Artificial intelligence joins hunt for human–animal diseases

Model predicts that the midwestern United States and Central Asia are at high risk for outbreaks of rodent-borne pathogens.

Boer Deng

18 May 2015

Stone Nature Photography/Alamy

The Northern flying squirrel carries diseases that can pass from animals to humans.

Lyme disease, Ebola and malaria all developed in animals before making the leap to infect humans. Predicting when such a 'zoonotic' disease will spark an outbreak remains difficult, but a new study suggests that artificial intelligence could give these efforts a boost.

A computer model that incorporates machine learning can pinpoint, with 90% accuracy, rodent species that are known to harbour pathogens that can spread to humans, researchers report this week in the *Proceedings of the National Academy of Sciences*¹. The model also identified more than 150 species that are likely to be disease reservoirs but have yet to be confirmed as such.

Central Asia and the midwestern United States were among the regions with the greatest concentrations of potential reservoir species. That surprised lead author Barbara Han, an ecologist at the Cary Institute of Ecosystem Studies in Millbrook, New York. "I had thought, 'Didn't we discover everything here already?'" she says.

Han and her colleagues began by training their model to identify characteristics that are common to the 217 rodent species that are known to carry zoonotic diseases. They set the model to analyse databases of animal traits, such as species' geographic ranges, reproductive behaviour and whether they are a reservoir for any zoonotic disease.

The model repeatedly divided species in those databases into groups based on arbitrarily selected traits — searching for patterns about which factors make a species more likely to carry pathogens that can infect humans. Eventually, the model developed a set of rules that could identify known carrier species with 90% accuracy.

Then the researchers used their model to analyse the disease-transmitting potential of the world's 2,277 rodent species. They identified more than 150 species — including some voles, squirrels and guinea pigs — that are not known reservoirs of zoonotic diseases but seem likely to be, based on factors identified by the model.

Felicia Keesing, a biologist at Bard College in New York, likes the study's use of animal traits such as litter size or frequency of

reproduction to identify disease risks. That could lead to more targeted disease surveillance than previous work that focused on the locations of past outbreaks. "We can predict not just where, but in what" species the next illness could arise, she says.

Sophisticated machine learning is beginning to be used more widely across ecology. Such work has uncovered hidden patterns in everything from the dispersal of invasive plants² to the flight paths of birds³.

Studying ecology mostly used to be about going into the field and gathering data, says Reuben Keller, an invasive-species researcher at Loyola University Chicago in Illinois. Now, he says, "ecologists are increasingly learning how to deal with all the data we've gathered in rigorous ways".

Nature | doi:10.1038/nature.2015.17568

References

1. Han, B. A., Schmidt, J. P., Bowden, S. E. & Drake, J. M. *Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1501598112> (2015).
2. Schmidt, J. P. & Drake, J. M. *PLoS ONE* **6**, e17391 (2011).
3. Guilford, T. *et al. PLoS ONE* **6**, e21336 (2011).