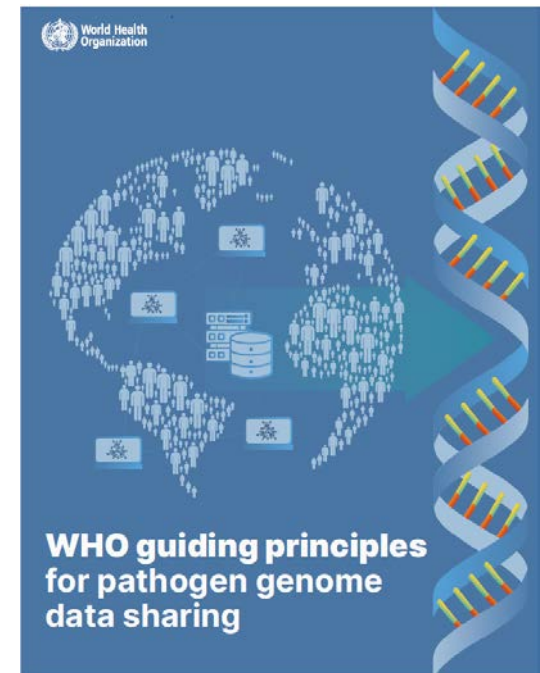


WHO Guiding Principles on Pathogen Genome Data Sharing

Vasee Moorthy MD PhD
Senior Advisor
WHO Science Division
moorthyv@who.int

[Link to WHO Guiding Principles](#)

[WHO Genomic Surveillance Strategy for Epidemic Pathogens](#)



Issues raised on data and results sharing during previous PHEICs

Different/related issues in each of these areas

Genome sequences

Epidemiological data

Natural history of disease

Clinical trials of experimental products

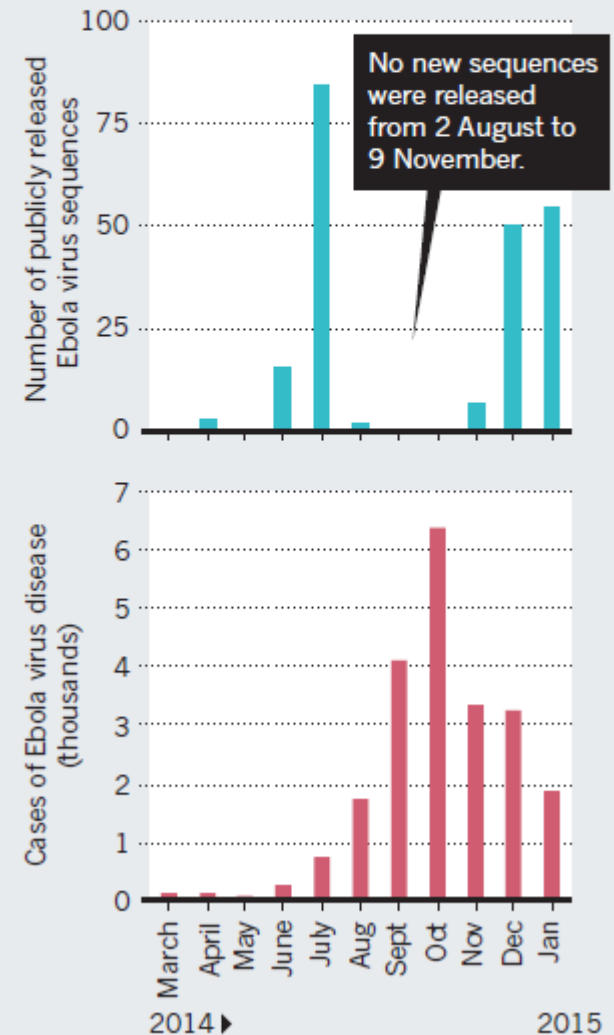
Biological material sharing



Example: genome sequences Ebola 2014-2015

GAPS IN THE DATA

Genome sequences from the West Africa outbreak of Ebola virus were first made publicly available in April 2014. Since 99 genomes were released in July, data sets have been shared sporadically, even though more are known to have been generated.



Process

- Face to face meetings to discuss principles Geneva 2017 and Bellagio 2022
- Virtual meeting 2021
- Scientists, public health agencies, medical journals, research funders, laboratory staff, analytical experts, bioethicists, hosting platforms

Common concerns raised

Will there be adverse repercussions to my sharing the data more openly?

Will I be scooped if I submit my sequence data to an international platform?

Will I be acknowledged?

Will my country have access to medical products developed based on sequences?

How can sequence sharing align with the open science movement?

What are the implications of the Nagoya protocol?

Information flow

Scientists

Governments

Funders

Timely

Representative

Quality

Focus on areas of discussion for sharing modalities

Capacity Development and Quality is an overarching theme

Ethical

As open as possible

Equity including
Access to
Interventions

Fairness eg
attribution

Transparency

Governance

Ease of Use

Data Curation

Timeliness

Not waiting for
manuscript publication

Models for sharing sequence data: WHO position as open as possible

- **No sharing**

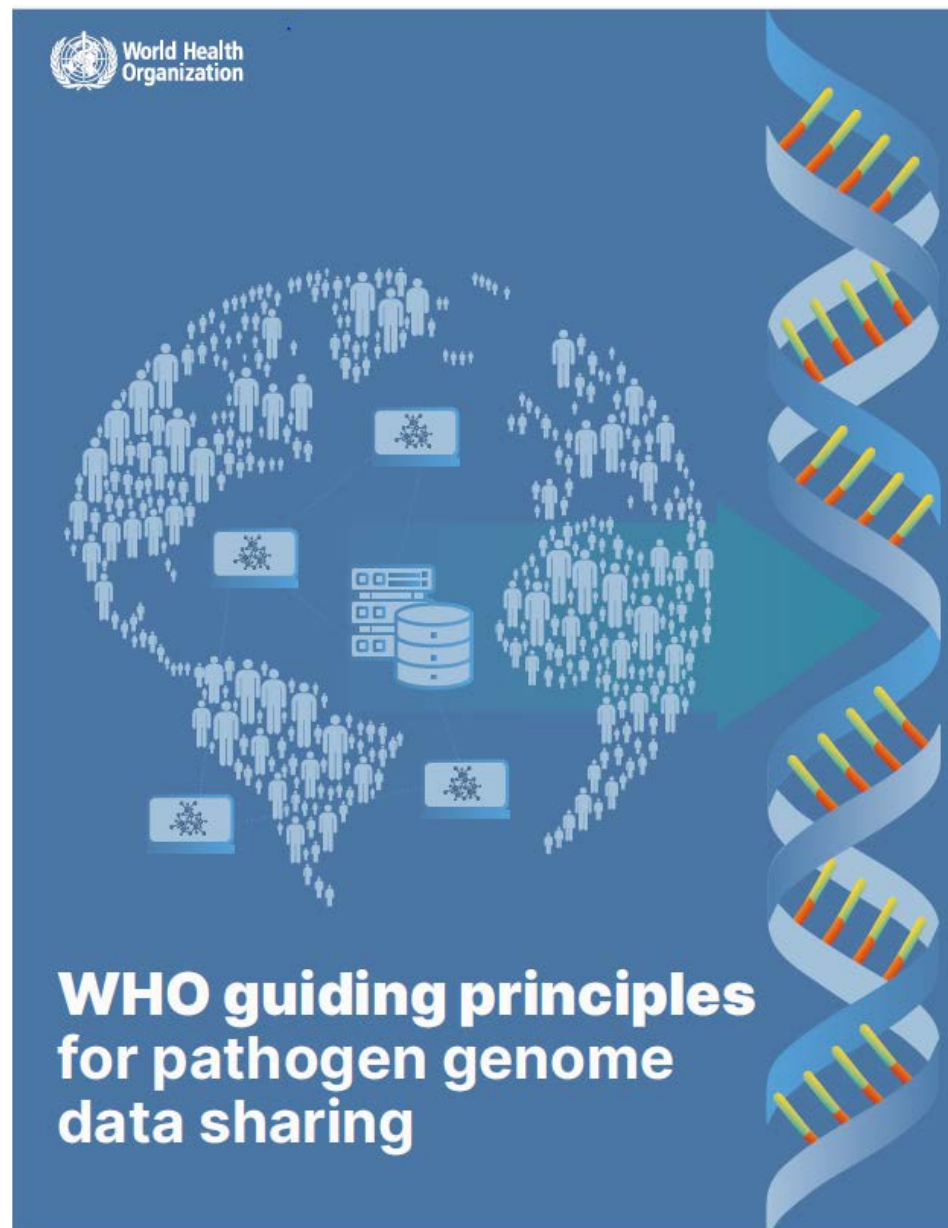
- **Closed
databases**

- **Unrestricted
Access**

- **Public access**

Audience

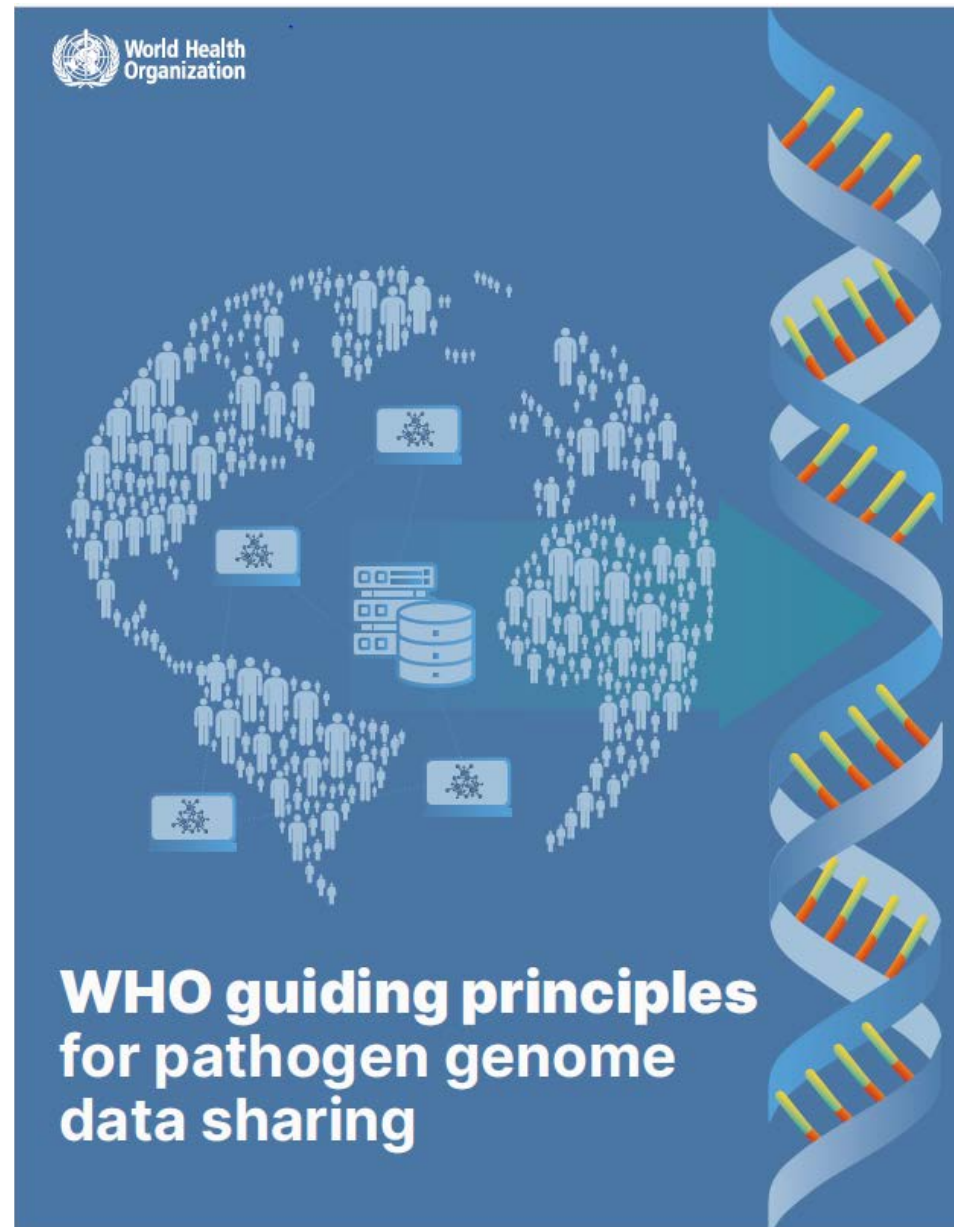
technical staff at public health agencies and researchers involved in the prevention and control of endemic, epidemic and pandemic diseases. The principles can be adopted by those who generate data, as well as those who operate or use data sharing platforms.



Scope

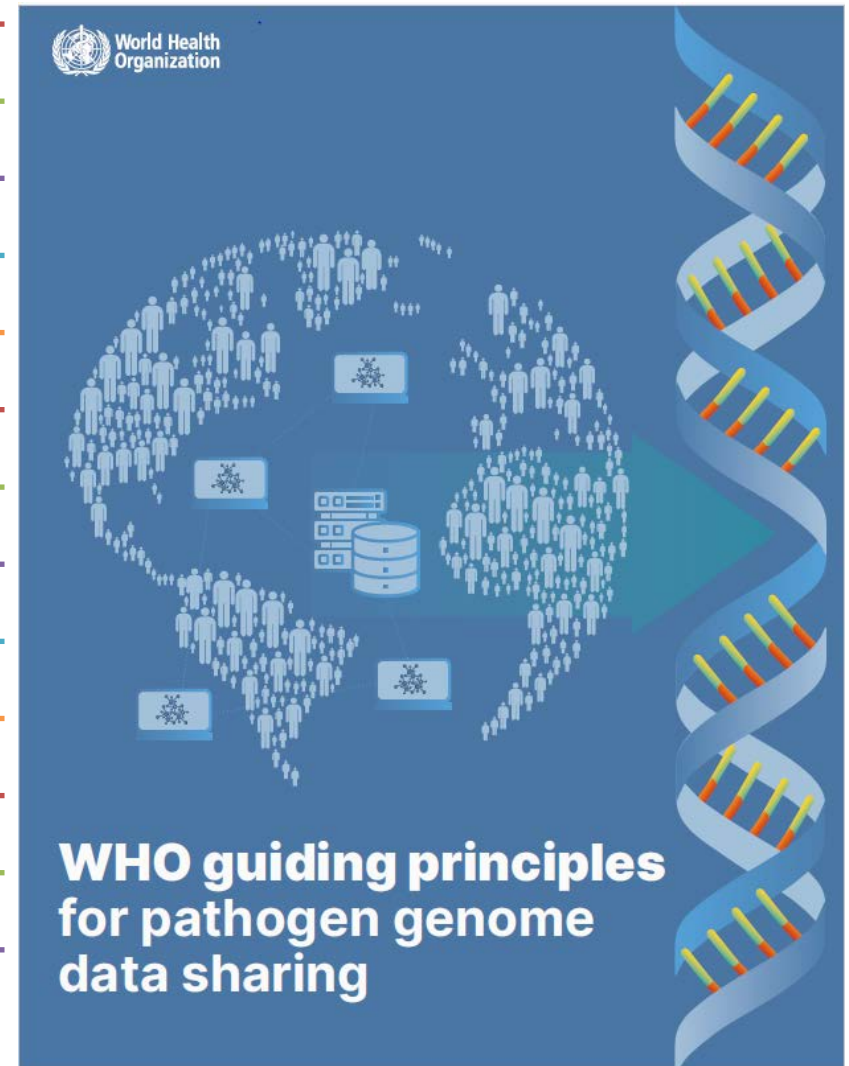
These principles apply to the rapid, public sharing of pathogen genome data, along with appropriate metadata required for interpretation (excluding any sensitive clinical data).

Genomes of organisms pathogenic to humans; also have some applicability to animal pathogens with the potential to spillover into humans.



Guiding principles

1. Capacity development
2. Collaboration and cooperation
3. High-quality, reproducible data
4. Global and regional representativeness
5. Timeliness
6. Acknowledgement and intellectual credit
7. Equitable access to health technologies as a benefit
8. As open as possible and as closed as necessary
9. Interoperability and relevance for national, regional and global decisionmakers
10. Trustworthiness and ease of use
11. Transparency
12. Consistency with applicable law and ethical regulations
13. Compliance and enforcement



Guiding principles

1. *Pathogen genomic initiatives should contribute to **capacity development*** that can establish and sustain data generation, processing, and submission, as well as reliable curation and annotation. In this manner, all countries will be able to analyze data.
2. *Promote **Collaboration and cooperation**:* between the laboratories and institutions that submit data to it and the scientists who analyze the data and report their analyses
3. ***Quality is central*** In emergencies, there is a trade-off between the time needed to attain very high-quality data and the ability to share data rapidly. In some cases, it may be appropriate to share lower quality data (clearly marked as “preliminary/not fully quality controlled” or similar), with the understanding that preliminary data need to be easily identifiable. The development, implementation, and maintenance of data standards should be prioritized for robust data sharing.
4. ***Global and regional representativeness**:* At the global level, the first few sequences submitted by countries that have not previously shared can be extremely informative. Timeliness and geographical and temporal representation relative to disease epidemiology govern the public health value of pathogen genomic data.

Guiding principles

5. ***Timeliness***: During epidemics and pandemics, ... sharing of ... data should be as timely as possible in order to make data available for analyses. The sharing of pathogen genome data should not be delayed because a data submitter wishes to prepare a scientific manuscript for publication. But neither should data submitters be obliged to waive their rights to the data they share before publication.

6. ***Acknowledgement and intellectual credit***: All contributions, including those of submitting laboratories, and where appropriate, laboratories from which clinical samples or pathogen isolates have originated, should be appropriately acknowledged in presentations and publications.

7. ***Equitable access to health technologies as a benefit***: Data-sharing policies and practices should contribute to equitable access to health technologies wherever possible. Although pathogen genome sequence sharing platforms may not be able to directly ensure access to medical interventions, such platforms should set expectations that all users of the data apply equity and fairness considerations in their use of the data for developing health technologies.

Guiding principles

8. *As open as possible and as closed as necessary:*

Unrestricted open access models for sharing should be used whenever data submitters do not wish rights to be reserved.

Where submitters opt to retain certain protections related to pathogen genome data, platforms that preserve data generators' rights should be used.

Guiding principles

9. Interoperability and relevance for decisionmakers:

The global system of pathogen genome data sharing should allow for large-scale standardized analyses that follow good practices developed in a rigorous manner by the global scientific community.

This will permit answers to critical public health questions such as epidemiological characteristics, transmission dynamics, emergence of variants, drug or vaccine resistance and the effectiveness of diagnostics.

Guiding principles

10. ***Trustworthiness and ease of use***: Establishing who controls data sharing and where data are stored is important for achieving trust.

11. ***Transparency***: Terms and conditions of databases in the global system of pathogen genome data sharing should be publicly available.

12. ***Consistency with applicable law and ethical regulations***: Pathogen genome data sharing platforms should be established and operated in a manner consistent with applicable national and international laws, regulations, rules, and standards, and ethical regulations, norms, and standards

13. ***Compliance and enforcement***: Where data sharing mechanisms permit, there may also be a need to raise awareness of instances and sanctions for breaches of the rules of data access as a means of supporting trustworthiness.



“When data related to research activities are shared ethically, equitably and efficiently, there are major gains for science and public health.”

Capacities for pathogen genomics around the world are advancing rapidly. The speed, scale and affordability of sequencing are all increasing at astonishing rates.

While these advances are welcome, inequities in access to new technologies leave blind spots in global surveillance that put us all at risk.

The commitment to equity and support for capacity development must therefore be the overriding principle of the international system for sharing of pathogen data.

