Mobilizing BIG DATA to Understand Mobility

By Esther Landhuis

In a win-win for patients and researchers, big biomechanics data has arrived.

One rainy day in the late 1990s, a biomechanics doctoral student stood sopping at the street corner, hoping his research participant would arrive for her gait analysis. "It was the third time she stood me up. I was sick and tired of waiting for people to come to me," says **Reed Ferber**, **PhD**, recalling his graduate years at the University of Oregon. "I wanted to create something where people stream in all the time and we have access to their data at our fingertips."

That yearning eventually birthed the Running Injury Clinic, which Ferber now heads as an associate professor of kinesiology at the University of Calgary, Canada. Unlike his graduate research subjects who grudgingly "gave two hours of their time and didn't get much out of it," he says, hordes of runners have flooded into the clinic, eager to fork over several hundred dollars to learn how they got hurt and what they can do about it. It's a win for data-hungry researchers, too. Cameras record every twist and turn of the runners' ankles, knees, and hips as they walk or jog on treadmills. All that data flows straight into a vast 3-D motion-capture repository, which stores data from the Calgary clinic as well as some 30 gait analysis centers around the globe.

It's a sure sign that biomechanics labs are starting to tap into the big data revolution.

Just as Amazon uses big data to deduce what a consumer might crave, or genomics researchers use it to identify genes that cause a specific disease, so too can biomechanics researchers like Ferber benefit from big data—mining it to distinguish between patients who will and won't respond to a particular sur-

gery or physical therapy treatment.

Currently, biomechanics datasets are growing, and big data methods to tackle those mountains of information are still under development. But there's enough of a track record to suggest the big data revolution can help injured runners get back on the road, keep people moving to fight obesity, and transform how physicians treat patients with movement disorders, such as cerebral palsy.

It will take hard work to overcome some of the challenges of biomechanics big data. More researchers need to share the data they have, some of which may be stashed on hard drives of individual graduate students. The heterogeneous nature of the data will require analytical tools that don't yet exist. Then there's the enormous and looming opportunity to take advantage of data from wearable sensors that track movement patterns of millions of users each day. But if these challenges can be overcome, the river of data could offer tremendous insight about human movement.

Big Data Challenge #1: Amassing Data

Ferber is not the only scientist who has longed for more research subjects. Most standard mobility studies involve just 10 to 100 people—hardly enough to populate a database for statistically meaningful comparisons. "To have over a million subjects is completely unheard of," says **Scott Delp**, **PhD**, professor of bioengineering at Stanford who is leading the University's effort in mobility data analysis.

To address that problem, Delp and his colleagues have established a consortium of four major clinical centers with mobility data from more than 20,000 people. "Over the last 10 years, we've been changing the culture of motion analysis labs. They feel comfortable and safe sharing their data," Delp says. He plans to expand the repository to include biomechanics data for over 10 million people.

Outside the controlled lab setting, **Ray Browning**, **PhD**, director of the Physical Activity Energetics/Mechanics Lab at Colorado State University, is four years into a five-year NIH-funded project that is accruing a massive quantity of data using sensors worn during what's called "free-living physical activity." His team is trying to determine if playground renovations and activity-promoting recess curricula encourage kids to be more active. They've fitted Denver area schoolchildren with wrist accelerometers that track their movements 75 times per second over a six-day period. Thus far, the researchers have gathered seven to eight gigabytes of mobility data from about 2,000 children—totaling in the terabyte range. "You could never process one of these files on a personal computer because there isn't enough memory to load it," Browning says.

In addition to big data collected by labs and free-living researchers, the field of biomechanics stands to benefit from a new frontier—mobility data captured daily by millions of people using wearable sensors such as FitBit and Jawbone as well as smartphone health apps such as Azumio's Argus. "Even if you pooled data from all the gait analysis labs in the world, it wouldn't be as big as the database that FitBit already has," says Browning. "People are wearing these devices. I don't have to recruit them. It removes a huge hurdle."

Big Data Challenge #2: Heterogeneous Data

With all this data in the pipeline, biomechanics researchers have to address one of the hallmark challenges of big data: its diversity.

In studies of gait abnormalities due to cerebral palsy (CP)—the most common movement disorder in children—researchers gather multitudinous data from patients: They stud their arms and legs with reflective spheres that light up high-resolution images captured by video cameras, use electrodes to track electricity patterns in their muscles, and attach sensors to measure forces generated against the ground each time the child takes a step. In the end, doctors scrutinize video clips and pore through several years' worth of graphs and charts to make their best guess as to how the patient might fare given a particular surgery.

It would be better, Delp says, if researchers could take all that data and compare it to a database of similar information for other patients. They could then identify which surgical treatment worked best for people whose muscle patterns and motor activities resemble those of the patient at hand.

Recent research suggests the challenges of integrating these multifarious data types are not insurmountable. Stanford mechanical engineer **Jennifer Hicks**, **PhD**, and colleagues have created biomechanical models that make useful predictions by integrating several different types of data—muscle strength measurements collected during physical exams and biomechanical data from motion-capture computer simulations. When applied to a group of children who had trouble walking due to cerebral palsy, the algorithms identify which patients will benefit from surgery 70 to 80 percent of the time. "It's not great—you'd like to be at 95 percent—but definitely better than human experts," says Delp. Research led by **Mike Schwartz**, **PhD**, associate professor of orthopedic surgery at the University of Minnesota—and reported this issue in *Computation in the Surgical Suite*—also suggests that modeling and simulation can help predict surgical outcomes.

Chris Re, **PhD**, an assistant professor of computer science at Stanford, and colleagues are creating computer algorithms that integrate heterogeneous biomedical data from motion-capture video, smart phones and other sensing platforms. They are borrowing strategies they've already applied in paleobiology, where they designed a machine-reading system that amasses and integrates fossil data more effectively than databases manually compiled by experts.

Calgary Running Clinic developed a 3-D gait analysis system to gather motion capture data. It is gathering big data not only at the Calgary clinic but also at about 30 clinics and universities around the world, according to 3dgait-analysis.com. Courtesy of Reed Ferber.

The onslaught of data from wearable sensors and smartphones adds yet another dimension, and collaborations are beginning to emerge between companies producing these devices and researchers who use high-resolution imaging and motion capture. For example, Stanford computer scientists are working with researchers at Palo Alto–based Azumio to extract biomechanics data from the company's mobile health app, Argus. Tracking 500 million steps per day, as well as heart rate and sleep patterns, Argus is "like FitBit on your smartphone," says **Bojan Bostjancic**, **PhD**, Azumio's CEO and founder.

Granted, smartphone recordings are noisier than high-resolution video footage from lab cameras. But "it's extremely cheap data from millions of people 24/7," says data mining expert **Jure Leskovec**, **PhD**, assistant professor of computer science at Stanford and a Delp collaborator. "We can connect these types of data so they complement each other. That is our aim—to take high-quality data from labs and integrate it with large-scale noisier data from cell phones."

Big Data Challenge #3: Extracting Insight

Once big data is in hand and integrated, the task of learning from it can begin, as it already has. In a recent study, Ferber and his team compared the gait patterns of 34 people with osteoarthritis to a large dataset of gait information and treatment outcomes for others with the disease and other musculoskeletal injuries. This big data approach allowed them to predict, with 94 percent accuracy, which participants would benefit from a six-week exercise program. As they forecast, nearly a third did not improve from the strength training, suggesting alternative treatments should be found. Ferber's team will present its findings this July at the World Congress of Biomechanics in Boston.

The gait analysis database has also helped physical therapists give better advice to injured runners. By comparing clients' gait data to that of runners in the database with a similar demographic profile (e.g., male elite runner), experts can not only confirm someone has patella-femoral syndrome (aka "runner's knee") but also tell them it's because their hip abductors are weak, for example. "The data help us go beyond relieving symptoms. They direct us to the source of the problem," says **Blair Shular**, head physiotherapist at Glen Sather Clinic in Edmonton, Canada, where 156 patients have undergone 3-D gait analysis since the clinic purchased the $25,000 system in early 2011.

Making use of movement data uploaded into the cloud by millions of people using wearable sensors and smartphone health apps will require much more powerful analysis tools, Browning says. But the payoffs could be staggering. He and others envision using wearable sensor data to encourage healthy physical activity in people at risk for obesity, or to warn runners of impending injury.

Many runners won't come in for gait analysis until they are hurt, Ferber notes. "We need to be able to tag them like wildlife, monitor them in their natural habitat, and analyze that data in a robust way to offer simple information like 'you're doing fine' or 'you're on the verge' or 'you'd better stop,'" he says.

Likewise, in studies of movement disorders, large datasets from wearable sensors might allow scientists to spot a pattern of shifts in center of mass just before people fall, and then use that insight for further research into fall prevention, Browning says.

And Delp thinks that large-scale, unprocessed mobility data from Argus and other wearable technology could help scientists understand the root causes of knee osteoarthritis and joint pain. "Instead of coming into the motion analysis lab once a year, they'd wear the unit every day for a year," he says.

One thing's for sure: Biomechanics researchers no longer have to stand in the rain to gather data. It's coming at them in a deluge. And they are ready for it. □

Fitbit, a wearable sensor (below), and Argus (right), a mobile phone app, are gathering data from millions of people. Biomechanics researchers are planning to tap into that treasure trove.