Countries and health providers are following Iceland's path and combining health and genetic data on large populations. They promise to deliver "personalized" medicine, but will they?

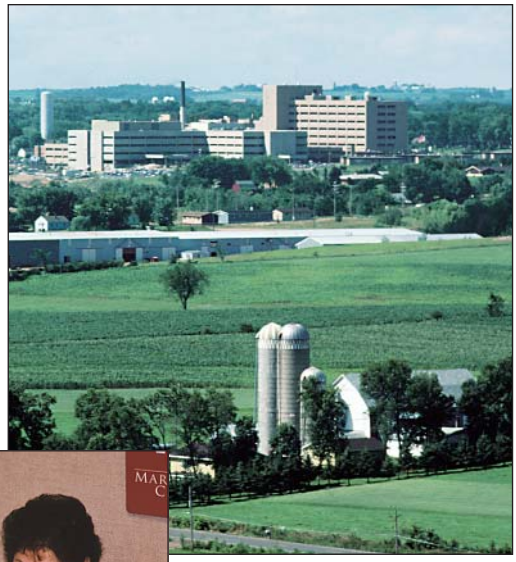
Population Databases Boom, From Iceland to the U.S.

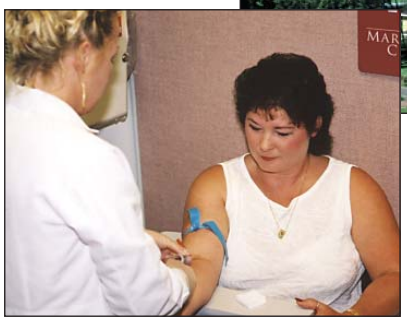
In August, residents of the dairy country of central Wisconsin received an unusual invitation from their local health care provider: an opportunity to donate their DNA for research. If they sign up, they will give blood and talk with a clinic staffer about their family disease history, diet, and exercise habits. The projected 40,000 participants will also give researchers extraordinary freedom to use this information—including details of their genetic makeup—to probe the complex interplay between genes, environment, and disease.

Once researchers have amassed a bank of blood samples, they will scan each subject's DNA for telltale markers of increased risk for various diseases. Ultimately, these data will be combined with the participants' electronic health records in a powerful new type of database. With a touch of a few keys, says Michael Caldwell, director of the Marshfield Medical Research Foundation, which will run the study, researchers will be able to mine the confidential data for links between genes, lifestyle factors, and illness. Caldwell's team hopes to find disease genes that have so far proved elusive and to sort out tricky epidemiologic questions, such as how much a particular combination of genes and exposures—sunlight, say, or drinking alcohol—is likely to raise the risk of cancer or heart disease. "The consensus is that such databases will be the key to unlocking the genetic basis of common disease," Caldwell says.

The project puts the medical clinic in Marshfield, Wisconsin, population 19,000, at the cutting edge of the new "genomic" medicine. It is in the vanguard along with countries such as Iceland and the United Kingdom, which believe that these new population databases are a sure-fire way to better health care. If the databases can find more disease genes and quantify risks, doctors believe they can then give patients personalized treatments and prevention plans. But ferreting out these links, say researchers involved, requires huge DNA collections, bigger than any gathered to date—some projects aim to sample a million people—plus long-term health data on each person who donates his or her DNA.

That's a dangerous combination, say some ethicists. They worry that data won't remain confidential and suggest that companies, which will play a role in some projects, should not be allowed to profit from people's genetic heritage. Indeed, Iceland's decision to give one biotech company, deCODE Genetics, exclusive rights to the nation's health records ignited a firestorm of controversy 4 years ago that continues even now.





Heartland biobank. This volunteer's DNA will go into a database to be probed by the Marshfield Clinic, which plans to study gene-environment interactions among residents of central Wisconsin.

CREDITS: MARSHFIELD CLINIC

And although many geneticists agree that these databases will yield a plethora of useful information, it is not clear that they will deliver on their most ambitious promises. "It's still mostly hype," says Stanford Law School ethicist Henry Greely. Nobody knows for sure, for instance, that bigger studies will be more successful at pinning down disease risks than previous, smaller studies. Researchers are also split on the best way to design population databases. "There are many, many opinions and not a lot of hard facts about this field," says human geneticist David Altshuler of Harvard Medical School in Boston and the Whitehead Institute in Cambridge, Massachusetts.

But that's not dampening enthusiasm. Encouraged by their native scientists, a half-dozen countries—as well as some U.S. health care providers—are laying plans to compile health information and collect DNA from a broad swath of the population. All are grappling with similar scientific and ethical issues.

From families to populations

Small DNA studies sufficed when the target was easier: a single gene that, when mutated, triggers a rare inherited disorder such as Huntington's disease. Common disorders such as arthritis or stroke—believed to be caused by defects in multiple genes in combination with lifestyle factors such as diet and smoking—pose a trickier challenge. Because each gene contributes just a small amount to overall risk, it emits a weaker signal, confounding efforts to find it. To compensate, researchers need to study genetic profiles of many more people and also incorporate information on phenotype, or health data.

One of the best-known ventures is Iceland's deCODE Genetics, which 5 years ago made the startling announcement that it proposed, under a contract with the Icelandic government, to put the health records of all 270,000 citizens into a single database. This health information—coded so it could not easily be traced back to individuals—would then be combined with Iceland's detailed genealogy and genetic data collected from volunteers. Under deCODE's 12-year license, drug companies could access the data for a fee; access would be free to academic researchers for "noncommercial" projects. Icelanders wouldn't learn their test results; the main benefit, supporters argued, was that the project would boost the country's economy.

SOME PROPOSED POPULATION DATABASES

Project	Company	DNA Sample Size	Budget	Status
Icelandic Health Sector Database	deCODE Genetics	280,000	$212M	Health database in 2003?; 80,000 DNA samples genotyped
Estonian Genome Project	EGeen International	1,000,000	$150M	3-year, $2.5M pilot (10,000 donors) began fall 2002
BioBank UK	?	500,000	$66M	Full enrollment in 2004
Marshfield Personalized Medicine	–	40,000	$3.8M+	Enrolling this fall
National Children's Study	–	100,000	?	Full study begins in 2004
Latvian Genome Database	?	60,000 pilot	$1.7M	Law passed in June; seeking funding
Quebec CARTaGENE	?	50,000+	$19M	Seeking funding

EXISTING BIOBANKS AND/OR HEALTH RECORDS

Västerbotten, Sweden	UmanGenomics	80,000	–	Data use agreement with county in 2002
Mayo Clinic	?	100,000	–	Prototype health database completed in July
EPIC	–	350,000	–	Pooling data for cancer studies through consortium
Nurses' Health Study	–	63,000	–	"
American Cancer Society CPS-II	–	110,000	–	"
CDC NHANES III	–	7300	–	Proposals to use individual data requested fall 2002

Researchers within and outside Iceland strongly objected. Perhaps the most contentious issue is that the project relies on "presumed consent": Government health records on every citizen are included in the database unless individuals specifically request otherwise. After safeguards were added to ensure privacy—for example, two government-appointed bodies will oversee encryption of data for research and database operations—the country voted to approve the project. About 7% of the population has opted out of the study.

DeCODE can't begin uploading medical records until it passes a final hurdle, expected next year: an outside expert's test of the database security system, says Icelandic health official Gudridur Thorsteinsdottir.

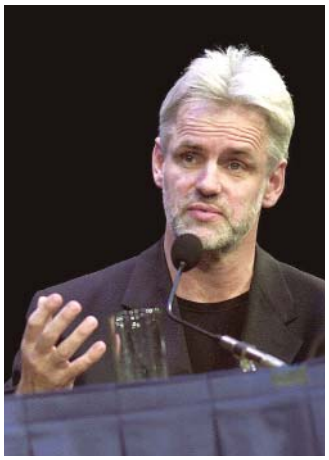
But in the interim, the company has compiled proprietary genetic data on a large chunk of Iceland's population by embarking on more traditional, although still ambitious, gene hunts for specific diseases. Through referrals from clinicians, deCODE researchers have identified 80,000 volunteers for these disease studies and have analyzed, or genotyped, their DNA (tagging at least 1000 markers on each genome). Already, the company says it has mapped or identified genes involved in arthritis, stroke, schizophrenia, and many other diseases and it is beginning to publish these findings.

Once the full database is ready, it can be used for new types of studies. "You will begin to see correlations you couldn't before," claims deCODE CEO Kari Stefansson—for example, whether a gene for diabetes also predisposes a carrier to hypertension or stroke.

Because Iceland's population is relatively homogeneous and has unique genealogical data, its power to find new genes might never be matched, says geneticist Stephen Warren of Emory University in Atlanta. But findings in Iceland won't necessarily apply directly to other ethnic groups and more diverse populations.

Beyond Iceland

Iceland's experience has informed the design of other population databases, such as one in Estonia. In September, the government-founded, nonprofit Estonian Genome Foundation began collecting DNA samples from 10,000 volunteers age 16 year and up. This 3-year pilot project, funded with $2.5 million by EGeen International, a U.S.-based company, will rely on a health questionnaire rather than medical records. Project founder Andres Metspalu of the University of Tartu, Estonia, who eventually hopes to enroll 1 million of the country's 1.4 million people, says that organizers have taken great pains to educate the public and allay ethical concerns. Participants can ask to see their genetic profile. And by feeding back to participants data that can be used in health care, Metspalu says, the project will give benefits "back to the people."



Maverick. DeCODE's Kari Stefansson started the trend in population databases.

The U.K.'s Medical Research Council (MRC) and the Wellcome Trust charity are planning to spend $66 million on a large cohort study with 500,000 volunteers (*Science*, 3 May, p. 824). "We're very different from Iceland in many ways," says Tom Meade of MRC.

Giving back. Andres Metspalu notes that Estonia's population database project will let participants see their own genetic profile.

Starting in 2004, BioBank UK plans to gather examination and interview data from volunteers 45 to 69 years old and then track them for at least 10 years. By starting with middle-aged volunteers—who aren't promised any direct benefit—the researchers expect to see enough cases of specific diseases to verify and quantify links with candidate genes. Access to BioBank UK will be open to "any bona fide researcher with a good idea," says Meade.

"I think it's a very exciting project," says cancer geneticist Nathaniel Rothman of the U.S. National Cancer Institute (NCI). He says with 500,000 samples, BioBank UK will be the largest population database in the world, and it will draw on resources—such as participants' national health care records

CREDITS: (TOP TO BOTTOM) ESTONIAN GENE BANK; GINO DOMENICO/AP

Private Biobanks Spark Ethical Concerns

As countries and health care providers rush to create huge new databases linked to genetic data (see main text), an increasing number of companies in the United States are amassing large, entirely private, "biobanks" of DNA and tissue samples. The companies either use this material themselves or sell it to drug companies or academic researchers for studying diseases, say, or tailoring drugs to individuals.

The companies assure donors that their privacy and interests will be protected, but some bioethics experts caution that such safeguards are hard to verify. Companies aren't necessarily subject to regulations, notes Mary Anderlik, a bioethicist at the University of Louisville, Kentucky, and "it's an extremely spotty area."

"The main issue is whether people [who donate samples] know what they're getting into," says George J. Annas, a law professor at Boston University.

One of the biggest commercial biobanks is Genomics Collaborative Inc. of Cambridge, Massachusetts, which claims to have health data and tissue and DNA samples from 120,000 people from all over the world. The firm maintains links with donors so it can update its "gold standard diagnostics," says medical director Kevin Krenitsky, but donor identity is not revealed. The company works with physicians to recruit patients to donate samples; Ardais Corp. of Lexington, Massachusetts, which also boasts 120,000 samples, contracts with major medical centers for disease tissue. Another company, DNA Sciences Inc. of Fremont, California, has collected 3000 of its 18,000 DNA samples from volunteers recruited through a Web site.

These companies say they provide secure data systems for ensuring patient confidentiality and voluntarily follow federal rules for protecting human subjects. But Anderlik says these procedures are often shrouded in secrecy. Some private biobanks, for example, consider key documents such as consent forms proprietary. And if companies go bankrupt, critics contend, tissue and DNA samples might be sold to practically anybody. Last year, for example, a court in Japan auctioned off a human cell collection that a scientific society had used as collateral on a loan.

Some U.S. states have laws specifying ethical standards for using DNA samples, but federal regulations are a "patchwork," Anderlik says. She favors federal licensing of biobanks, with an exception for short-term collections for specific studies. Iceland enacted such a law in 2000, but no such legislation is pending in the U.S. Congress. **–J.K.**

for prescription histories and comprehensive disease registries—that are "just not available in the U.S."

To the east in Latvia, researchers in June got parliamentary approval for a planned pilot database. Even scientists in Germany, which has been wary of some areas of genetic research, are contemplating an Estonia-like project, says Spiros Simitis, an ethics law professor at the University of Frankfurt. Researchers in Quebec are seeking funding for a $19 million, 5-year project that would initially enroll 50,000 adults, says Claude Laberge of the University of Laval in Quebec City. And Singapore is taking the first steps toward a population database with five new disease registries and a linked cancer tissue databank (*Science*, 30 August, p. 1470).

Pros and cons

Proponents believe these databases will be a gold mine for improving health care. Identifying the genes involved in common diseases will eventually yield new treatments, they say. And quantifying genetic risks—for instance, how much a certain combination of mutations ups the risks of cancer—could help patients decide whether to have invasive procedures, such as a colonoscopy. Companies could use these databases to design drugs suited for an individual's genetic profile.

Some of these goals are out of reach today, database designers concede. Finding new disease genes, for instance, requires scanning the entire genome for markers. But the cost—10 cents per marker, when 50,000 markers per person might be needed—is prohibitive, says Metspalu. He and others are banking on technological advances—at least a year away—to lower the cost to 1 cent per marker, as well as a new kind of genome map that will reduce the number of markers needed.

Bottleneck. The cost of sequencing DNA samples like these has to drop considerably before gene-discovery studies in large populations will be affordable.

CREDIT: MARSHFIELD CLINIC

Yale geneticist Kenneth Kidd sees another obstacle: The databases will be only as good as the individual clinical or exposure information they contain. "The quality of diagnosis is a sine qua non of doing these kinds of studies," Kidd says. "Are these individuals going to be well worked up?" Opinions vary over whether a routine exam and a patient's health record are sufficient, or whether more detailed measures are really needed—such as insulin metabolism tests to study diabetes.

Harvard epidemiologist Walter Willett has a more fundamental complaint. "We already know that most variation in human disease is due to diet and lifestyle factors," he says, and quantifying how the risks vary with one's genetic makeup usually won't change the solution: encouraging healthier lifestyles. Willett worries that the zeal for genomic medicine will divert resources from prevention (*Science*, 26 April, p. 695).

Willett is also part of a camp that argues that new population studies could be reinventing the wheel, because existing studies with DNA samples could provide similar information (see table). Funded by NCI, he and colleagues are pooling data from many large cohort studies, such as Harvard's nurses and physicians studies and EPIC, a European cancer study; the combined database will have more than 1 million DNA samples for cancer research. True, there are hurdles to studying additional diseases: Participants might have to be tracked down for fresh DNA samples or new informed consent. But Willett thinks that these efforts, as well as new population databases, should be supported.

And 4 years after deCODE sparked international debate on population databases, ethical questions still loom large. One issue is "how much of a blanket consent you can create" for studying unspecified diseases, says Wylie Burke of the University of Washington, Seattle. Meade says there is "still a lot of discussion" about whether BioBank UK participants should be able to give consent only for specific diseases. Estonia and Marshfield will rely on ethics review boards to decide if new informed consent is needed to undertake potentially controversial studies—on behavior, for instance.

Despite claims to the contrary, some critics charge that privacy is still not assured. Jane Kaye, a doctoral student at the University of Oxford, U.K., says that although Iceland's data system is "quite tight," BioBank UK has not yet outlined a plan that will adequately protect data. The role of companies, which is still in flux, remains contentious: The advocacy group Human Genetics Alert, for instance, is opposed to allowing companies to patent findings from BioBank UK.

Biobanks, American style

Genomics leaders in the United States think the benefits of population databases will likely outweigh these risks. But federally funded projects are still in the early planning stages. At the National Institutes of Health, officials are thinking about a project like BioBank UK but even bigger, says Lisa Brooks of the genome institute there: "Something that looks at a lot of people and a lot of diseases. Something that's big and pretty comprehensive."

The obvious way to create a large population database in a country without a national health care system is to work with health care providers, as Marshfield is doing on a small scale, says Stanford geneticist Neil Risch. Indeed, in some ways, the Marshfield Personalized Medicine Research Project is out in front, because Marshfield Clinic—whose research foundation is conducting the study—already has electronic health records on more than 1.2 million patients and began collecting DNA samples this fall. Patients won't learn their results, but they will help advance health care in general, the clinic tells donors. The project has strong support in Wisconsin, where the state has contributed $2 million of $3.8 million in initial funding. Although the nonprofit clinic expects to patent discoveries, it will funnel any profits back into research or donate them. Companies will not be directly involved: "The hope is to keep funding in the public domain and have this become a national resource," Caldwell says.

Some other health care providers are also moving ahead on their own: The Mayo Clinic is building a database of the health records for 4 million of its patients and members; it plans eventually to add genetic data stored in the clinic's many tissue banks. A research database is also "in the early discussion stages" at Kaiser Permanente's division in Northern California, which has 3.1 million members, says Kaiser Permanente epidemiologist Cathy Schaefer.

But U.S. researchers are proceeding cautiously, wary of running into the controversy that Iceland's deCODE and other projects have encountered. Says Risch: "We're not going to have many opportunities. It will be very expensive, and it really needs to be done right."

–JOCELYN KAISER

PROFILE BRIAN TUCKER

Bracing for the Shocks Of the Future

From Ecuador to Nepal, a geoscientist leads a hands-on crusade to help city dwellers survive earthquakes

The news photos from San Giuliano di Puglia were heart-wrenching: a rainbow of backpacks brightening the cold ground; piles of concrete rubble strewn amid unscathed apartment buildings. Until last week the litter was an elementary school. Then an earthquake struck southern Italy, leveling the building and killing dozens of its youngest pupils.

To Brian Tucker, the horror was all too familiar. He spends his days trying to stave off such disaster. Tucker heads GeoHazards International (GHI), a nonprofit group that helps the developing world brace for earthquakes. With a staff of five and a few phones based in spartan offices above a tuxedo rental shop in Palo Alto, California, GHI has taught masons how to strengthen schools in Quito, Ecuador, and in Katmandu, Nepal. The group has counseled community leaders about quake risk across Latin America and Asia. And there is plenty of risk: Over 85% of the world's quake-prone cities are in developing countries, often overlooked by researchers.

At 57, Tucker, a seismologist, looks like a softer version of actor Anthony Perkins. He's tall and lanky, with a sharp nose, brown eyes, and short hair in need of a comb. Tucker chooses his words carefully, although his voice cracks with enthusiasm. Colleagues call him a dreamer with a practical streak. "Brian is completely committed to the safety of people we'll never know," says L. Thomas Tobin, a GHI adviser and earthquake consultant in Mill Valley, California.

That commitment has begun to pay off. In September, Tucker was named a MacArthur Fellow, winning a $500,000 genius award for his work at GHI. Still, Tucker describes his success as "a very slow creep." And there have been surprises along the way. "When I started, I felt I could do this because I knew something about seismology in California," Tucker says. "That took me about a tenth of the way."

Catalyst for change. Seismologist Brian Tucker left research to save lives in developing countries.

CREDIT: GEORGE NIKITIN/AP

A formula for success

Tucker's journey began 27 years ago, with a cup of tea. In 1975, while a postdoctoral associate at the Massachusetts Institute of Technology, he studied quake prediction with Soviet scientists in Tajikistan. One afternoon, he visited a Tajik villager at home. Sitting inside the adobe house, Tucker noticed telephone pole–like posts supporting each corner—a uniquely sturdy design. "My dad was a real kook, and he thought houses should be built this way," his friend shrugged.

In fact, the posted house had withstood a 1949 quake that crushed neighboring homes with snow and rock. Impressed, Tucker quizzed his Tajik friend: How would *he* build a house now? The startled reply: "In the traditional way, of course! I'm not nuts." That exchange, Tucker says, convinced him that earthquake defense is about belief as much as building codes: "The moment was seared into my brain."

Tucker brought that lesson to his next job, a 7-year stint heading the geohazards program for what is now called the California Geological Survey. While he was improving quake readiness at home, however, his thoughts often traveled. "Year after year, conference after conference, my colleagues and I would shake our heads and say, 'Gosh, the real problem is in developing countries,' " Tucker says. "I was almost 45 when I began to think I'd spend the rest of my life saying, 'Somebody should help these people.' "

Finally, disaster jolted Tucker into action. In the winter of 1988, an earthquake struck Armenia, killing over 30,000 people. A year later, a comparable temblor—the 1989 Loma Prieta quake—hit the San Francisco Bay area. It killed 66. Haunted by the asymmetry, Tucker left his job, enrolled at Harvard University to learn public policy and business basics, and then leased a San Fran-