Shelters from the Storm
Sex-Based Defenses Against Disease
As the truck carrying Tom Kirsch crossed the Dominican Republic’s border into Haiti, the nation’s telltale signs of poverty—the primitive shacks, bare brown earth—came into view. There was no evidence of further catastrophe, of things that are even worse than the usual, grim day-to-day reality. That would change a few hours later when he wheeled into downtown Port-au-Prince, the Haitian capital. Two weeks earlier, on January 12, 2010, an earthquake had shaken the capital to its core, killing 237,000 people and leaving more than half of the city’s 2.5 million people homeless.

“The closer you got to downtown the more you saw people sleeping in any open space, including highway medians,” says Kirsch, MD, MPH ’87, an associate professor of International Health at the Bloomberg School and of Emergency Medicine at the Johns Hopkins School of Medicine. He serves both schools as an expert in disaster response. “Downtown itself was just horrible,” he recalls. Buildings were leveled. Streets were blocked by rubble, further hampering any relief efforts. The human toll was much worse, with harrowing examples of tragedy around each corner. An entire school of nursing students immersed in an exam—save one—was killed by the quake, which struck seven minutes before test time was up. Even two weeks after the quake, one
in three injured or ill Haitians had not been treated. “We’re talking about terrible bone breaks and deeply infected wounds,” Kirsch says. People needed water. Food riots had erupted when some earthquake victims thought aid wasn’t being distributed equitably enough.

In the past year, Kirsch has traveled three times to Haiti, the poorest nation in the Western Hemisphere, not just to deliver emergency medical care but to ask questions about how the world and the Haitian leadership, also decimated by the quake, had responded to the crisis. Had there been enough concern about finding housing for people? How was aid being distributed? Why was health care being delivered so slowly? What were the nonprofit relief groups that regularly swarm to disaster sites doing right in Haiti? Where were they coming up short?

As co-director of the Center for Refugee and Disaster Response, a collaboration of the School of Medicine and the Bloomberg School, where it is housed, Kirsch oversees grant-funded research projects aimed at finding out what goes wrong after everything goes wrong. Which is to say, Center investigators look for ways to evaluate and, ultimately, improve the quality of relief efforts and government services during refugee crises and after tragic mega-events.

It’s painstaking work. “The trouble with trying to do research is that everyone is over-
whelmed by the need to respond immediately,” says Kirsch. “It’s hardly an ideal situation to investigate things. What we want to do at the Center is improve the way we collect data in the midst of that chaos.” Because “disaster science” is an emerging field of inquiry, the Center is still working to figure out how best to do that, he adds.

In search of a method for evaluating how dire situations can be rapidly improved, Kirsch and the other 20 or so Center-affiliated investigators study the effects of disaster relief months, even years, later in hope of uncovering similarities between emergency responses in far-flung nations. They also train disaster responders to take advantage of the latest knowledge.

Too often, programs devised and carried out by NGOs, governments or the United Nations don’t do enough to get people back on their feet. Experts say that many countries don’t have the “capacity”—money or human resources—to deal with mass tragedies. Many developing nations lack disaster plans that would take full advantage of coordination and logistics. Security concerns slow the flow of aid to places where it is most needed. In many cases, getting people back to work isn’t emphasized, leaving people in poverty for longer than necessary. And even when a disaster elicits an outpouring of support, much of it never ends up on the ground. For example, only 20 percent of the $10 billion pledged by individuals, groups and governments to post-earthquake Haiti has been delivered.

What’s more, the lessons learned from one disaster often aren’t remembered during subsequent catastrophes. “What we’ll see is that different groups and nations will collect post-disaster information in different ways,” Kirsch says. “The field cries out for standardization.”

“‘When a disaster happens, there’s a lot of attention paid. But after the news trucks leave, people are often left homeless and jobless for years. People forget that.’”

Adds Courtland Robinson, PhD ’04, an assistant professor in International Health and at the Center: “What I’m hoping we can bring to this field are measures that go beyond profiles of a population or risk assessments, measures that can give us what I call a durable solution. Each time we do this, we shouldn’t have to reinvent the wheel. We should be able to take those approaches that have been validated by research and put them to work.”

Even though an earthquake in the Caucasus hardly resembles a drought in the Horn of Africa, the range of disasters presents many of the same challenges. “We want to be able to standardize what we do while having a customizable approach built in,” adds Robinson.

The Right Questions
In the course of their research, Center faculty and students find approaches that work. In Indonesia, for example, Robinson and his colleagues continue to measure the effects of emergency aid that followed the tsunami that steamrolled the Aceh region six years ago.
They found that an NGO program designed to put people back to work cleaning up their neighborhoods in exchange for cash did much to encourage people to return to their ravaged communities, and accounted for 93 percent of their household income, on average, as they put their lives back together.

Specialization within the disaster research field also appears to have value, Robinson adds. “What we’re learning is that if you’re going to measure the effect of an emergency on a household, you use demography to determine a rate of mortality, morbidity or other things,” he says. “For food security in a crisis, you can look to the assessments of nutritionists. The next question is whether there is a set of modules we can use across disasters that would prove effective at measuring response.”

The aim of the studies in Indonesia, as well as Haiti and Pakistan, is to find that set of comparative metrics. Investigators asked basic questions: How many people died following a disaster? How many households were affected? How many meals has each household eaten during an average day? How did the catastrophe affect a family’s finances and ability to earn a living?

Using cluster surveys, demography and other tools, scientists mine the surviving population for information that points up when help was made available and when it wasn’t, and whether people feel their needs have been served overall. In Haiti, for example, early results from a January study of households show that people there are greatly disappointed in the quality of relief services.

Others at the Center have continued years-long investigative stretches in war-torn areas, including Afghanistan and Iraq, and seven nations in East Africa that are regularly inundated by floods and landslides. In those countries, their work has led to concrete recommendations. A handful of doctors and public health scientists track disaster and response efforts in Japan, Singapore and South Korea, or examine food aid strategies in South Sudan, or monitor outflows of refugees from starved North Korea into China, or work on ways to improve the health and household economies of Iraqi refugees and displaced people in Jordan and Lebanon, as well as Gaza and the West Bank.

The Center has also helped educate disaster relief workers and planners. The School of Public Health and the Center authored *The Johns Hopkins Red Cross Red Crescent Public Health Guide in Emergencies*, now in its second edition. An ongoing program—Health Emergencies in Large Populations, or HELP—run by the Bloomberg School in conjunction with the International Federation of the Red Cross, has trained 500 people from a variety of countries and universities in the latest disaster response techniques, as well as in relief planning and handling refugee crises.

Although Kirsch and others will visit disaster-stricken areas right after an event hits, the Center does not function as an emergency responder, notes Gilbert Burnham, MD, PhD, MS, an International Health professor and Center co-director. “Our purpose is to analyze the depths of a crisis and point out where the response to it may be lacking, and how it might be improved,” he says.

The Center was founded in 2004 with
“Our purpose is to analyze the depths of a crisis and point out where the response to
the merger of separate programs—each run jointly by Kirsch and Burnham—at the
schools of Medicine and Public Health. Since then, it has grown markedly. Before
1998, when Burnham began his own group within the School of Public Health, “there
was little in Baltimore that dealt with refugees and disasters, even though our students
were very interested,” says Burnham.

“We knew we needed to grow something like the Center because some students
see working internationally with refugees or during disasters as a major part of their
careers,” he says. “Then there are other students who know they’ll need to learn
something about working overseas in general. And there’s a third group that has al-
ready been overseas, and [the students] were floored by what they saw and need to under-
tand it more before they go back.”

Kirsch himself became drawn to disaster scenes at an early age. (When a tornado
hit Omaha in 1975, Kirsch, then in high school there, joined the response efforts.)
His interest continued as he pursued his MD. “Like any medical student, I wanted to
save the world, so when I had the chance to work in Cambodian refugee camps in 1984,
I jumped at it,” Kirsch says.

He returned to the U.S. to get his MPH
at the School, steered there by James Cobey,
MPH ’71, who had organized the Red
Cross' refugee relief efforts along the Thai-
Cambodian border (and who still teaches in
the HELP course). “Jim told me, ‘If you
want to do this kind of work, you have to
go to Hopkins,’” Kirsch recalls. In 1986,
itook a seminar course put together by
Melvyn Thorne, a School of Public Health
professor and former Peace Corps doctor. “The students literally met in his living
room,” Kirsch remembers. “It was one of the
first attempts to quantify and teach refugee
response from a public health perspective—
I mean anywhere, not just Hopkins.”

The type of work the Center does is
still rarely found in academia. Governments
mostly welcome Center researchers—even
despite leaders want to appear competent
in the eyes of the public during times of
distress, Burnham says. To grease the skids,
Center faculty will often arrange partnerships
with government agencies to help them
set up and run their research programs. Not
that they’re always greeted with open arms.

“We have had staff questioned by au-
thorities,” Robinson says of his work in Chi-
na interviewing traumatized North Korean
refugees. “We tell them we’re working with
vulnerable populations.” China shut down
data collection during the 2008 Summer
Olympics. When such edicts are handed
down to researchers, they have to do their
work under the radar, Robinson says: “You
have to hide in plain sight.”

Elsewhere, their work has spun into con-
troversy. Burnham’s 2006 study that pinned
the numbers of Iraqi civilian war deaths
during the U.S. occupation at 655,000,
published in the British journal The Lancet,
became a political football among those de-
fending U.S. policies. His work in Iraq since
then has focused on health care systems.

“Our intention is not to get into the
political side but to improve situations,” he
says. Burnham’s team investigated the ef-
facts of the war on social sciences and medi-
cine to see whether they were recovering
from the war. The team found that 29 per-
cent of Iraq’s medical specialists had left the
country during 2006.

“Many had left because of assassina-
tion attempts,” Burnham says. “If you want
to destabilize a country, you go after the in-
tellectuals.” Fortunately, the situation is now
looking slightly better in 2011, he adds.

The same can’t be said for Afghanistan, which
Burnham calls “an ongoing humanita-
tarian disaster.” Relief workers and disaster
researchers around the world can’t do any-
thing to shore up dicey political situations,
Burnham notes. But they believe they can
learn enough to prevent some types of catast-
rophes, or at least minimize how many peo-
ple are affected by them and for how long.

He and a small team of International
Health department scientists, supported by
90 full-time workers on the ground, have
tracked the performance of the Afghan
health system for the past nine years. Some
of the team’s research on water sources and
diarrhea has led to the development of na-
tional water protection policies. An assess-
mom of hospital performance speeded up
hospital reform in the nation. Their work on
the spread of HIV led to major changes in
Afghanistan’s HIV policies.

Early Warning Systems
Although countries riven by civil strife,
refugee crises or war take up much of the
Center’s attention, it is well acquainted with
natural disasters. A five-year Center project
involving several countries in East Africa has
encouraged public health professionals there
to develop their own plans for local disaster
preparation and response.
“There are a lot of floods and landslides in these countries,” explains Daniela Lewy, MPH ’06, a research associate for the Center who spends part of each year in East Africa working on the project. “And they are increasing, possibly because of climate change.” Lewy and Center personnel train district-level public health practitioners to assess their country’s risk for disasters and develop action plans for dealing with them. They make sure that those practitioners, who are also teachers at public health schools, have some link to national ministries, so that government policymakers take their plans seriously.

Last spring, when landslides were scarring the faces of mountains and killing villagers in eastern Uganda’s Bududa district, Lewy’s project intersected with another one run by Shannon Doocy, PhD ’04, an assistant professor in International Health and at the Center. Students from Lewy’s East African program and others began to devise inexpensive ways to measure shifts in hillsides that could portend landslides nearby. “We taught people how far apart to place two poles on a hill,” she says. “If the [poles] move a certain distance apart, that’s a sign of instability. It’s something that local people can use.”

Doocy’s group had traveled to Uganda to see if studying the terrain could somehow foretell landslides—a lifesaving early warning system of sorts that could allow people to move to safe ground before disaster strikes. A newly developed (and decidedly higher-level) technology that utilizes geographical data taken during space shuttle missions, along with a topographical mapping scheme that was aided by global positioning satellites, allowed them to find a crack in a village hill adjacent to a camp for people displaced by another landslide. “We warned the local authorities of the risk and advised them that if there was further change in the slopes, they should move people out,” says Doocy, who adds that landslides did indeed occur nearby later.

Many who have come to study in the School of Public Health enroll in the master’s-level Health in Crisis concentration, which Doocy co-directs. But the Center’s reach extends beyond Hopkins students, thanks to the Health Emergencies in Large Populations course, run by Burnham.

Craig Jaques, now a program strategy consultant with a U.S. Department of Defense agency that prepares health professionals for disaster response, took the intensive three-week course at Hopkins during the summer of 2009. Since then, he has trained dozens of people from Southeast Asia on how to construct health and medical programs during an emergency—including lessons he learned during his time at Hopkins.

“The HELP course broadened my views on the international response to disasters and got me thinking about how to deal with public health concerns,” Jaques says.

Now he regularly teaches other HELP alumni. “They utilize the course’s tools and knowledge to train others,” he adds. “They respond with confidence to health emergencies.” Recent HELP graduates have gone on to devise health systems in strife-torn Somalia, develop preventive health programs in rural Pakistan and run HIV programs in South Sudan.

Yet others, like Paul Perrin, a doctoral candidate in International Health, come to the School to study so that they can continue their work amid disasters. Perrin, formerly a missionary for the Mormon Church, began disaster work when he was 19, helping victims of an Armenian earthquake.

Now 30, Perrin believes he and others are on the cusp of some exciting findings. “It’s a young field,” he says. “Most of the research done by NGOs isn’t robust because they measure their own idea of success. They’ll ask people who get their services what they think of their work but often don’t do the same for those who haven’t. They don’t use population-based methodologies. That’s the strength of what Hopkins does.”

Biostatistics and epidemiology classes give students a broad knowledge base with which to construct studies and work them out in the field, he adds. “Hopkins teaches you that this is a scientific discipline above all, one that takes a hard look at individual disasters and how they are dealt with.”

In Haiti, Perrin personally trained locals to do interviews in the field, overcoming a language barrier. (He doesn’t speak Creole.) He says that gathering such data from the 60 sites that Center personnel systematically chose is vital in improving NGO performance. “I suspect many relief groups think we’re looking over their shoulders, but I also think they’ll be very interested in our findings,” he says. “When you get down to it, they really want to do the right thing.”

"It may be lacking, and how it might be improved." —Gilbert Burnham
Sabra Klein forgoes a handshake, coughs by way of greeting. “Would you like one?” she rasps, sharing a bag of mentholated lozenges. “It’s not the flu. I don’t think.”

Flushed and weary, Klein plans to retreat early from her office and lab in the Bloomberg School, quarantine herself at home, rest and drink lots of fluids. But not before she infects you, via this article, with the germ of an idea she has long championed: Sex matters. In ways we never fathomed.

It matters, for instance, with the flu that Klein hopes she doesn’t have. Her ongoing studies show that females have a bigger and badder inflammatory response. They don’t just feel worse. They don’t just visit doctors more or complain more. They literally experience worse disease than males. Klein’s talking sex-based biology, not gender issues. (Although “sex” and “gender” often are used interchangeably, sex is biology while gender refers to the social constructs related to one’s sex.) Molecularly speaking, females respond differently to flu than males. They mount a more robust immune response—which sounds like a good thing, until you delve into Klein’s data and see that this heightened immunity contributes to tissue damage and even death. Females respond so differently to immunizations, she concluded, that a woman needs about half the flu vaccine dose of a similarly sized male.

Klein coughs, pops another lozenge and launches into why she loves the flu. For a dozen years, the assistant professor in the W. Harry Feinstone Department of Molecular Microbiology and Immunology has investigated infectious diseases, first focusing on hantaviruses, then malaria, now flu. The constant throughout her career has been her dogged insistence that sex matters.

That concept, neither new nor original, holds that every cell in us—indeed, every cell in the H1N1-infected mice languishing in Klein’s lab—has a sex. Advocates of sex-based biology contend that maleness or femaleness in humans as well as rodents needs to be considered, compared and contrasted in order to uncover basic biological truths about everything from heart disease and depression to lupus and liver cancer.

Klein’s data always have spoken louder to her than naysayers. As an infectious disease expert working in nonreproductive tissues and cells, Klein sensed for years that her grant submissions or research papers focusing on sex differences didn’t so much pique reviewers’ interest as annoy them. Fellow researchers who ignore sex differences have distinguished careers and mountains of data invested in their way of doing things. Some told her flat out: Sex did not matter.

Still, she stayed the course, giving sex differences center stage instead of sloughing them off. She used both male and female animal models in her hormone-centric studies. She manipulated estrogen and testosterone, surgically removing the bird-seed-size ovaries and testes of mice, and then put hormones back, always looking for cause-and-effect relationships between sex and disease. She analyzed her results by separating the sexes in the statistics instead of lumping males and females together in one big androgynous data set, as was—and still remains—conventional practice. (In top-tier journals, it’s common for authors of clinical studies to demonstrate demographic correctness in Table 1, showing 50 percent of their subjects were male and 50 percent female. However, after that obligatory nod, the breakdown by sex simply goes away, never to appear again in Tables 2, 3 or 4, Klein says: “There’s no more mention of sex. No statistical comparison. It’s sooooo frustrating!”)

Klein saw firsthand, time and again, that sex was remarkably relevant to her research on animals’ susceptibility to infection and their response to pathogens. Indeed, sex often was the only difference between those that recovered and those that succumbed to disease. With influenza, for example, when she gave male and female mice a standard dose of virus, none of the females would survive past two weeks, while more than half of the males would recover, surviving the infection. For many infectious disease mouse models, if fewer than half of the mice of a particular strain survive an infection, then that strain is defined as “susceptible” and if more than half of the mice live, that defines the strain as “resistant.” But while working with two sexes of the same strain, Klein noted stark differences.

There are other people—not so many in infectious disease but more studying the...
heart and brain—who are as convinced as Klein that males and females differ in their basic physiology and, therefore, in the susceptibility to and progression of diseases. Arthur P. Arnold, a professor and chair of physiological science at UCLA who assumed editorship of the just-launched journal *Biology of Sex Differences*, has been doing sex-based biology since the mid-1970s when he was a neurophysiology postdoc at The Rockefeller University. He realized early on that gonadal hormones explained only part of the reason a brain structure in male songbirds was six times larger than in nonsinging females. Some sex differences, he has since discovered, are the result of direct action of genes encoded on the sex chromosomes: XX and XY cells differ functionally because of the action of X and Y genes intrinsic to the cells.

The fact that there are sex differences in disease implies that one sex has something protective about it, Arnold says. If that something could be enhanced or modified, it might affect the disease. That something might make an attractive drug target, for instance, and only a lack of understanding of the biological basis of sex differences in disease keeps us from hitting that target with new therapies.

Researchers like Arnold and Klein quietly celebrated a milestone 10 years ago when the Institute of Medicine issued a report ("Exploring the Biological Contributions to Human Health: Does Sex Matter?") concluding that every cell has a sex, and therefore sex matters in health “from womb to tomb.”

Heartening as that was, it more or less preached to the choir, Klein says. That choir subsequently formed a new academic society—the Organization for the Study of Sex Differences—to promote the interests of this emerging field, not least of which was to address the fact that only a pittance of NIH grants supported the study of sex differences.

Klein, with a small cadre of colleagues across various disciplines, proposed in 2008 to establish a center for the study of sex-based biology at Hopkins. Helping lead that effort was Pam Ouyang, a cardiologist based at Bayview Medical Center, who says, “Men and women are different in lots of ways that we don’t necessarily understand yet because we don’t study them. I thought it would be really nice to have a place where general conversations about sex differences could be discussed with people from various spheres of knowledge who would approach questions—such as why are men’s and women’s risks of heart disease different—from all angles.”

Ouyang no doubt would have been intrigued by DeLisa Fairweather’s perspective on autoimmune heart disease in men. A staunch advocate of sex-based biology who is based in the Bloomberg School, Fairweather is teasing out how inflammation induces chronic conditions in males and females.

Despite their common interests, Ouyang and Fairweather were not yet destined to meet. The proposal for a center was denied. The upstart field of sex-based biology limped along on its uphill trek.
Then the flu hit.
Not just any old flu, but the H1N1 pandemic of 2009. The fact that it hit in the midst of a panic about vaccine shortages proved fortuitous for Klein. Steeped in sex-based biology à la influenza, she was ready with a kill shot. She leveraged interest in the pathogen, finessing a low-grade professional buzz about sex differences into a very public debate. Flu, once and for all, confirmed the legitimacy of her convictions.

“Sabra was the prepared mind in the right place at the right time,” says Florence Haseltine, MD, PhD, director of the Center for Population Research at the National Institutes of Health and a founding member of the Organization for the Study of Sex Differences. “H1N1 threw her—and sex-based biology—into the limelight.”

Klein, a mother of two girls, promptly coauthored an op-ed (“Do Women Need Such Big Flu Shots?”) published in *The New York Times* in October 2009. She wrote, “In all likelihood, we’d have a better H1N1 vaccine—and more of it—if in our preparations we had accounted for the biological differences between men and women.” That article sparked the WHO to enlist Klein’s help in preparing “Sex, Gender and Influenza,” a report issued in July 2010 that examines the 2009 H1N1 pandemic through the prism of sex differences in immunology.

Meanwhile, Klein also coauthored a review that appeared in the May 2010 edition of *The Lancet Infectious Diseases* (“The Xs and Ys of Immune Responses to Viral Vaccines”) that re-examined published data from a high-profile paper by analyzing it according to sex. The review revealed conclusions strikingly different from those of the original authors who had ignored sex. When Klein re-analyzed the genomic data by sex, she found that the transcriptional activity along immunological pathways—pathways that supposedly predict long-term protection following, in this case, yellow fever virus vaccination—was 10-fold higher in samples collected from female than from male volunteers, suggesting that females may be better protected than males.

About that same time, a landmark reference book Klein co-edited was published: *Sex Hormones and Immunity to Infection*.

Amid her publishing flurry, Klein accepted invitations to speak locally and abroad. She described how her female mice were mounting inflammatory responses up to a hundred-fold higher than males in the first week after flu infection. She cautioned audiences against assuming that a bigger immune response is better. Take the 1918 flu or the avian flu, for example: They caused profound sickness and death, she explained, not because of out-of-control viral replication but because the human hosts—and hostesses—initiated excessive immune responses to those pathogens.

Despite preaching about a pervasive lack of consideration for sex differences in the design of scientific studies, and therefore in the analyses of data, Klein never expected people to suddenly fixate on sex differences. She just wanted them to pause and question the assumption that males and females weren’t going to be different.

And then, out of the blue, a Hopkins
“In biomedical science, the dogma is that there are no differences between men and women. People like myself who design studies looking at both sexes are left with that uphill battle of challenging the dogma.”

—Sabra Klein

physician investigating Lyme disease contacted Klein about analyzing his patient data according to sex. On the heels of that request, a scientist from New York University conducting malaria research in Peru sought her out, wanting to discuss intriguing sex-based trends. Most recently, a Harvard researcher offered to share unpublished epidemiological data and asked to pick her brain about sex differences in flu immune response.

“Flu is of such great public health importance,” Klein says, “that it puts all of this sex-based biology business in a context people suddenly care about.”

Sex differences were not a part of John Aucott’s original research plan when he set out three years ago to study the natural history of Lyme disease. In his field, there was no precedent for separating out male cases and female cases. It simply hadn’t occurred to him, a Hopkins-trained physician and fellow in infectious disease, that sex had anything to do with his patients’ biological reactions to a tiny tick bite.

Aucott is the go-to guy in the Mid-Atlantic region for Lyme patients with chronic post-treatment issues. Considering that Lyme is an emerged epidemic, and he’s based at its epicenter, Aucott is busy. A steady stream of acutely ill patients sporting telltale rashes show up daily at the urgent care center next door to his office in suburban Baltimore. From their ranks, he recruits study subjects “by the gobs.” Already he’s beginning to glean valuable information about the “normal” human immune response to an acute infectious disease over time.

“Actually, we’re going to get a two-fer,” Aucott enthuses. “We are going to describe the normal human response in men and in women. It can’t be the same. Why would it be? I’m convinced of that now. I’m a believer.”

Aucott credits his conversion to study coordinator Alison Rebman. It was her key observations that led to finding differences in antibody responses between men and women—and, ultimately, to Sabra Klein. About a year ago, she noticed significant trends that appeared to vary by sex and wondered if anyone had ever looked at sex differences and Lyme before, so she dug into the scientific literature. “I came across a review Sabra Klein had written about viral infection, and then a chapter of her book. I kept coming across her name, and after reading a few more articles, realized she’s at Hopkins!” Rebman says.

She contacted Klein, whose excitement about their Lyme findings was, well, infectious. “Sabra talked about how the female immune system is designed differently than a male’s,” Aucott says. “It has to do with all these crazy immunologic maneuvers to tolerate being pregnant. She’s helping us to think about how to design our study and how to analyze the data.”

Rebman’s initial discovery—that sex may affect the antibody response measured by the diagnostic test—landed her and Aucott in unexplored territory. It was known that the standard test for Lyme is not very sensitive, and apparently it may be even less so for women. Possibly, many, many more women go undiagnosed than ever suspected.

People who don’t get treated are known to develop months or years later a condition characterized by big swollen knees; it’s known as late Lyme arthritis, and twice as
many men get it as women, according to Aucott.

“We know that equal numbers of men and women get Lyme to begin with, and twice as many men get late Lyme arthritis, so what happened to all the women?” Aucott asks. “Did they all just get better without treatment?”

He thinks not. Rather, he speculates that maybe the women who didn’t make enough antibodies to fit the diagnostic criteria for Lyme in the first place ended up getting diagnosed with something else: fibromyalgia, for instance.

Anecdotal information, now backed up by sex-based biology, informs his hunch: Acute Lyme disease is easily treated with an antibiotic. The rash and other symptoms disappear. However, 20 percent of people post-treatment—the majority of them women—develop a fibromyalgia-like syndrome some months later. The overall ratio of women to men with fibromyalgia is 7:1.

Aucott sees lots of patients who report having felt fine before their bout with acute Lyme, then recover, only to develop a vague constellation of disabling symptoms some time later. The majority of them are females.

“These are people who are told by their physicians that they are ‘just depressed,’” Rebman says. “Well if they weren’t depressed before the rash, there’s got to be more to it. That can’t be the end of the story.”

Sex differences complicate the story. No doubt about it. Marc Lipsitch, director of the Center for Communicable Disease Dynamics at the Harvard School of Public Health, will attest to that.

His lab was looking for evidence that early exposure to certain strains of influenza proved protective against infection in the 2009 pandemic. Lipsitch’s group noticed, as others had, that older people seemed at lower risk of being a confirmed flu case in 2009. Why would that be? Did they have antibodies from immune responses earlier in life? Did they have fewer children at home and therefore were less exposed to school kids, known transmitters of flu?

“We thought it would be important to know differences between men and women in the context of this drop-off in risk,” Lipsitch says, “because you might expect a bigger drop-off in women than men if it was due to exposure.” When his team separated out the data by sex, they realized that although the drop-off in risk is much stronger in women, it does not appear to be due to any lack of contact with flu-infected people.

Sabra Klein came to mind: Lipsitch recalled seeing a poster at a conference by one of Klein’s students, showing how the immune system behaved differently according to sex.

“We thought she’d be interested in our unpublished data,” he says, admitting that he originally intended to ignore sex. “We actually had hoped to reject the idea that what we were seeing had something to do with sex differences in the immune experience.”

“But it was really helpful that Sabra pushed us to look at both age and sex at the same time and make sense out of data of this sort, even though now it’s a more complicated story.”

OraLee Branch, an assistant professor of medical parasitology at New York University, studies the immune response to malaria infection. When her students, who
accompanied her to Peru for a field investigation, noticed sex-based trends in their data, she scrambled to find anything even remotely related in the malaria literature, finally unearthing a paper Klein published in 2006.

“I right away picked up the phone and explained to her what we were finding,” Branch recalls. “I wondered if she had followed up, if there was more information. Her work had since veered toward viral pathogens. But still, she offered to help.”

One of the things Branch had spent time investigating was differences in malarial symptoms based on differences in exposure.

Now she wants—no, she needs, she says—to go back into her data to reanalyze it, taking sex differences into account. “It’s really opened up a can of worms, it has,” Branch says. “We might find immunity is developing better in males than females, or maybe the other way around. But pretending sex differences don’t exist is just going to obscure the real mechanisms we’re all trying to find. Pretending differences don’t exist is not the answer.”

All was unusually quiet throughout the Bloomberg School of Public Health on the snowy holiday commemorating Dr. Martin Luther King.

Michael Coronado—winner of the Florence P. Haseltine Award for young investigators at the 2010 annual meeting of the Organization for the Study of Sex Differences—was hard at work in DeLisa Fairweather’s lab, however.

First, he was extracting blood from male and female mice that were infected with the same virus but producing different levels of cytokines—the signaling molecules that cells generate to communicate with each other. Then, he was removing their hearts to study the diseased tissue. Later, on other live mice, there were gonadectomies to perform.

By de-sexing the males and females, he was attempting to make them biological equals, at least in terms of the type of heart disease he’s studying. It’s apparent that testosterone drives the disease. But it also could be that estrogen protects from it. Or both could be true.

One thing’s for sure: Male and female mice are anything but equals in terms of their hearts, as the males’ floppy, scarred organs clearly show.

Some investigators would justify using only male mice to study disease that affects mainly males. They’d argue that it’s expensive and inconvenient to deal with female rodents because it means having to account for hormonal cycles.

For that matter, lots of researchers who study equal opportunity diseases such as cancer don’t use both sexes of mice either, no matter that half the population whom the research is ultimately meant to serve happens to cycle, too.

The converse is also true. Lots of researchers studying so-called women’s diseases such as lupus and multiple sclerosis use only female mice and therefore risk missing half the picture.
Both males and females always are used for comparative purposes in the Fairweather lab where sex differences have been shown to drive autoimmune disease.

Autoimmune diseases are notoriously lopsided in terms of whom they strike. Thyroid disease affects mainly women while dilated cardiomyopathy (the chronic type of heart disease that Fairweather’s studying) affects men more.

It’s not simply good practice to acknowledge sex differences, it’s outright dangerous to ignore them, says UCLA’s Art Arnold: “Treating one sex like the other may be as inappropriate as treating a child like an adult. The equitable treatment of females and males requires an understanding of their differences.”

When Sabra Klein gave presentations early on in her career and talked about the differences she saw between the sexes, she was invariably followed by a more senior colleague who would not mention any differences between males and females. Inevitably, an audience member would ask if he saw the same types of differences that Dr. Klein reported.

The answer, Klein recalls, always was curt and final: No. Absolutely not.

“That would be it, with no further explanation,” Klein says. “And I’d feel everything sink. I wasn’t born with thick skin. I’d take it personally and worry: What are people trying to say about me, my data, the quality of my work?

“But as soon as I’d get back into the lab, into what we were doing, I’d get over the intimidation because I was excited by the research. I believed in it.”

It didn’t hurt that senior faculty in her department believed in her work as much as she did: “Diane Griffin, Al Scott and Greg Glass had a lot to do with my ability to stand up and be that lone voice in the wilderness. They took a risk hiring me. They told me I needed to lay claim to a field.”

These days, the very same people who made presentations after Klein and disavowed sex differences now are sending their data to her, reporting that they too are seeing trends. They now allow Klein to present their unpublished data in talks, which is a giant leap forward. But out there in the published world, sex differences remain largely ignored. “When you put your data out there, you are choosing to enter a debate,” Klein says. “It’s not some definitive conclusion; the book is not closed, the story not over. You put yourself out there to be judged, and to be open to people’s interpretations. I am always working to hone my arguments, to improve on my logic.”

But as the long, cold flu season gives way to spring, Sabra Klein senses a thawing of attitudes toward sex-based biology. Her team recently resubmitted a research paper for consideration by a prestigious journal. Because the reviewers’ comments and concerns had been so constructive, Klein has reason to believe that she’ll hear good news back soon: “I think they might actually be rooting for us.”

Sabra Klein describes “his and her” heart attacks: magazine.jhsph.edu/extras

obscure the real mechanisms we’re all trying to find. “ —OraLee Branch
How to bring bullying into the light

It’s a cold, gray January morning in Edgewater, Maryland, and the 1,100 students in Central Middle School are rambunctious. Wearing coats, hoodies, T-shirts and jeans, they stop first at their powder-blue lockers and then head to their homerooms while carrying on conversations that compete with the din of a building filled with eager adolescents. They may be extra-amped because a snowstorm closed the school yesterday and delayed today’s opening; plus, it’s Friday—the weekend’s coming up. So, as Principal Mildred Beall kicks off the morning announcements, teachers have to spend extra time settling everyone down.

Soon, four eighth-grade girls gather round the microphone and wait for a fellow classmate to play the opening chords of Journey’s “Don’t Stop Believin’” on the piano. They sing utterly new lyrics:

I’m just a new kid at Central Middle School
I get laughed at every day at school
It’s just not fair, I’ve got no one here
Why doesn’t anybody care?

The lead-in to the chorus quickly answers:
BAC, we’ve got your back
We are right here, if you need us.

When the song’s over, two more members of BAC (Bully Awareness Crew) make announcements—one for a lunchtime art contest, the other a poem about “kindness” urging listeners to “pass it on.”

If this sounds too good to be true—a Hallmark moment in a Lake Wobegon-like middle school—consider this: The students in Beall’s office are just a handful of the 86 eighth-graders who have joined BAC since its inception in the spring of 2010. That’s when the class’s guidance counselor, Sandra Seward, asked several of its then seventh-grade members to come up with a student-led means to combat what the federal government considers an urgent problem.

“About 30 percent of youth report they’ve been involved in bullying, as a victim or perpetrator, on a frequent basis within the past month,” says the Bloomberg School’s Catherine Bradshaw, PhD, associate director of the Johns Hopkins Center for the Prevention of Youth Violence, and a renowned expert on bullying. In August 2010, she spoke at the first-ever Federal Partners in Bullying Prevention Summit in Washington, D.C.

By the time students graduate from high school, Bradshaw notes, 80 percent will have witnessed at least one bullying incident. At the middle school level—the peak time for bullying—more than 30 percent of students don’t feel safe from physical, verbal and/or indirect abuse, including cyber-bullying, she says.

Central Middle, like other schools in Anne Arundel County, Maryland, is doing something about the problem—in part because of state and school district mandates. Maryland is one of 45 states that, in the past few years, have passed anti-bullying laws, many requiring school systems to implement reporting, intervention and prevention procedures. While there are many reasons for the current focus on an age-old problem, headline-grabbing incidents have raised its profile. Last year, a Massachusetts high school freshman named Phoebe Prince committed suicide after being bullied for months by a handful of classmates. Her death prompted the state government to pass sweeping anti-bullying legislation in May 2010.

“This is a watershed moment,” says Deborah Temkin, research and policy coordinator for Bullying Prevention Initiatives at the U.S. Department of Education. “We’ve hit a point where the effects of bullying have struck such a chord with people that they’re really taking notice.”

The choice is to act or to continue to suffer adverse public health consequences, says Philip Leaf, PhD, director of the Johns Hopkins Center for the Prevention of Youth Violence. “On the one end, you have people dying from it, committing suicide,” he says of some bullying victims. As for the bullies themselves, he notes, “children in adolescence engaged in aggressive behavior are at much higher risk for both subsequent juvenile issues and substance use.”

Bradshaw has put together a list of potential effects of bullying. For victims, they include anxiety, depression, lack of sleep and dislike of school. Perpetrators tend to feel the same way about school, while assuming that aggressive behavior is acceptable. And both groups are at a higher risk than their classmates for low academic performance and/or dropping out.
Which is why Bradshaw, who’s been researching the subject for 10 years, has also advised federal, state and district officials on bullying prevention. Her relationship with Anne Arundel schools, in particular, is unique. With the district’s help, she designed a Web-based bullying survey, which serves as an annual data collection system on bullying’s effects on students, staff and parents in the district’s 120 schools.

The questions vary—depending on who, anonymously, takes the survey—but they cover common ground. Both students and teachers, for example, are asked where bullying occurred (e.g., classroom, hallway, playground, bus), in what form (name-calling, rumors, teasing, pushing), and what the student’s reaction was (ignored it, told an adult, bullied back). Among the questions asked of parents is how they reacted when their children were bullied (talked to the child, the bully, the bully’s parents or someone at the school).

“It’s a research-based measure [to describe] what bullying looks like,” explains Bradshaw, who’s featured the findings in numerous papers written since the survey was introduced in 2005. Theoretically, each school can use the survey results to target problems and improve prevention strategies, but they’re not required to, and many schools, like Central Middle, focus more on the number of incidents officially filed as “bullying.”

Chuck Buckler, director of Student Services and Alternative Programs in the Maryland State Department of Education, says, in many schools, that number has increased—and for good reason. “The data may suggest that schools have more incidents [than in previous years],” he explains, “but it’s really a matter that they have more kids, parents and friends willing to file a report. And that means the awareness is there.”

Awareness, Bradshaw says, is the first step in tackling the problem. But it’s not an easy first step for many schools, where the design and implementation of prevention programs vary. Part of the challenge, she says, is that the causes of bullying and the incidents themselves are extremely complicated and, thus, demand more than a disciplinary approach. Ideally, the entire school community—students, teachers, administrators and parents—should participate in an effort that’s sustained over time and not changed from year to year, depending on resources and who’s in charge.

It should also begin, at the very least, in elementary school. “You want to try to get in there early to get kids on track.” Bradshaw says, “because if you wait ’til they have problems, then [behaviors] might be entrenched and harder to change.”

At Central Middle School, the effort includes what she considers key to a whole-school approach—students. Says one BAC member, “We’re just trying to be a good influence for other students, the lower grades and stuff, and say, ‘Hey, bullying’s not cool. You should try to be better people.’”

Lunchtime for the 300-plus sixth-graders at Central Middle School may play havoc on the ears, but there is order in the cafeteria, where one of the monitors is Seward, a former art teacher wearing a black dress, matching boots and a big smile. She grabs a microphone and lets the packed lunchroom know that it’s time to check out the BAC table, where a couple of members will explain the “Two Hearts Contest.”

Soon, a handful of girls are lined up, each holding a pink flier instructing them to draw a design within a heart that “portrays kindness.” They’re told that once two winners are chosen, Seward and a few BAC members will decorate two large wooden hearts using the designs as blueprints. Plus, each winner will get a $15 iTunes gift card.

There’s no pitch for the BAC or any attempt to engage the sixth-graders in a bullying discussion. “Oh, no,” Seward says, “these kids already know about the BAC.” Aside from the weekly announcements, there’s also visual evidence: the laminated hearts, smiley faces and peace signs students purchased from the BAC, at 25 cents each, and inscribed with friendly messages. They share space in the hallways with bright yellow banners featuring BAC-composed slogans such as, “Step Up! So Others Don’t Get Stepped On,” “Friends Don’t Let Friends Be Bullies” and “Bullying Is Cruel and NOT Cool.”

Meeting in a conference room with 15 of the BAC members is a good way to sample the enthusiasm firsthand. While they take their role-model responsibilities seriously, they’re also having fun, in part because Seward facilitates biweekly after-school meetings in which members discuss and formulate new ways to get the antibullying message out. The group, which has made presentations to parents and district officials, hopes to cap the school year with

“Sending kids [who’ve bullied] home for three or five or 10 days has no remedial impact in terms of changing their behavior. If they’re not in school, and not learning anything, there won’t be any changed behaviors.”

—Philip Leaf
an assembly featuring performances and a PowerPoint presentation. “This is our first year doing this,” says Seward, whose aim is to have half the eighth-grade class, or 150 students, join by year’s end. “So it’s very much a work-in-progress.”

Even so, the students say they’ve seen the school’s atmosphere change since implementation of the BAC, which welcomes new students—often prime bullying targets—and helps Seward recruit “borderline bullies”—those whose behavior tends toward aggression but can be changed, she says.

To appreciate how far schools like Central Middle have come, it’s worth looking back 10 years, when a series of school shootings, including one that took the lives of 12 students and one teacher at Columbine High School in Colorado, were met with extreme responses: metal detectors and zero-tolerance policies, among them. “We’ve since learned that the punitive route is not the way to go,” says Temkin of the U.S. Department of Education.

“Sending kids [who’ve bullied] home for three or five or 10 days has no remedial impact in terms of changing their behavior,” concurs Leaf, who, along with Bradshaw, is a co-director of the youth-focused Johns Hopkins Center for Prevention and Early Intervention. “If they’re not in school, and not learning anything, there won’t be any changed behaviors.”

The numbers play this out. The rate of bullying, according to Bradshaw, has remained stable, even as other forms of school violence have declined in the past decade.

Bradshaw first took up the anti-bullying crusade in the late ’90s, when she realized, as a graduate student counseling youth in detention centers, that they’d been subjected to various forms of violence, including bullying in schools. The attitude then was that bullying was a rite of passage to be endured. And the research was scant—one or two papers a year. Post-Columbine, however, there was a rush to investigate bullying, accompanied by state and federal grants. “Now, you see at least 100 papers a year,” she says.

BULLYING DEFINED ONLINE AND OFF

SO WHAT, EXACTLY, IS BULLYING? Most of the definitions used at the district, state and federal levels fall in line with three basic features that bullying expert Catherine Bradshaw, PhD, uses in much of her research:

• Involves aggressive behavior that intends to cause harm or distress
• Is usually chronic or repeated over time
• Occurs in a relationship where there is an imbalance of power or strength

In addition, its forms are verbal (e.g., threatening, name-calling), physical (hitting, kicking) and indirect (spreading rumors, influencing relationships), which includes cyber-bullying.

The latter, though just roughly 10 percent of the total, is on the rise—in part, according to Philip Leaf, PhD, and director of the Johns Hopkins Center for the Prevention of Youth Violence, “because most youth have cell phones, so they’re at risk 24/7.”

In a study on cyber-bullying at the middle school level (grades 6–8), the peak time for bullying activity, Bradshaw reports that 25 percent of girls and 11 percent of boys have been cyber-bullied at least once. The most common methods: instant messaging, chat rooms, emails, websites and texting. And the types range from “flaming” (online fights with angry language) to “outing” (sharing secrets or embarrassing information) to “cyberstalking” (intense harassment, including threats).

—RS
In general, Bradshaw reports, “the data suggest that you start to see bullying pick up in late elementary school—grades 4 and 5. Middle school tends to be the peak, and around 10th, 11th grade, it peters out.”

The exception is cyber-bullying, which is on the rise, even though it’s still a small slice of the pie, with roughly 10 percent, on average, claiming to be victims (see sidebar on page 35).

Whatever the form bullying takes, research indicates that a positive school environment is key to prevention. And, in 2007, Bradshaw and two colleagues reported disconnects between staff and student perceptions on this score. For example, more than 70 percent of staff in elementary schools, 40 percent in middle schools and 57 percent in high schools assumed that the number of students bullied in the previous month was 10 percent or less. But students at those grade levels indicated that 34, 33 and 23 percent, respectively, had been bullied in that time period. And at the middle school level, more than 30 percent of students felt that staff did nothing to follow up reports of bullying.

What makes middle school such a challenge is the timing. Sixth- through eighth-graders are sandwiched between elementary school, when students are supervised round-the-clock, and high school, where the primary focus is academics. Meantime, they’re encouraged to take on a host of responsibilities—juggling the demands of multiple classes and teachers, for example—even as their maturity levels are in flux.

“I don’t think [middle school] kids are natural bullies,” Principal Beall explains. “But, developmentally, they have so many social issues to deal with. Intimacy is important, but they’re very immature. They have feelings of jealousy and competition and sort of love-hate things going on. Generally, they get all mixed up.”

So it’s perhaps no surprise that Beall and her staff recently had to handle a “sticky situation.” A new eighth-grade student, a girl who’d just moved to the area, made the mistake of poking fun at a classmate who, with the aid of friends, retaliated verbally with the aid of friends, retaliated verbally.

“With the announcements, the banners, the songs, what they’re doing is creating a culture. We’ve set up an environment that says, ‘This is not OK. We do not accept this in our school.’”

—Sandra Seward

A whole-school approach to bullying is exactly what district and state policies—and the U.S. Department of Education, which released policy recommendations this past December—endorse. Bradshaw, in fact, goes a step further, recommending a schoolwide framework, developed by the University of Oregon, called Positive Behavioral Interventions and Supports, or PBIS. Although it doesn’t target bullying specifically, PBIS is aimed at establishing—via data collection and positive reinforcements—a safe environment in which students thrive socially and academically.

“We love PBIS,” Buckler says of Maryland, where more than 800 schools have signed on, including half the schools in Anne Arundel County. He cautions, however, that if specific anti-bullying measures aren’t included, “PBIS isn’t going to cut it. It isn’t a stand-alone.” Schools, he says, need to make sure that they’re tailoring anti-bullying efforts toward the needs of their students.

At Central Middle, Beall has asked her best classroom managers—those skilled at keeping students engaged and out of trouble—to share their techniques with colleagues while giving counselors like Seward free rein to involve the students.

“With the announcements, the banners, the songs, what they’re doing is creating a culture,” explains Seward, who, like her fellow counselors, teaches anti-bullying lessons throughout the year. “What I’ve noticed with BAC is, kids are informally reporting things to me that they would not have done before. We’ve set up an environment that says, ‘This is not OK. We do not accept this in our school.’”

This is crucial for Beall. “Formal bullying is not a huge problem [at Central],” she says. “It’s the annoying kinds of behaviors that can lead to bullying that get to be a problem. And the guidance department and student involvement help curb it.”

A drop in bullying referrals—from 13 last year to two this year—is proof, as are Central’s latest bullying-survey results, released in January. In most categories, the school bettered the district average. Only 54 percent of students (compared to 62 percent districtwide) feel that bullying is a moderate-to-serious problem, and 78 percent (versus 72 percent) feel safe in school. Would Beall like to see those numbers improve? “Of course,” she says, “but we’re moving in the right direction.”

Bradshaw, who spoke at her second White House bullying summit earlier this year, is of the same mind. “Many anti-bullying policies focus more on documentation and responses than on prevention,” she explains. “They need more training opportunities and evidence-based interventions. And from the research field’s perspective, we’d like to know more about what really works and how to change behaviors. But it’s a start.”
The Monday Campaigns—dedicated to “the day all health breaks loose”—started with one great idea and are changing how people eat, exercise and make decisions about a host of other health-related behaviors.

Humble Beginnings
It all started with Meatless Mondays—the brainchild of former ad executive Sid Lerner, who in 2003 decided he wanted to do something about rising meat consumption. He remembered “meatless Mondays” from a World War II rationing campaign during his childhood and decided it was a catchy way to get Americans to give up meat—the source of most of the saturated fat in our diet.

In 2005, Lerner partnered with the Bloomberg School and Columbia’s Mailman School of Public Health to start a nonprofit to persuade Americans to adopt healthy behaviors on the first day of the week. And so Healthy Mondays were born. The Johns Hopkins Healthy Monday Project (jhsph.edu/clf/monday) provides scientific and technical assistance to the national Monday Campaigns (mondaycampaigns.org).

The Data to Back It Up
The wisdom of homing in on Monday was confirmed by a literature review in 2010. Researchers Jillian Fry, MPH, and Roni Neff, PhD ’06, MS, at the Bloomberg School’s Center for a Livable Future (CLF), looked at the effectiveness of periodic public health messaging prompts and the cultural significance of Mondays. A literature review of 19 studies, with a combined sample size of 15,655 participants, found that frequent, periodic messages are an effective way to get adults to adopt healthy behaviors.

A survey by a research firm also found that more than half of 1,500 surveyed adults over the age of 25 viewed Monday as “a day to get their act together.” It was the day most would start a diet or exercise regime. As Lerner has said, Monday is the “January of every week.”

Reinventing Monday
That Monday comes 52 times a year means the campaign has multiple chances to appeal to our better natures. The Johns Hopkins Healthy Monday Project has partnered with everyone from health insurance companies, like Wellpoint/Anthem Blue Cross Blue Shield in Virginia, to the food giant Sodexo to create wellness plans and promote lifestyle changes. “It’s about changing people’s default self-image,” says Ralph Loglisci, project director for the Healthy Monday Project at CLF. “The Monday concept works perfectly for that. If you say, ‘I am someone who is healthy, it’s just that every once in a while I fall off the wagon,’ you can always hit restart on Monday.”

A survey by the Meatless Monday campaign in 2010 found that 30 percent of all Americans were aware of the campaign—more than double the number from two years previously. “Meatless Monday has just exploded,” says Loglisci.

The national Monday Campaigns promote seven Monday projects, including Man Up Monday, which encourages STD testing, and Quit and Stay Quit Monday, which targets smokers.

Getting the Oprah Shout-Out
Healthy Mondays got the coveted Oprah nod in February. The media mogul hosted a show about forgoing meat, at one point cheering to the camera, “Go Meatless Monday, Meatless Monday!” Loglisci says it was the most exciting pickup the project has gotten so far.

“The nontraditional approaches to push the message have included an appealing website, contests for meatless chili recipes, clever and humorous signage, and recruiting prominent health professionals, celebrity chefs and others,” says Robert Lawrence, MD, CLF’s director. Such pop culture venues offer exposure for health messages that traditional public health channels cannot match.

What’s Next
The Johns Hopkins Healthy Monday Project has partnered with Baltimore City Schools and its Education Channel 77 to launch the next Monday program: Kids Cook Mondays. It encourages children to take charge of food preparation and their own health. The Baltimore pilot program—including television spots with local kids—began in March.

by Phoebe Connelly
In the late 1980s and early 1990s, investigators at the Medical College of Wisconsin wondered if popular men in gay bars could effectively promote safe-sex messages. They designed a randomized-controlled trial (RCT) that was conducted in three small Southern cities. In Biloxi, Mississippi, where the intervention took place, the popular men were recruited to endorse condom use to bar patrons. The two comparison cities received no specific intervention and served as controls. The results were impressive. In Biloxi, the number of risky sexual encounters fell by about 30 percent over two months. In the other two cities, the risky behavior stayed the same. Another trial, this time in 16 cities, yielded the same results. A new, effective HIV/AIDS intervention had been identified.

Ten years later, epidemiologist David Celentano, a veteran in the field of HIV/AIDS prevention and of RCT design, did a similar trial with the same intervention in five countries on four continents. Over 24 months, the number of risky sexual acts decreased by 33 percent in the intervention group. But the number of risky sexual acts decreased by the same amount in the comparison group. Clearly, risk had been reduced all around, but there were no differences between the two study arms. (With dramatic downward shifts in risk in all five countries, it seemed unlikely that anything other than the trial itself had effected the risk reduction.) The results surprised Celentano. Because the intervention known to work seemed no more efficacious in the trial, it might be cast aside.

What happened?

The standards for ethical conduct of trials had changed, says Celentano, the Charles Armstrong Chair in Epidemiology at the Bloomberg School. In the intervening years, many new interventions had been proven effective, and ethical obligations required that they be offered to the control arm of the later trial. Celentano lists the services offered to participants: educational materials, free condoms, HIV and STI testing and treatment, pre- and post-test counseling, and extensive interviews about risk behavior. “And that’s just the control group,” he says.

The stakes for finding effective HIV preventions are high: An estimated 33 million people live with HIV, another 2.6 million are newly infected every year, and 1.8 million die of AIDS annually. Now 30 years old, the field of HIV/AIDS prevention draws scores of researchers who spend billions of dollars in a race to find ways to prevent transmission. Some want to identify biomedical interventions, such as microbicides, vaccines and male circumcision. Others are counting on behavior change programs—safe sex education, peer counseling, media campaigns—to slow down the epidemic.

But the field of behavior change, in particular, is tricky terrain for evaluation. As new interventions are shown to be effective, ethical obligations and aspirations evolve, making evaluations more challenging. And as real-world HIV/AIDS conditions become more complex, RCTs begin to seem less feasible.

Do the Right Thing

In HIV prevention trials, investigators compare the incidence of new infections among the control arm to those among the intervention arm. For an intervention to be deemed effective, the intervention arm must show significantly fewer infections than the control, which reveals the intervention’s impact. “No researcher wants people to get infected,” says Maria Merritt, a core faculty member in the Johns Hopkins Berman Institute of Bioethics, “but the expectation is that some participants will get infected.”

Researchers and ethicists agree: Investigators have an obligation to minimize risk to all participants. And in this field, there is an acceptable level of protection for all participants in a trial that has been generally agreed upon by experts. This standard of prevention guides what is offered to participants—known as “the prevention package”—in order to minimize risk. (Prevention packages often include counseling, testing and treatment, like the suite of services offered in Celentano’s trial.) But, notes Merritt, PhD, an assistant professor in International Health, some people would add that there is an obligation to maximize benefits for participants. This might take the form of making available as many known effective interventions as possible to all participants.

Not surprisingly, robust prevention packages tend to dilute the results of a trial. Overall, there may be fewer participants with new
Trials of behavior change interventions are expensive. With full-monty prevention packages, the results may be unimpressive. “Then the basic science folks say, ‘See, behavior change doesn’t work,’” says David Celentano.

HIV infections, which is a wonderful thing. “The paradox,” says Jeremy Sugarman, deputy director for medicine at the Berman Institute, “is that the more effective the basic package, the less likely that the research question will be answered.”

When faced with this challenge, some investigators and bioethicists invoke the “real world” or the notion of “usual care.” In theory, the control arm of a trial represents the local standard of care in the setting. “What is usual care in some developing countries? Nada,” says Celentano, ScD ’77, MHS ’75. “You want to do what’s right for the community, but these ethical imperatives make the science incredibly hard to do.”

Community trials of behavior change interventions are expensive and take years. With full-monty prevention packages, the results are sometimes unimpressive. “And then the basic science folks say, ‘See, behavior change doesn’t work,’” Celentano says.

Sugarman, MD, MPH, MA, who chairs the Ethics Working Group of the HIV Prevention Trials Network, cites ethical and pragmatic reasons for examining carefully the standard of prevention. Ethically, it might be irresponsible to introduce a prevention package that cannot be sustained or implemented locally after the trial has finished. Practically, he argues, it’s important to remember that just because something works in one setting doesn’t mean it will work in another setting. For example, while an antiviral drug may be useful in preventing HIV among men who have sex with men, it remains unclear whether the same drug will be effective in preventing heterosexual transmission. To include that in a prevention package among heterosexuals might be premature and presumptuous.

Yet another challenge is the adding on of new interventions midtrial as they are shown to be effective—to maximize benefits to participants. “I worry that the ethics discussion got ahead of itself by not asking what we need to know before adding new interventions. You can’t always change a study design on the fly,” says Merritt.

“Heaping on interventions in the prevention package isn’t necessarily the right thing to do,” says Sugarman. “You need to make sure there’s an explicit reason for doing so and that there is reason to assume that more will necessarily be better.”

The Gold Standard
For the last 30 years, RCTs have been considered the gold standard in evaluating HIV preventive interventions. The RCT has been linked in our hearts and minds with the term “evidence-based medicine,” and thus, argues Steve Goodman, a core faculty member in both the School’s Center for Clinical Trials and the Berman Institute of Bioethics, the gauntlet has been thrown down. To be taken seriously, an intervention must prove itself in a randomized trial.

In biomedical interventions such as male circumcision, designing an RCT is challenging enough. In the circumcision trials, participants consented to take part in the trial without knowing if they would be assigned to the intervention arm (circumcision) or the control arm (circumcision only after it was shown effective). Because the assignments are randomized to reduce bias, neither the participant nor the provider has any say in who gets what done. “One could argue that the reason clinical trials are a new entry in the tools of medical investigation is that doctors and patients couldn’t countenance the notion of randomization,” says Goodman, a professor in Epidemiology. “There’s an ethical calculus in every trial.”

But at least with a biomedical intervention—circumcision, for example—investigators can more or less control both arms. One group gets the circumcision; the other doesn’t. It’s either yes or no, cut or uncut. No gray area.

Behavior change interventions, on the other hand, create gray areas. Take the example of a mass media campaign that encourages people to have fewer sexual partners. The first complication is that there’s no way to control who sees a billboard or hears a radio commercial; diffusion is inevitable.

Deanna Kerrigan, who directs a USAID R2P (Research to Prevention) project to evaluate HIV/AIDS intervention programs under way in several African countries, finds diffusion to be one of the many challenges in evaluating these types of interventions. “With a pill, you know—yes or no,” she says. “When you deal with a mass media campaign, it’s impossible to say this group got it, this group didn’t.”

There’s another wrinkle: many behavior change interventions are aimed at communities, not at individuals. (Individuals get circumcised; communities get the billboard.) And tracking outcomes in individuals is much more cut-clean than tracking outcomes on a village or town level.

Historically, the government agencies or NGOs that implement interventions such as media campaigns evaluate their own programs. Sometimes programs overlap, often inefficiently, with several interventions targeting the same people, and the effectiveness of the interventions is at the mercy of social factors such as migration or civil war. But, to avoid any perception of bias, independent evaluations should be made.

(continued on page 46)
Behavior on Trial (continued from page 43)

The goal of Kerrigan’s project is to provide objective, academic evaluation. She and colleagues will study dozens of interventions already under way in various villages and districts, which are implemented by dozens of different partners. Randomization is out of the question: in such complex situations, it’s impossible to conduct a conventional RCT. It’s messy, says Kerrigan, an associate professor in Health, Behavior and Society. “But this is real-world stuff,” she says. “We can’t slow down this train.”

Almost Gold?
The best she and her colleagues can hope for, says Kerrigan, is to identify solid control arms and conduct a valid observational study [see sidebar]. But there is some help for the herculean task ahead for her and other researchers. Where randomization is not possible, recently introduced statistical tools such as propensity score matching may help close the gap between observed and randomized trials, compensating somewhat for the loss of comparability between study arms when randomization is not possible. (Propensity score matching is one of many statistical tools used to make intervention and control arms similar enough to allow fair comparisons.)

“We have statisticians bringing some observational studies very close to RCTs in terms of confidence in their findings,” says Goodman, MD, PhD. Adds Kerrigan, PhD, MPH, “We want the most rigorous design feasible.”

But Kay Dickersin, who directs the Center for Clinical Trials at the Bloomberg School, advises caution when relying on observational studies to determine the effectiveness of new interventions. “Certainly we should use whatever data we have. But we’ve seen major mistakes that make us shy about using observational data” to determine intervention effectiveness.

The hormone replacement therapy controversy is a good example: Many observational studies showed a cardiac benefit for postmenopausal women treated with estrogen plus progesterone; when the treatment was tested in the context of a large RCT, however, it showed no cardiac benefit, and perhaps even a higher risk of heart disease. “That trial was like, ‘Oops, we blew it.’ … I’m intrigued that we might be able to emulate RCT findings using observational studies and special statistics. I’d like to see studies validating the modeling alternative, comparing findings to those of RCTs,” says Dickersin, PhD, professor in Epidemiology. “As far as I know, it’s still an open question.”

Goodman notes that a recent reanalysis of the largest observational study on estrogen showed that the observational results were quite similar to the clinical trial’s.

What’s an Investigator to Do?
Albert Einstein is credited with saying, “Things should be as simple as possible, but no simpler.”

Ethicists believe that studies should be designed to accommodate ethical obligations, and investigators agree. Celentano thinks that a larger sample size would help: “With reduced risk in the control arm, the anticipated difference between the two arms is smaller than anticipated, and so you need a larger sample size to demonstrate that one arm’s intervention is more effective than the other. … But we can’t always afford a larger sample size.”

Goodman, MD, PhD, suggests that in some cases, an RCT might be overkill; sometimes what we learn from an observational study is good enough. Dickersin acknowledges that “randomized-controlled trials are hugely expensive, and they can only address one question at a time, usually over the short term, whereas with huge data sets you can address multiple questions and rare, longer term outcomes. [But] I think we better be very careful if we rely on observational data to determine the effectiveness of an intervention.”

There’s no right or simple solution for measuring the effectiveness of behavior change interventions in the field of HIV/AIDS prevention. The virus is wily, and the epidemic entrenched. What kind of evaluation is the best evaluation? Says Goodman, “What you’ll learn is defined by the purposes at hand. … You measure the risk, the cost, the stakes, the consequences of being wrong. … A clinical trial cannot be done in all situations, regardless of the stakes.”

MAY 31 IS WORLD NO TOBACCO DAY
HELP MAKE A HEALTHIER WORLD
GIVE TO THE INSTITUTE FOR GLOBAL TOBACCO CONTROL
JHSPH.EDU/GLOBAL_TOBACCO