The mathematics behind the Ebola epidemic

October 9 2014

The highly contagious Ebola virus has been running rampant in West Africa for months – this photo was taken in Monrovia, Liberia. Credit: Ahmed Jallanzo / Keystone

Researchers in the Department of Biosystems Science and Engineering at ETH Zurich have calculated new benchmark figures to precisely describe the Ebola epidemic in West Africa from a mathematical perspective. Their results may help health authorities to contain the epidemic.

The Ebola epidemic in West Africa appears to be spiralling out of control. More than ever, local and global health authorities want to know how the epidemic will develop and, above all, how to prevent it from spreading further. Certain parameters help them to determine this, such

as the reproductive number, which is the average number of infections caused by a single infected individual. The incubation and infectious periods are also highly relevant; i.e. the time from infection to the onset of symptoms and the time from onset of symptoms to the clearance of the pathogen.

In the current Ebola epidemic, several estimates based on official data of recorded cases of illness were used to derive these figures. A team led by Tanja Stadler, professor of computational evolution in the Department of Biosystems Science and Engineering at ETH Zurich in Basel, has now calculated these parameters based on the gene sequence of the virus in various patient samples, using a statistical computer programme developed by the group.

Increase in unreported cases

The virus sequences were obtained by US-American, British and Sierra Leonean researchers from blood samples taken from patients in Sierra Leone in the first few weeks after the epidemic migrated to the country from neighbouring Guinea in May and June 2014. Newer sequences are currently not publicly available, says Stadler. From the data, the researchers calculated a viral reproductive number of 2.18. This value is in the range of the previous estimated values based on the incidence and prevalence of the disease, which are between 1.2 and 8.2.

"A major benefit of our method is that we can use it to calculate unreported cases and therefore the true scale of the epidemic," asserts Stadler. Official patient figures only take into account those cases reported to the health authorities. The actual number of infected persons is generally significantly higher. Using the data made available to them, the ETH researchers were able to calculate an unreported case rate of 30% (i.e. patients of which blood samples were not taken). "However, this applies only to the situation analysed in Sierra Leone in May and

June. We do not have any blood samples since June at all," claims Stadler.

Virus family tree created

The researchers were also able to calculate the incubation period for Ebola (five days – this value is subject to significant uncertainty) and the infectious time. Patients can pass on the virus from 1.2 to 7 days after becoming infected.

To obtain these values, the researchers created a phylogenetic tree based on the gene sequences of the virus samples. "The Ebola virus changes in the body of the patient from day to day, meaning that the virus sequence varies slightly from patient to patient," explains Stadler. With the knowledge of the different sequences, the researchers were able to determine at what point in the past infection events happened between patients. From this, they were able to calculate the epidemiological parameters.

Already tested for HIV

These epidemiological values are important in developing strategies to contain the epidemic and evaluate the effectiveness of these measures. Imposing a curfew is one measure. "If the curfew lasts longer than the incubation period, then only those people who continue to show symptoms of Ebola are carriers of the disease," says Stadler. In turn, the reproductive number is one of the most important benchmarks used by health services. The most pressing aim of these authorities is to reduce the reproductive number to a value lower than 1, as this would imply that the epidemic has been contained.

The ETH researchers developed the computer programme used to

calculate these figures during the past few years and applied it to data collected from HIV and hepatitis C patients. They now hope that new sequences of the currently circulating Ebola virus become available, despite the adverse conditions in the areas affected by the epidemic. As Stadler states, "our programme is ready. If we are given access to current Ebola sequences, we will be able to gain a detailed insight into the spread of the epidemic literally overnight."

More information: Stadler T, Kühnert D, Rasmussen DA, du Plessis L: Insights into the Early Epidemic Spread of Ebola in Sierra Leone Provided by Viral Sequence Data, PLOS Currents: Outbreaks, online publication 6 October 2014. currents.plos.org/outbreaks/ar … viral-sequence-data/

Provided by ETH Zurich