

Hi, I'm Senjuti Saha from the Child Health Research Foundation in Dhaka, Bangladesh. I wanted to talk a little bit about the ethics of data sharing in the next couple of minutes. With a lot of sequencing capacity across the world, with an increasing amount of data that is being generated, from many countries from high income countries to low income countries, it is becoming extremely important that we think carefully about the ethics of data sharing and we come up with certain guidelines of how the data should be shared and how the data should be used by the wider scientific community.

So I really think that there are two kind of scenarios that call for different data sharing policies or effects. One is a scenario of war time. For example, during a pandemic, and the other scenario is during peacetime when there is no outbreak ongoing, but there are certain endemic diseases that we are dealing with.

During wartime, during a pandemic like COVID or let's say Ebola, it's extremely important that data is shed as quickly as possible. Speed is the most important thing. So let's say, I know I am in Bangladesh and my group and I figure out that there is an outbreak ongoing and we sequenced by the pathogen behind the outbreak. I think it's a must that we shared the data with the world immediately so that the wider scientific community can use that data and as a team, we can figure out how to resolve the war, how to design new drugs, how to design new vaccines, so speed is of essence.

At the same time, the scientific community should also utilise that data to resolve the war not necessarily for personal benefits. I think it's unethical that during war time armchair epidemiologists or bioinformaticians sitting in a high income country, be it US, UK use data that are generated by other scientists and publish that in prestigious journals without including the primary data generators or understanding the context properly. So there were responsibilities, both on the scientists, the groups who are generating the data, and at the same time, the wider scientific community across the world who are using the data that is being shared as quickly as possible.

Now, a lot of these things really change when we are at peace time when we are just dealing with endemics. We have a little more time in our hand to generate the data, analyse the data, to figure out how to design epidemiological studies, take all the metrics into consideration, resolve the issues we have, write a good manuscript and publish that manuscript analysis and at the same time share all the data that were used to let's say write that paper. When scientists are using that data or after the scientists have published a paper, release all the data, I think it's totally fine for the wider scientific community to be able to use the data.

To be able to use the data for meta analysis and publish new papers. However, even in that scenario, at some times, I think it's important that the data is put into context. That primary data generators are brought into the conversation and is ensured that the data is being analysed within the proper context. I also think there are certain roles and responsibilities of the donors who fund the study, the journals who publish the study. They must ensure that, not only papers are published, but the data that were used for analysis are also all released and shared openly.

Now, with increasing conversations of data sharing, especially during peacetime, a lot of our agencies and a lot of donors are coming up with this very stringent policies for data sharing, giving scientists certain strict deadlines by when data must be shared. But I think it's very important to respect and acknowledge that every country, every individual, every group are going to have different practises of data sharing, and oftentimes, those practises are based on certain reasons. And reasons could include access to compute resources, language barrier, access to expertise of analysing the data, et cetera.

So let me give you an example over here too. In our group in Bangladesh, we currently have access to much less compute resources than, say, my colleagues in California. Something that takes us two weeks to do can sometimes take just a few hours for my colleagues to do in California, so it's really important that it's taken into consideration that not all groups can work at the same speed, so policies should be different for different groups. There is no way one policy can fit all.

So I think it's important that, when certain policies are made, local communities are engaged. It is important that scientists of both the global North and global South are included in these conversations so that we can come up with certain rules, regulations, or guidelines or recommendations that can actually be used to help all the scientists across the world. I think we are still at infancy in understanding on how data should be actually shared that is ethical and equitable and really ensures that everybody is equitably able to practise science and that science is beneficial for all across the world. Thank you.