated. We had phenomenal speakers. The thing they all had in common is they all felt like fish out of water, which means that the workshop has accomplished exactly what I hope it does by bringing together people who wouldn't otherwise see each other, converse with each other, share ideas, and to me that's what motivates this workshop.

For 2007, we will plan to hold another workshop, here in Chicago at this time. I think it will be a day and-a-half; although since it's intended to be a global workshop - bringing together major population-based studies from around the world that are incorporating both rigorous sociological measures and biological measures, particularly studies that are looking at aging populations - I'm wondering whether we maybe should do two days because people are going to be coming from a long ways away. So if you have any thoughts, about that, please let us know on the evaluation forms or offer an alternative.

We've covered a great deal of territory. I was so excited this year that we could include -- actually we kicked off the workshop with a discussion coming from research around non-human primates and how we might use the approaches of researchers who study monkeys to think about how we can make observations about human beings beyond self-report, and I found those conversations, the ones that followed, fascinating. We talked about chemicals of all kinds, social chemosignals, what kind of messages we send from our body odors and other external characteristics. We talked about chemosignals that go from the brain to the adrenal glands and to the ovaries and throughout the body when we talked about physiological measures that we could use either as substitutes or

complements to the more chemical measures. And we talked a lot about genetics, and I really appreciate Thom McDade's comments about that. The precision of language not just around genetics but around all of the things we discussed is so important for this workshop because we do come from different disciplines and the terms of art that are comfortable within our own circle become sources of confusion when we're in a very interdisciplinary setting. So my perspective on the genetics, the language around genetics, is that really I believe almost everything we're measuring on the biological side comes from genetics at some level. It comes from expression of our most basic biological material whether we're looking at the most downstream expression of that or we're really trying to measure the amino acids themselves, these things are all very much interrelated.

So, with that, I promised we'd finish before noon. Unless there are any questions at this time, then I want to offer you all my gratitude and congratulations and a safe trip home, wherever that may be. You are now excused. Thank you.