

## Rethinking barriers to big data

When it comes to investigating what ails us, thinking outside the box means looking beyond our national borders. Big data — the kind of diverse, population-wide information that can shed new light on the underlying nature of health and illness — simply will not be big enough if we limit ourselves to Canadian sources.

For Bartha Knoppers, director of the Centre of Genomics and Policy at McGill University in Montréal, this approach is becoming a defining feature of 21st-century medicine. Speaking to an audience in Ottawa as part of the recent Big Thinking lecture series, she said, “You can’t have good science without some sort of solidarity and sharing of information. You certainly can’t have personalized medicine.”

According to Knoppers, who holds a Canada Research Chair in Law and Medicine, the path to this kind of medicine starts when medical researchers stop relying on volunteer participants. “The problem is that people select themselves into research,” she said at the Oct. 30 lecture. “We need to have a broader database where we’re getting all different types of ethnicities, of age groups, and so on.”

We are beginning to appreciate that health is not only determined by what can be measured in a doctor’s office, but by such extraneous factors as whether we take public transit, how often we fly on airplanes and how much we use electronic media, Knoppers said. She argued that medical researchers now have access to information technology that can enable them to extract practical insights from large databases, obtaining findings that are more representative of society as a whole, as opposed to a small, arbitrarily chosen subset.

Above all, everyone in our society would stand to benefit from comprehensive mining of as many databases as can be secured.

“Since everyone who is a citizen can become a patient, that is how data should be collected.”



Les photos de Penelope (www.photospenelope.com)

**Everyone can potentially benefit from comprehensive mining of international databases, says Bartha Knoppers, Canada Research Chair in Law and Medicine.**

Such talk usually raises fears of privacy violations, perhaps leading people to being denied jobs or insurance coverage because of their known genetic traits. Although she did not deny that prospect, Knoppers said that ongoing expansion of social media shows that the bar for what constitutes loss of privacy is being set ever higher. Nor has there been a litany of complaints about medical research compromising individual privacy, even though it could happen.

“That is no reason to not build infrastructure or have the quality data that we need for medical care,” she said.

Canadians are not alone with these concerns. Countries around the world are struggling with regulations that would free up useful information about our lives in ways that further the goals of research and provide legal protection

for individuals. All are motivated by the prospect that our big data will get even bigger if we can learn how to share detailed medical information as easily as we share the myriad other forms of data that criss-cross the globe at an accelerating pace and volume.

Where it might take years to round up a few thousand people for a study in Canada, several hundred thousand participants could be added if databases in the United States and the European Union could be incorporated.

Knoppers described how she finds herself caught up in efforts to wrangle this kind of data into forms that can help us deal with the challenge of access. Among the most ambitious of these initiatives has been the Global Alliance for Genomics and Health, an international coalition founded in 2013 that now includes more than 170 institutional members. This body’s mandate stems from Article 27 of

the 1948 Universal Declaration of Human Rights, which guarantees all people the right to share in scientific advancement and its benefits.

It may sound paradoxical that medicine will become ever more personal as we add more and more people to these massive databases, but this approach, which goes by the formal name of “stratification,” breaks down groups of people into categories to determine more precisely what makes us sick and what keeps us well.

“We need to understand those influences, which means sound data from everywhere,” Knoppers concluded. “The more data we can put together, the more we can move from populations to subpopulations to personalized medicine.” — Tim Loughheed, Ottawa, Ont.

CMAJ 2014. DOI:10.1503/cmaj.109-4942