



Taking on the **EXPOSOME**

*Bringing Bioinformatics Tools to the
Environmental Side of the Health Equation*

BY KRISTIN SAINANI, PhD





When it comes to what kills people, Nurture trumps Nature: Chronic diseases with overwhelmingly environmental (rather than genetic) causes are responsible for the deaths of two-thirds of the world's population. Yet the investment made in unraveling the environmental side of the health equation pales by comparison to the investment in human genome research.

"In the past 20 years, a lot of effort and funding have pointed toward genome research," says **Paolo Vineis, PhD**, professor of medicine and chair of environmental epidemiology at Imperial College London. "Now, people are suggesting that a similar effort should be put into exposure research, and also that exposures should be investigated systematically as has been done for the genome, such as with Genome-Wide Association Studies, or GWAS."

Though we know some of the biggest players in chronic diseases—air pollution, smoking, poor diet, and lack of exercise—an estimated 50 percent of the environmental drivers remain unknown. "I'm not going to argue that diet or physical activity or smoking don't have a role to play," says **Chirag Patel, PhD**, assistant professor of biomedical informatics at Harvard University. "But it behooves us to explain more of the variation than can be explained by classical environmental factors. We need to look beyond the proverbial lamppost."

Environment-disease research suffers from the same problems that gene-disease research did 20 years ago: Individual labs study hand-picked risk factors one at a time in small studies with inconsistent methodologies; and they are incentivized to report positive findings. The result: a literature rife with spurious findings. "There's a now-famous number being punted around in genetic epidemiology that, prior to GWAS, over 95 percent of the findings from candidate gene studies—that is, your favorite gene in connection with a trait—are false," Patel says. In a 2011 review, researchers found that only 13 of 1,151 purported loci-phenotype associations for eight conditions were replicated in large-scale studies. It took GWAS and related approaches—which consider a multitude of

genes simultaneously in an unbiased, standardized way—to clean up this literature.

We need a similar revolution in the study of environment-disease associations, Patel and others say. In 2005, **Christopher Wild, PhD**, now director of the International Agency on Cancer Research, coined the term “exposome” as a call for high-throughput, systematic approaches to studying how the environment impacts health. Echoing this call, Patel and others coined the term EWAS, or Environment-Wide Association Study, to encourage researchers to apply GWAS-like methods to study health-environment associations.

The exposome encompasses the entirety of a person’s exposures from birth to death. Thus, the first challenge is how to measure it. Fortunately, technological advances are making it possible to measure the exposome at higher resolutions and on larger scales than ever before. Metabolomics measures the chemical ghosts of exposures in our blood; wearable sensors and smartphones track where we go, what we breathe and eat, how we move, and how we feel; social media sites amass records of our moods and social connections; electronic health records store our clinical, personal, and demographic attributes; and geographical information systems and survey data reveal the wider societal factors that influence our health.

The sheer volume and complexity of these data are overwhelming. According to **Gary Miller, PhD**, professor of environmental health at Emory University in Atlanta, Georgia, a geneticist on his staff once commented that after he saw how complicated the exposure data were, she felt like “a wimp” for studying genetics. Whereas genomic data consist of stable linear sequences, exposome data are heterogeneous, non-linear variables that change over time and space. Dense webs of correlation among environmental variables make it hard to tease out causation. And, due to the highly personal nature of the data, privacy and security concerns abound. Exposome researchers can draw heavily on the bioinformatics tools developed for GWAS, but to fully realize the promise of the exposome, they will need new tools for storing, integrating, and analyzing the data.

“It’s daunting. It’s definitely hard,”

Miller says. But it’s also an opportunity for bioinformaticians and computational biologists, he adds. “For people who like wrangling with data, the exposome offers some great challenges.” This article reviews recent progress in exposome research and the challenges that remain for studying everything from the chemicals in our bodies to the quality of our neighborhoods.

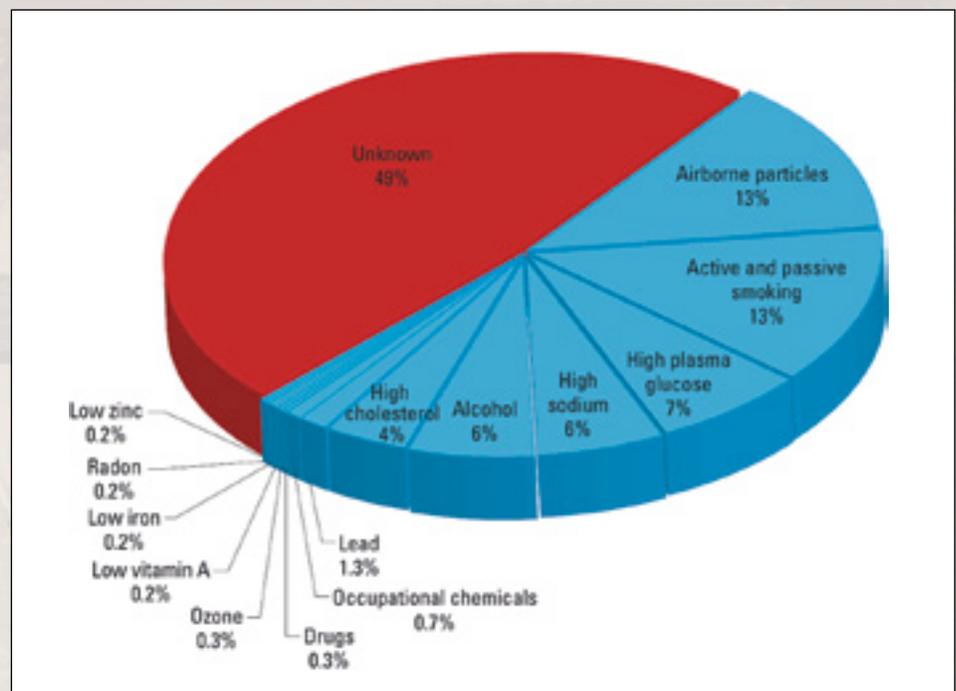
WHAT’S INSIDE: METABOLOMICS

External exposures leave chemical traces in our bodies. These can provide a convenient window into how those exposures affect health. “Exposures are inherently chemical in nature,” says **Stephen M. Rappaport, PhD**, professor of environmental health sciences at the University of California, Berkeley. “Anything that causes a health effect is either a chemical or is mediated through chemicals.” Food, drugs, and pollutants leave behind metals and small molecules in the blood. “Even psychosocial stress produces hormones and other biologically relevant molecules in the body,” Rappaport says.

Fortunately, researchers who want to

perform large-scale exposome studies can access troves of specimens and associated health outcome data that have been collected and archived by epidemiologic studies and national surveys. In 2010, when Patel was a doctoral student at Stanford, he and his mentors performed the first proof-of-principle EWAS using publicly available data from the National Health and Nutritional Examination Survey (NHANES), which includes data on chemicals in the blood and urine of thousands of participants. When they compared 266 chemicals across participants with and without type 2 diabetes, they turned up four hits: the pollutants polychlorinated biphenyls (PCBs) and heptachlor epoxide and the nutrients vitamin E and beta-carotene (the latter was inversely associated with diabetes). Follow-up studies are needed to determine if any of these factors is causally related to diabetes, Patel stresses. “But by taking a data-driven, agnostic, unbiased approach, EWAS leads to a more reproducible list of hypotheses to prioritize for further study.”

Rappaport concurs: “All we want to do with EWAS is to sort through the thousands of chemicals to which people



INTO THE UNKNOWN: Though we know many of the environmental risk factors for chronic diseases, about half remain unknown. This chart shows the percent of total global chronic disease deaths that are believed to be explained by each factor, according to 2010 data from the World Health Organization. Reprinted from Rappaport SM, Barupal DK, Wishart D, Vineis P, Scalbert A. 2014. The blood exposome and its role in discovering causes of disease. *Environ Health Perspect* 122:769-774.

are exposed during life and identify those few exposures that may be causes of disease. Then epidemiologists can follow up with focused studies to establish causality. Thus, the exposome paradigm begins with a data-driven EWAS to generate hypotheses and ends with tests of these hypotheses in subsequent stages.”

Patel’s team had developed publicly available software for EWAS (<http://www.chiragjgroup.org/exposome-analytics-course/>) that combines off-the-shelf GWAS tools with cutting-edge machine-learning techniques. “There’s nothing novel in the methods. Rather, we are taking existing methods that statisticians and informaticians have developed for different domains and introducing them to people doing exposure science and epidemiology,” Patel says.

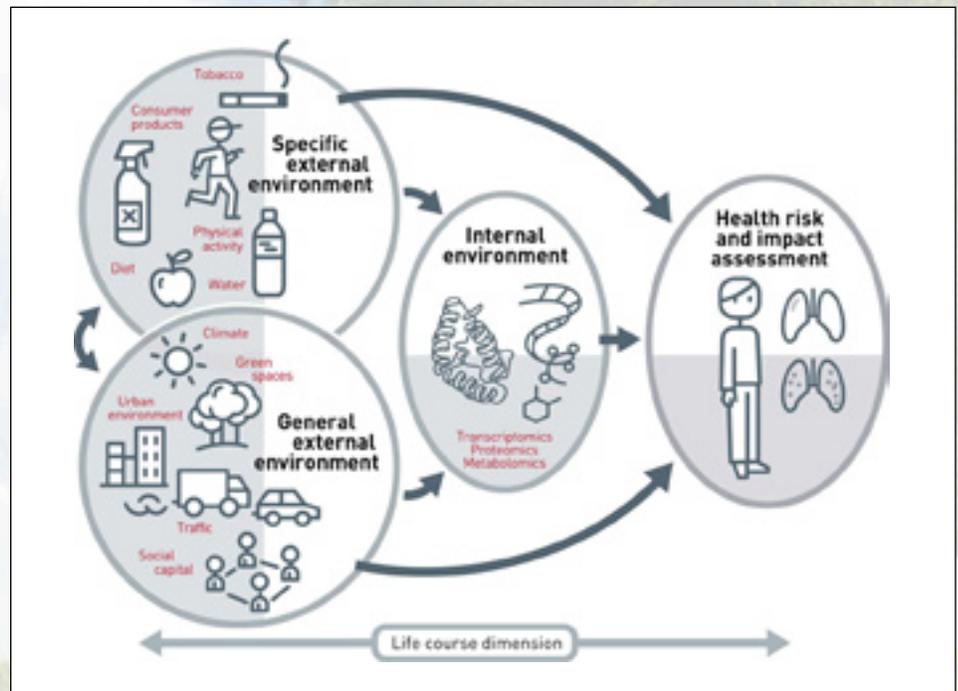
Exposome researchers dream of a day when there is a cost-effective exposome chip akin to the SNP (single nucleotide polymorphism) chips that enabled GWAS studies. “If you could measure even 500 chemicals consistently in human plasma, and you could do it in a cost-effective way at the scale of a GWAS, you would start finding things,” Miller says.

To look for novel triggers of disease, many exposome researchers are also widening their search beyond known chemical markers. They are turning to untargeted metabolomics—using mass spectrometry to explore the vast landscape of unknown chemicals in the blood. Platforms can now measure 100,000 small molecules from a few microliters of blood in 20 minutes, Rappaport says. The catch: Mass spectrometry just gives signatures of chemicals, or spectral peaks; so, once researchers have fished out the most interesting peaks, they still need to work out the identity of the chemicals. Spectral reference libraries exist, but they cover only a small fraction of the metabolome, so chemical identification remains a challenge.

Rappaport’s lab is nevertheless taking this approach. To ensure they are picking up causes rather than effects of disease, they use archived samples from cohorts of people who were healthy at the time of the blood draw. For example, to look for clues to childhood leukemia, Rappaport’s team is using neonatal blood spots

collected on all babies born in California since the mid-1980s. By comparing the metabolomic profiles of 1,000 babies who later developed childhood leukemia with those of 1000 comparable controls, they hope to identify possible pre-natal causes of leukemia. They are also looking for evidence of exposure to damaging reactive molecules by measuring telltale alterations of the blood protein serum albumin (called adductomics). “Adducts from albumin are interesting because they stick around for

eggs—bacteria in our guts convert the fat into trimethylamine N-oxide, or TMAO. Animal studies showed that TMAO clogs arteries. And subsequent human studies showed that individuals with high levels of TMAO are 2.5 times more likely to have major cardiovascular events (heart attack, stroke, or death) than those with low levels. The American Heart Association and American Stroke Association listed TMAO as one of the top 10 advances in heart disease and stroke science for 2013.



THE EXPOSOME: The exposome encompasses the entirety of a person’s exposures from birth to death, including internal exposures such as gut bacteria, lifestyle choices such as smoking, and social determinants such as poverty. Reprinted from M Vrijhied, *The exposome: a new paradigm to study the impact of environment on health*, *Thorax* 69:876-878 (2014) with permission from BMJ Publishing Group Ltd.

a month. So we’ll get a picture of what babies were exposed to during the month preceding delivery,” Rappaport says.

It’s too early to know what Rappaport’s study will turn up. But the power of the metabolomic approach is illustrated by a series of studies from the Cleveland Clinic, including 2010 and 2013 papers in *Nature* and the *New England Journal of Medicine*, respectively. Researchers compared stored blood samples from 150 people who developed a heart attack or stroke with 150 age and gender-matched controls. Following up on the strongest signals from mass spectrometry, they uncovered a key metabolic pathway: When we eat lecithin—a fatty acid found in meat and

“If their hypothesis is correct, I think we’re going to see that this has a major impact on how people diagnose and treat heart disease in the future,” Rappaport says.

Success stories like this have been limited, however, due to the lack of informatics infrastructure. Exposome initiatives in Europe and the United States are building infrastructure such as spectral reference libraries, shared data platforms, and analysis tools. For example, Vineis leads a consortium of 12 European institutions, called EXPOsOMICS (<http://www.exposomicsproject.eu/>), while Miller leads The Emory Health and Exposome Research Center: Understanding Lifetime Exposures (HERCULES, <http://emoryhercules.com/>).

HERCULES hosted the first-ever exposome course this past June, which trained diverse researchers to collect, integrate, and analyze metabolomics and other omics data.

WHERE WE GO AND WHAT WE DO: PERSONAL SENSORS

Internal markers provide clues to the exposome, but they are still several steps removed from the exposures themselves. “We’re trying to prevent that exposure in the first place. If you wait until it’s in the body, it’s too late to see where it occurred and where one could intervene,” says **Jacqueline Kerr, PhD**, associate professor of family medicine and public health at the University of California, San Diego. Internal markers also capture just a single moment in time. In contrast, wearable sensors allow exposome researchers to quantify external exposures with unprecedented precision, and to pinpoint where and when they occur.

For example, air pollution can be crudely estimated from a person’s home address—by referencing data from local air monitoring stations. But two people who live in the same vicinity may be exposed to disparate pollution levels due to differences in their indoor environments, places of work, and modes of transportation. “All these studies are being done on people’s home addresses. But where we live is not what we’re exposed to,” Kerr says. Wearable air pollution sensors offer a minute-to-minute accounting.

To illustrate the importance of individual-level monitoring, **Geoffrey Jacquez, PhD**, professor of geography at the State University of New York at Buffalo, points to a study in which researchers outfitted children with personal air pollution monitors. There was a surprising spike in pollution levels at the end of each school day—it turns out that children sitting on idling school buses were breathing in large amounts of exhaust. From this realization, policy makers came up with an easy solution: Close the doors on idling buses. Stationary sensors on the tops of buildings could not have detected this health threat.

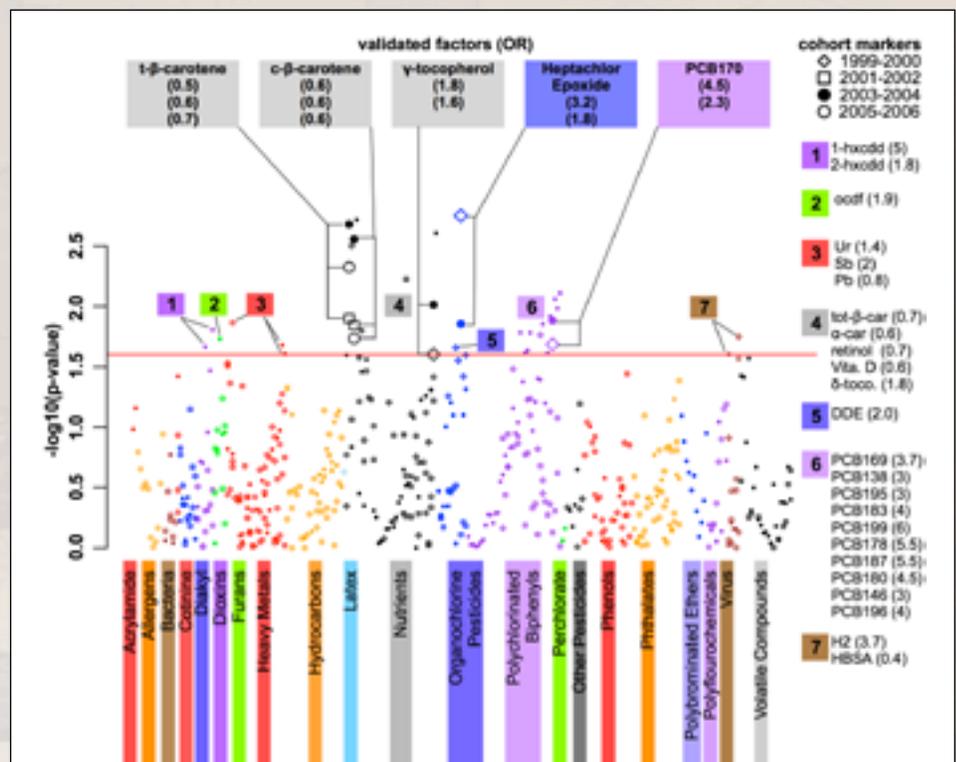
Personal sensors can also measure UV

light, humidity, temperature, and noise. But most sensors remain too bulky and costly to deploy on the thousands of participants needed for EWAS-type studies. For example, one of the largest studies to deploy personal monitors for air pollution is EXPOsOMICS, which involved just a few hundred volunteers wearing backpacks equipped with ultra-fine particle sensors. But because the EXPOsOMICS volunteers were sampled from other large cohort studies in Europe, Vineis’ team was able to leverage data for the smaller subsample (age, county of residence and job, for example), to predict the air pollution exposures of the larger group.

GPS technology can also provide detailed exposure profiles. GPS-enabled

University in the United Kingdom. People’s spatial-temporal paths (also called “space-time cubes”) can be intersected with spatial-temporal maps of environmental hazards—such as particular pollutants, radon, or even the density of liquor stores or fast food restaurants—to quantify individual exposures, he says.

Besides the physical environment, smartphones and personal monitors also measure individual behaviors, such as sleep, exercise, and diet. Jacquez and Sabel coined the term “behavome” to draw attention to these factors, which are at least partly in our control. Accelerometers count steps and sleep times; heart rate monitors gauge exercise intensity; smartphone cameras snap photographs



SCANNING THE EXPOSOME: Environment-Wide Association Studies (EWAS) systematically scan the exposome for environment-disease links in the same way that Genome-Wide Association Studies (GWAS) systematically scan the genome for gene-disease links. Here, a Manhattan plot reveals significant associations between type 2 diabetes and certain pollutants (PCBs and dioxins, for example), as well as the nutrient γ -tocopherol (vitamin E). The nutrient beta-carotene was inversely related to the disease (i.e., seems to protect against it). Reprinted from Patel CJ, Bhattacharya J, Butte AJ (2010) An Environment-Wide Association Study (EWAS) on Type 2 Diabetes Mellitus. PLoS ONE 5(5): e10746.

smartphones can track exactly when and where a person travels throughout the day. “It’s only quite recently that the technology has been good enough that we can do that with some confidence,” says **Clive Sabel, PhD**, professor of quantitative geography at Bristol

University in the United Kingdom. People’s spatial-temporal paths (also called “space-time cubes”) can be intersected with spatial-temporal maps of environmental hazards—such as particular pollutants, radon, or even the density of liquor stores or fast food restaurants—to quantify individual exposures, he says.

NHANES has collected accelerometer data on thousands of participants since

2003. But large-scale exposome studies using behavior trackers remain rare. Since many technologies have only become available recently, scientists are still testing their usability and accuracy. “We’ve spent so much time investigating the reliability of the devices,” Kerr says. Researchers are also grappling with how to deal with the quantity of data. NHANES has seven terabytes worth of accelerometer data, including 150 million data points per person. Besides issues of storage, it’s unclear how to process such data. How do we extract meaning out of 150 million data points—do we look at averages, slopes, standard deviations, or more complicated statistical measures? Two of NIH’s Big Data to Knowledge (BD2K) centers—The Mobilize Center at Stanford and Mobile Sensor Data-to-Knowledge (MD2K) center—are grappling directly with this issue (See *BCR* story: “Wearing Your Health on Your Sleeve”). Privacy is another concern. Kerr outfits study participants with personal cameras, which end up photographing people who are not involved in the study. “Because we have that type of information, we have to handle it in a very secure way. We have to be very careful about our ethical framework,” Kerr says.

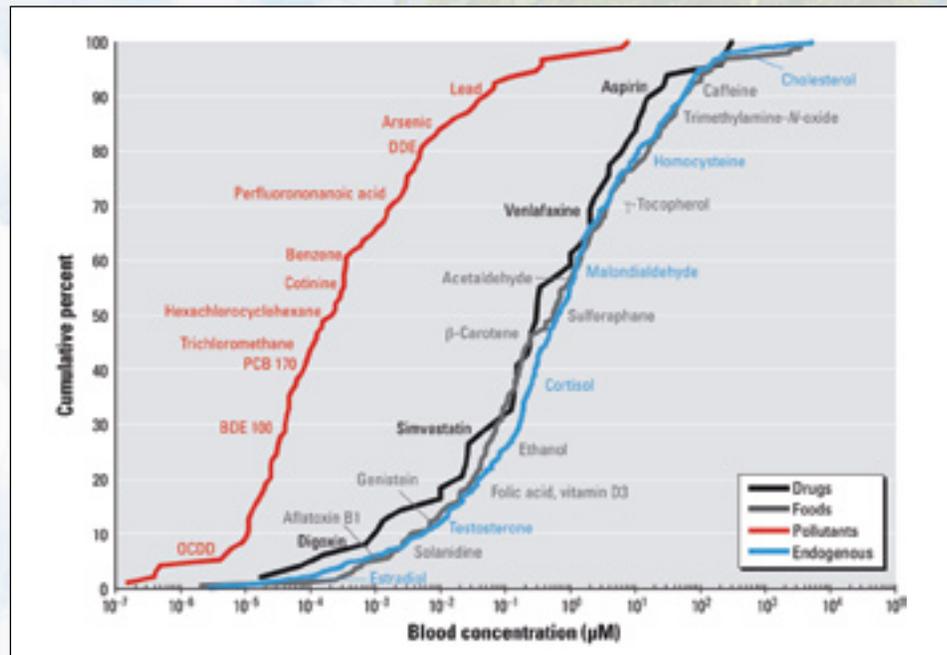
Exposome researchers are also hoping to tap into the massive amounts of personal health data being collected outside of mainstream research. Twenty percent of Americans own a health wearable, such as a fitness band or smartwatch. If just a small fraction is willing to share these data, this translates to huge sample sizes. Many challenges in using and accessing these data remain, however. For one thing, people who are willing to share their data tend to be very different from the average American. “We’ve looked at typical journeys that you might be able to get from Strava, the GPS-based biking system. And they look nothing like the typical journeys that we get in our study participants,” Kerr says. “The data probably don’t represent a lot of the underserved groups that we’re trying to reach.”

Also, the commercial companies that own the data are often unwilling to share, Jacquez says. He hopes to see more “benefit corporations,” or “B-corporations” set up to sell these devices. B-corporations

blend traditional for-profit and non-profit business models—they make money, but are also committed to serving society. Such companies could make user-generated data

predicted mortality as well as high cholesterol and high blood pressure.

Constructs such as stress and social isolation may seem “squishy” and hard



*EXPOSURES IN OUR BLOOD: Summary of small molecules and metals in human blood. Each curve represents the cumulative distribution of chemical concentrations from a particular source. Concentrations of drugs, foods, and endogenous chemicals are several orders of magnitude higher than concentrations of pollutants. Reprinted from Rappaport SM, Barupal DK, Wishart D, Vineis P, Scalbert A. 2014. The blood exposome and its role in discovering causes of disease. *Environ Health Perspect* 122:769-774.*

freely available to research scientists. “This would be a model for people sharing their data for the greater good,” Jacquez says.

HOW WE FEEL AND RELATE: ELECTRONIC FOOTPRINTS

The exposome encompasses a wider set of psychological, social, and behavioral variables that include stress, subjective well-being, personality traits, resilience, social connectedness, and social support. It would be a mistake to neglect these risk factors, says **Nancy Adler, PhD**, professor of medical psychology at the University of California, San Francisco. “The physical environment is concrete and it is related to health, but the effect sizes are small. The associations for some of the social and behavioral variables are actually more powerful.” In one study, her team showed that social isolation

to pin down, but we have well-validated instruments for measuring them from social science and psychology. “We know what the factors are, and we know how to measure them with self-report,” says **Elissa Epel, PhD**, professor of psychiatry at the University of California, San Francisco. The ability to measure these constructs electronically—via mobile phones, social media, and electronic health records—opens the door for their widespread inclusion in exposome research.

Smartphones can measure stress and other emotional states and behaviors in real-time. In Ecological Momentary Assessment (EMA), people are randomly pinged throughout the day and asked questions such as: What’s your mood? How stressed are you? Who are you talking to? Do you have a craving for food? Did you overeat? “We can characterize people in their natural environment in a fresher, closer way to their actual experience,” Epel says.

EMA gives a much richer set of

data than could be obtained from a few questions on a survey. But it also presents challenges for data analysts. “We’re good at collecting masses of data and we haven’t caught up to being able to use it well and create meaning out of it,” Epel says. “We’re in need of data scientists who can manage and make sense of these data. It is a hot new area that we need to be training more scientists in.”

Others are gathering data from social media sites. Sabel uses Twitter to study emotions, for example. People’s tweets objectively reveal their moods, Sabel says. “The idea of mining data from

Many large epidemiologic surveys also include stress-related variables. For example, the Health and Retirement Study—which has been following 20,000 older adults in the United States for nearly a quarter-century—has periodically queried participants about socioeconomic stressors, such as unemployment and financial hardship. Participants also filled out a one-time survey in 2004 that asked about their exposure to stressful life events—such as divorce, loss, or trauma—in both childhood and adulthood. Using an EWAS approach, **Eli Puterman, PhD**, assistant professor of psychiatry at the

University of California, San Francisco, is asking which of 92 variables available in the Health and Retirement Study is most strongly linked to mortality. “I think what’s really exciting about it is that we’re allowing the data to speak for themselves,” Puterman says.

Epel co-leads the Stress Measurement Network, a consortium that aims to deploy more and better measurements of stress in large epidemiologic studies. In particular, more subjective measures of stress are needed, Epel says. “You cannot know how someone

is feeling unless you ask them. That’s one case where we absolutely need self-report.”

Beyond epidemiologic studies, electronic health records (EHRs) offer a huge opportunity for exposome researchers. “If we had interoperable EHR records that had these data in them, we could really start to study the exposome,” Adler says. She participated in an Institute of Medicine panel tasked with recommending social and behavioral measures for inclusion into EHRs. The panel devised an 11-item battery that included one or two questions each on smoking, physical activity, education, race/ethnicity, and home address, as well as four questions on social connection and isolation.

Getting health care providers to implement the battery is challenging, but Adler notes that doctors are increasingly being held accountable for patient outcomes. “Once doctors are on the hook for keeping

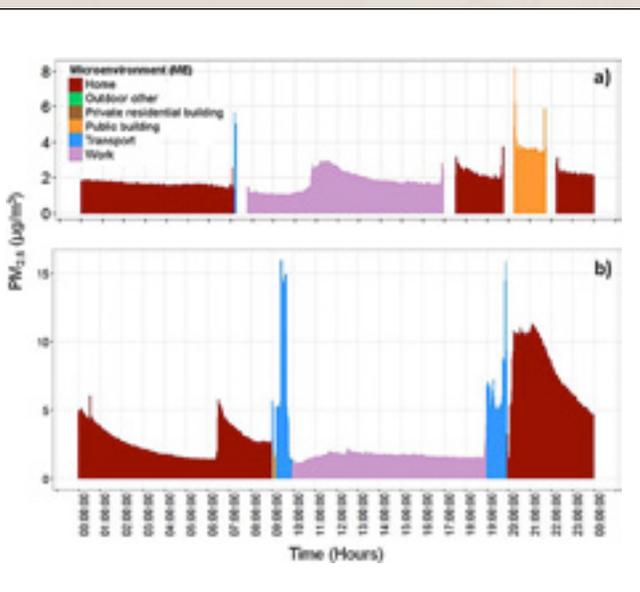
people well, they start to pay much more attention to the things that really drive their health, many of which are social,” she says.

HOW THE DECK IS STACKED: GEOGRAPHICAL INFORMATION SYSTEMS

Many factors that influence our health operate at the societal rather than individual level: what culture we come from, whether we live in poverty, whether we have access to health care and high-quality education. “There’s a bit of a paradigm shift to say behavior is not just an individual choice. It’s also constrained by the social environment this person is in and their financial resources,” Adler says. To get at these macro-level factors, some researchers are using geographical information systems. “Geo-coding is really opening up possibilities of linking what’s going on in neighborhoods and communities to disease outcomes,” Adler says.

For example, **Paul Juarez, PhD**, professor of family and community medicine at Meharry Medical College in Nashville, Tennessee, uses mapping technology to study health disparities. Juarez and his team created the Public Health Exposome Database, which contains 15,000 data points on each of 3,100 counties in the U.S.—including data on water and air pollution; availability of sidewalks and grocery stores; education and poverty; local, state, and federal laws pertinent to health; and health outcomes. “With county level data, you can do some great maps and show the hotspots and patterns,” Juarez says. “People understand maps better they do spreadsheets.”

To analyze the data, “we’ve had to go out and recruit people who have big data skill sets,” Juarez says. For example, he collaborates with **Michael A. Langston, PhD**, professor of electrical engineering and computer science at the University of Tennessee, who uses graph theory to analyze big datasets. “We have these tools that we’ve built over decades and applied to problems that arise in many disciplines. We just need to map them over to the



PERSONAL EXPOSURE MONITORING: This figure shows daylong recordings from personal air pollution monitors (PM_{2.5} = particulate matter smaller than 2.5 micrograms) for two different people. Colors indicate different microenvironments. For example, the person pictured in (b) experienced high levels of pollution while traveling between home and work (transport periods are in blue). Reprinted from S Steinle, S Reis, CE Sabel, et al., *Personal Exposure Monitoring of PM_{2.5} in indoor and outdoor microenvironments*, *Science of the Total Environment* 508:383-394 (2015).

Twitter is that it’s like you’re looking at them without them knowing that you’re listening.” He looks for positive emotions expressed in tweets and links these to the locations people are tweeting from (from GPS). One drawback with Twitter data is that only about two percent of Twitter users agree to make their location data publicly available, so the sample may not be representative, Sabel says.

exposome setting rather than redesigning them from scratch,” Langston says.

In graph theory, variables are viewed as points in space. Langston’s algorithm examines all pairs of variables in the dataset; if two variables are highly correlated, he’ll put an edge between them. “I have all these points and edges floating around in space and what our algorithms do is find the dense regions—areas where there are a whole bunch of edges, meaning all these variables are moving together.” These dense regions, called paracliques, can then be correlated with disease outcomes. More refined statistical analyses are then applied to try to isolate the causative factors from the mere confounders.

In one example, Juarez and Langston studied variations in the rates of premature births across counties. The lowest prematurity was found in Marin County, California, and the highest in Hinds County, Mississippi. They considered 590 variables, representing indicators from the economic, health care, physical, and social environments. Of 48 paracliques extracted, 17 correlated highly with prematurity rates. From there, traditional regression techniques identified race, obesity and diabetes, sexually transmitted disease rates, mother’s age, income, marriage rates, pollution, and health insurance as key drivers of disparities in prematurity rates.

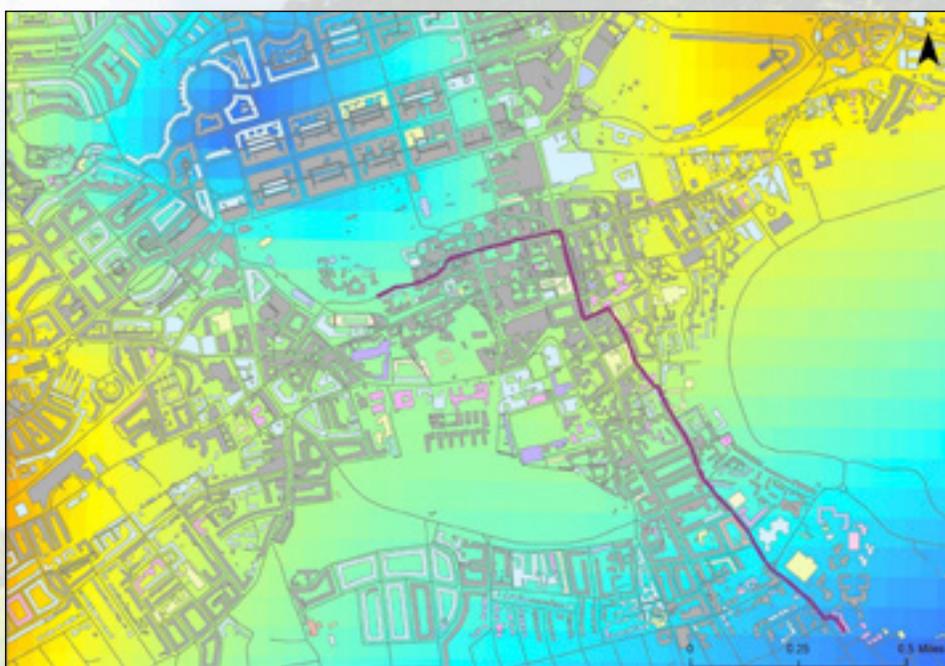
In another example, Juarez and Langston showed that disparities in lung cancer mortality for white men and women were largely driven by variations in smoking rates; but, surprisingly, disparities in lung cancer mortality for black men and women were driven more by differences in poverty, overall health, and access to health care. “The advantage of this data-driven approach is that it allows you to see patterns that you may not have thought about before with a hypothesis-driven approach,” Juarez says.

The lack of high-quality data management tools remains a critical obstacle. “The up-front handling of the data is still back in the stone ages,” Langston says. “Research scientists are going through files by hand, trying to move columns around,” he says. “We learned in biology years ago, if you’re going to deal with large volumes of data, then you’ve got to

bring on board a database administrator and a data curator so the domain experts can concentrate on the science,” he says.

ASSEMBLING THE EXPOSOME

Bit by bit, researchers are making inroads into the human exposome. But much remains to be done. Besides meeting the challenges already detailed, researchers also must figure out how to integrate all the layers of data—from the chemicals in our blood to the laws in our countries—and also link them to genome data, to get at gene-environment interactions.



HAZARD MAP: This map overlays a person’s GPS-recorded travels with a hazard map showing the concentrations of the pollutant nitrogen dioxide (as measured at fixed pollution stations). The map can be used to estimate a cumulative daily exposure for the individual. Concentrations of nitrogen dioxide are lowest in light blue areas and highest in dark yellow/orange areas. Courtesy of Clive Sabel, Bristol University.

The exposome community needs to adopt a “big science” approach akin to the Human Genome Project, comments **Christopher Austin, MD**, director of the National Center for Advancing Translational Sciences at the National Institutes of Health. To “kick it up to this level,” he advises exposome researchers to heed some lessons from the genome community. For example, he says, the exposome community should invest in improving measurement technologies, just as the Human Genome Project did for sequencing technologies; establish a public data

repository similar to GenBank, but for exposures; and agree on standards such as for variable names, meta-data, and security.

The key is to make the data easy to access and use, Austin says. “Otherwise, it becomes what a friend of mine calls ‘data composting’—you just put it on a pile and hope that, if it sits there long enough, something magic will happen.”

On top of all that, Austin says, the exposome community needs strong project management and leadership. With individual-investigator projects, you can make things up as you go along, Austin says. “The building isn’t that big, so if you need to build a foundation halfway through,

you just do it.” But big science projects need to be methodically planned and executed or they risk catastrophic collapse.

Understanding the exposome is an ambitious idea, Miller says. But it is far from impossible. In the early 1990s, people estimated that it would take 130 years to sequence the human genome. “But once the scientific community said, ‘Okay, we’re going to do it, and we’re going to invest money in it,’ they were able to rapidly accelerate progress and get it done under budget and under time,” he says. “It was really amazing what happened.” □