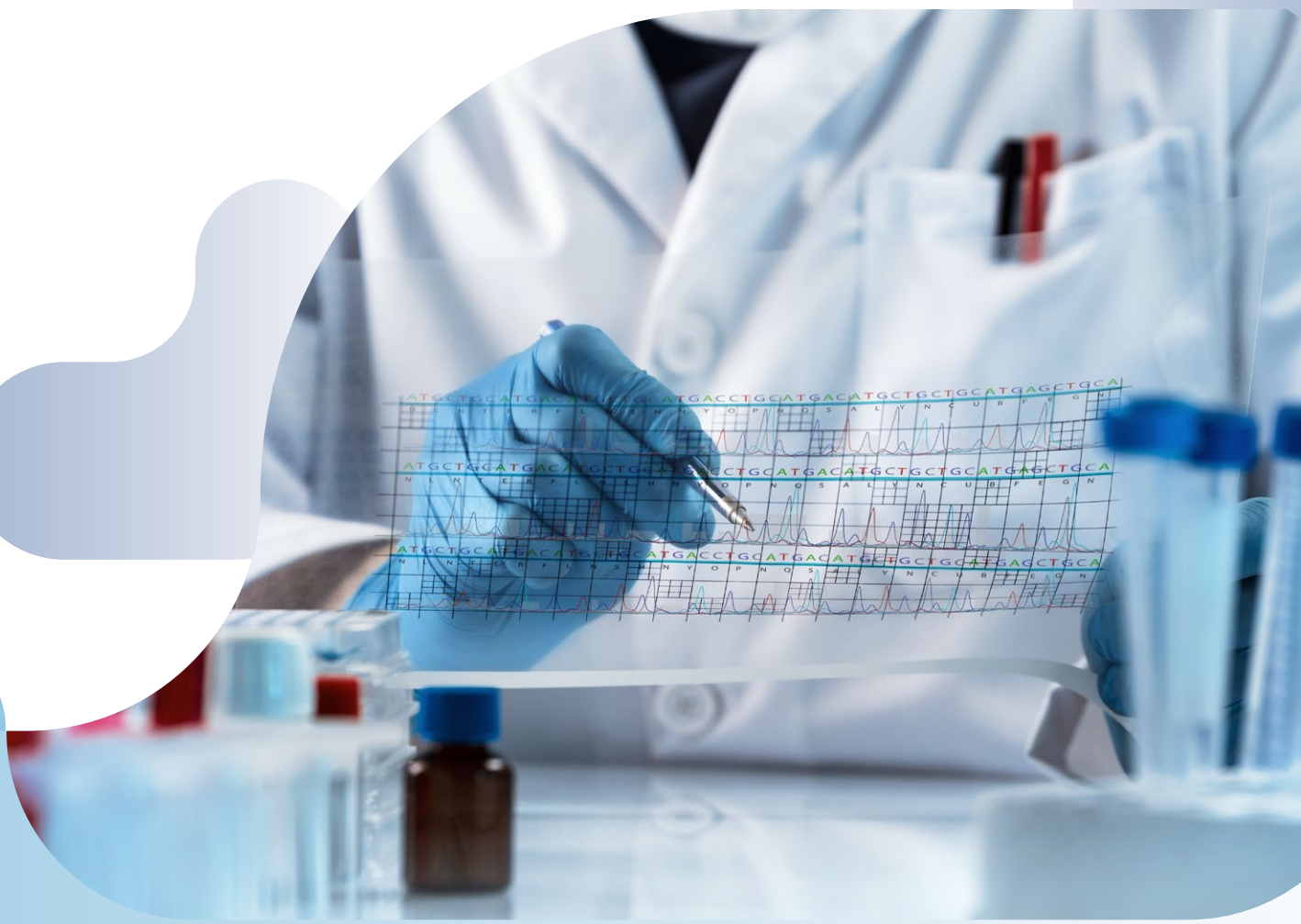
World Health Organization

WHO BioHub System technical consultation sessions

Session Two: Genetic sequence data

Background note and discussion questions



I. Introduction

In the context of the continued development of the World Health Organization (WHO) BioHub System,[1] the WHO Secretariat is conducting a series of virtual technical consultation sessions. The aim is to gather views and ideas on specific topics that will support WHO's work, to ensure that the system meets its global public health objectives.

This consultation session, scheduled for 19 November 2021, focuses on the handling of genetic sequence data (GSD), an issue of crucial importance to the WHO BioHub System's design and implementation.

WHO expects that this session will stimulate robust discussion and presentation of views and ideas on the topic, to support WHO's development of options for handling GSD in the context for the WHO BioHub System. These options will then be submitted for further consideration and validation by Member States.

The purpose of this background note is to facilitate such discussion and debate; thus, the document does not present a conclusive view or reflect the position of WHO on any specific subject or topic. Definitions for any terms used in this note but not defined can be found in the BioHub documents package (see footnote 1).

With the above in mind, this background note provides:

- a working **definition of GSD** for the purposes of the session;
- a discussion of certain **factors that may affect access to and use of GSD in the WHO BioHub System**; and
- **questions for discussion**, intended to steer the consultation towards identifying options and associated implications for handling GSD in the context of the WHO BioHub System.

II. GSD: A working definition for purposes of the WHO BioHub System

GSD results from the sequencing of the genes of biological material. Such sequencing furnishes the nucleotide base pairs encoding for the proteins of the organism, in informational form.

In relation to the WHO BioHub System, GSD relates directly to the current WHO glossary definition of biological materials with epidemic or pandemic potential[2] (BMEPP). The glossary defines BMEPP as *clinical samples, specimens, isolates, and cultures, either original or processed, of a novel pathogen with epidemic or pandemic potential.*

GSD are a necessary step for the WHO BioHub System's goal of quickly characterizing emerging pathogens that pose a risk to global health security. GSD can arise from the processing of BMEPP.

1 See https://www.who.int/initiatives/who-biohub.

2 BMEPP do not include influenza viruses or other non-BMEPP pathogens that may be contained in clinical samples or specimens shared under the WHO BioHub System.

III. Factors potentially impacting the access and use of GSD in the context of the WHO BioHub System

The ability of potential users to access the GSD depends on the database(s) used by the GSD provider. Different GSD databases have different conditions for access and use. Some databases are publicly accessible with data access conditions (e.g. GISAID[3]), some have unrestricted access (e.g. the databases in the International Nucleotide Sequence Database Collaboration [INSDC][4]) and some are commercial private databases.[5]

GSD are accessed and used by different kinds of researchers, who often use one or more databases (broad or specialized) that best facilitate their research activities. Although the use of GSD is heavily concentrated in high-income countries, researchers around the world rely on GSD to advance both curiosity-driven biomedical science and analyses for public health purposes.

At the beginning of a new event, high-quality sequences are essential to allow diagnostics, therapeutics and vaccine development to be initiated. Later in outbreaks, GSD has most value in aggregate. Here, timeliness of sharing, high quality and data that are representative of as many countries as possible are the critical aspects that determine the value of aggregate sequencing datasets in local, regional and global analyses.

GSD directly enables the development of medical countermeasures such as vaccines, diagnostics and therapeutics. Despite its significance, the use of GSD alone may not yet be sufficient to enable licensing and market authorization of products such as vaccines because, in some jurisdictions, regulatory or testing requirements mean that access to physical BMEPP is still required.

Under the proposed WHO BioHub System, BMEPP providers will share material with the WHO BioHub System, on the understanding that the use of such BMEPP will give rise to fair and equitable benefit sharing. The question of how GSD are to be handled under the WHO BioHub System is the focus of this technical consultation session.

The WHO Secretariat's proposal for discussion is that any GSD generated from processing of BMEPP are to be made publicly accessible, so that everyone committed to supporting a public health response can access this information in a timely manner. This could be done through either a public access database with data access conditions (e.g. GISAID) or through databases with unrestricted access (e.g. one of the INSDC databases). A point for discussion is that the WHO Secretariat proposes to leave this decision to the provider (i.e. the party depositing BMEPP into the system), who can choose whether one or both options will be used. We invite discussion on these options, taking into account biorisk implications and potential undermining of the guiding principles of the WHO BioHub System.

---

3 See https://www.gisaid.org/.

4 See https://www.insdc.org/. INSDC comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA) and GenBank at NCBI.

5 From a study on public and private databases for the CBD Secretariat, see https://www.cbd.int/.

IV. Questions for discussion at the consultation session

Mindful of the BioHub's guiding principles of the WHO BioHub System and its core objectives of promoting *rapid and timely sharing of BMEPP, facilitating rapid characterization of such pathogens and rapid risk assessments, facilitating the development of safe, protective countermeasures and other public health products and ensuring fair and equitable access to such products by all countries based on public health needs:*

1. What do you consider to be the most important factors and considerations relating to GSD vis-à-vis the BioHub project?

2. What are your views on the above proposed approach, by which GSD generated from processing of BMEPP should be made publicly accessible?

3. What are your views on whether the Provider of BMEPP should decide whether the access to GSD should be done through either a public access database with data access conditions (e.g. GISAID) or through databases with unrestricted access (e.g. one of the INSDC databases), or both, or through other modalities?

4. How can we address the handling and use of BMEPP GSD to promote fair and equitable sharing of public health benefits arising from its use?