On 23 March 2014, the World Health Organization (WHO) released a report confirming that an Ebola outbreak was under way in West Africa. Prior to 2014, this highly lethal haemorrhagic fever was largely contained to Central Africa.

_Ebolavirus_ – to use its scientific name – was first documented among humans in 1976, in Zaire (Zaire is now known as the Democratic Republic of the Congo). Transmission of the virus has been traced to contaminated bushmeat, and while all Ebola outbreaks have zoonotic points of origin such as this, it can also be passed from person to person through contaminated bodily fluids. It is these human-to-human infections that typically determine the rate of escalation of the disease.

At the onset of the current outbreak in West Africa, transmission was relatively limited. Between March and May 2014, only a handful of cases were reported per day in Guinea and, later on, Sierra Leone (Figure 1). But in June, the deadly virus crossed the border into Liberia and everything changed (Figure 2). The outbreak began growing at an exponential rate, and
five months later (at the time of writing), it still is (Figure 3).

The West African Ebola epidemic is already 33 times greater in size than any other Ebola outbreak in the past, and it is not immediately obvious how much further it will grow or when it will end (Figure 4).

Because of this uncertainty, projection models that can predict the number of cases to come – as well as when and where they will occur – are critical for planning, preparedness, and prevention: the three core components of controlling an infectious disease outbreak.

Without knowing how many people will get sick in the coming weeks, it is impossible to know how many Ebola wards – including hospital beds, intravenous drips, healthcare workers, and personal protection kits – will be necessary to treat them. And without enough Ebola wards, the sick are inevitably turned away, which often results in transmission of the disease to friends and family once they have returned home.

Providing care to those who seek it therefore prevents further transmission at the community level, thus serving as an essential component of outbreak control. But accurately predicting the number of new cases is easier said than done.

Conflicting reports

The problem is not a lack of projection models – we have plenty of those, thanks to well-meaning computational epidemiologists at the WHO and the Centers for Disease Control and Prevention (CDC), as well as university-affiliated labs. Rather the problem is that these models spit out projections that do not necessarily agree with each other.

For example, in late September the WHO Ebola Response Team published its projection outputs in the New England Journal of Medicine, predicting 20,000 cumulative cases in countries with sustained transmission – that is, Guinea, Sierra Leone and Liberia – by November 2014.4 Around the same time, the CDC published its own projections: a whopping 1.4 million cumulative cases by late January 2015 in Sierra Leone and Liberia (with the authors citing slowing growth in Guinea as the reason for its exclusion from the model).5 Meanwhile, at HealthMap – the lab where I work – our most recent model outputs are a decent match to the WHO predictions; using calibration data from late October 2014, we expect about 16,000 cumulative cases by mid-November and 59,000 by late January (Figure 5).

Given that both the WHO and the CDC are using the same aggregate reporting data and neither assumes that transmission conditions will change, why are their results so different? Methodological approaches might account for some of the discrepancy, but the most significant difference is that the CDC considers under-reporting of Ebola cases while the WHO does not.

Aggregate reporting data only reflects patients who are actually hospitalized; it does not account for the cases that never make it
into an Ebola ward, which may happen for any of the following reasons:

1. **Fear and mistrust.** Often after witnessing loved ones going into hospitals alive and coming out in body bags, the sick do not necessarily believe that their likelihood of surviving will increase by seeking care.

2. **Time and money.** Given that treatment cannot be provided at hospitals that lack isolation facilities, seeking care can become resource-intensive for those who have to travel far to reach an Ebola-ready clinic.

3. **Death en route.** Even when the sick try to seek care, they can sometimes deteriorate quickly and die before they can make it to a treatment facility.

4. **Lack of beds.** When Ebola wards are full, the sick are turned away without diagnosis.

The CDC incorporates an “under-reporting factor” (UF) of 2.5 in its model, suggesting that for every case that is reported, 1.5 are not counted. Upon closer inspection, however, it becomes apparent that the approach taken to calculate the value of this constant multiplier is subjective. The method looks something like this:

\[
UF = \frac{B(t)_{\text{Expected}}}{B(t)_{\text{Observed}}}
\]

where \( B(t)_{\text{Expected}} \) is the number of beds that the CDC model predicted would be in use at time \( t \) and \( B(t)_{\text{Observed}} \) is the number of beds that were actually in use at time \( t \) according to ‘expert opinion estimates’. This guess is as good as any, but scientifically speaking, it is only that – a guess (and not a particularly objective one, at that).

**A balancing act**

These differing projections complicate how we handle planning for the outbreak in the months ahead, giving way to two major considerations. Firstly, how do we properly allocate resources to ensure that our hospitals can treat those who do seek care? Secondly, how do we allocate resources to prevent community-level transmission among those who choose not to seek care – that is, the cases that will go unreported for the first two reasons cited earlier?

Given that resources are finite, an inherent trade-off exists in this situation. Resources spent on clinic-centric preparedness cannot also be spent on community-centric preparedness. So do we prepare our hospitals for the number of cases we expect to be reported, or for all cases – reported and unreported alike? The latter option would lessen the likelihood of turning patients away who are actively seeking treatment, but it would also be more cost-intensive. The former, meanwhile, would provide greater resource availability for community preparedness – but at the cost of not being able to treat every patient who shows up at an Ebola clinic.

To optimize this trade-off, we need an improved probabilistic understanding of the reasons why under-reporting occurs in the first place. As mentioned before, there appear to be four primary reasons why cases go unreported – being turned away at the doors of an Ebola clinic is only one of them. Even if we assume that the CDC’s macro-scale assessment of under-reporting is correct, we need to understand the probability distribution associated with each of the four reasons cited earlier.
do not know what percentage of cases go unreported because there are not enough beds for them, as opposed to one of the three other listed reasons.

If we prepare for the future with the assumption that all unreported cases would have sought treatment had it been available in the past, we will inevitably misallocate resources. We may end up over-funding clinical preparedness in lieu of safe household care-giving campaigns, despite the fact that the latter might very well be more effective at stemming community-level transmission.

With this in mind, we should assign conditional probabilities to each of the four possible reasons for under-reporting to help us plan better for both reported and unreported cases alike (as outlined in the box on page 12). To do so correctly would require either a survey-based study of currently affected communities or a simulation-based study using information garnered from past outbreaks.

Unfortunately, both are difficult to implement and pose their own set of unique limitations. Classical behavioural surveys, for example, rarely generate a large enough sample to draw population-level conclusions, while simulations rarely serve as a good approximation of reality because the past probabilities used to set parameters do not necessarily apply to the present. Nevertheless, either approach would be better than what we currently have at our disposal.

Refining our probabilistic understanding of under-reporting is crucial to curbing the growth of this outbreak in a way that is both time- and resource-efficient. Proper allocation of money and human resources – in clinical settings and for community interventions alike – is of the essence. For our projection models to responsibly inform planning and preparedness, we need to predict not only how many cases will be reported as the outbreak progresses, but also how many will not and for what reasons. Our ability to stop this epidemic may very well depend on it.

References


Maimuna (Maia) Majumder is an engineering systems PhD student at Massachusetts Institute of Technology and computational epidemiology research fellow at HealthMap. Before coming to MIT, she earned a Bachelor of Science in engineering science and a Master of Public Health in epidemiology and biostatistics at Tufts University. Her research interests involve probabilistic modelling, Bayesian statistics, and systems epidemiology in the context of emerging infectious diseases

Ebola — the key facts

• Ebola virus disease (EVD), formerly known as Ebola haemorrhagic fever, is a severe, often fatal illness in humans.
• The virus is transmitted to people from wild animals and spreads in the human population through human-to-human transmission.
• The average EVD case fatality rate is around 50%. Case fatality rates have varied from 25% to 90% in past outbreaks.
• The first EVD outbreaks occurred in remote villages in Central Africa, near tropical rainforests, but the most recent outbreak in West Africa has involved major urban as well as rural areas.
• Community engagement is key to successfully controlling outbreaks. Good outbreak control relies on applying a package of interventions, namely case management, surveillance and contact tracing, a good laboratory service, safe burials and social mobilisation.
• Early supportive care with rehydration, symptomatic treatment improves survival. There is as yet no licensed treatment proven to neutralise the virus but a range of blood, immunological and drug therapies are under development.
• There are currently no licensed Ebola vaccines but two potential candidates are undergoing evaluation.

Source: WHO Fact Sheet no. 103 (bit.ly/1x0Wup2)