BEST PROCESS TO HANDLE GENETIC SEQUENCE DATA FROM INFLUENZA VIRUSES WITH HUMAN PANDEMIC POTENTIAL (IVPP GSD) UNDER THE PIP FRAMEWORK

Options to monitor the use of genetic sequence data from influenza viruses with human pandemic potential (IVPP GSD) in end-products

This version replaces the draft of 29 September 2015 which was shared with the PIP Advisory Group in October 2015 and shared with stakeholders, for comments, from October to December 2015. This revised document takes into account comments received. To view the previous draft or the comments received, please go to http://www.who.int/influenza/pip/advisory_group/gsd/en/.

I) BACKGROUND

The PIP Framework

The PIP Framework is an international arrangement, adopted in 2011 by the 194 Member States of the World Health Organization (WHO), that seeks:

i) to improve and strengthen the sharing of influenza viruses with human pandemic potential ('IVPP') through a WHO-coordinated network of public health laboratories (known as ‘GISRS’), and;

ii) to promote the fair and equitable access, by developing countries, to the benefits arising from such sharing.

Under the PIP Framework, IVPPs are part of a broader set of materials called ‘PIP Biological Materials’ or ‘PIP BM’, which include human clinical specimens, influenza virus isolates, extracted RNA, cDNA, and influenza candidate vaccine viruses developed from IVPPs by GISRS laboratories.1 Under their Terms of Reference, GISRS laboratories must share PIP BM in a “rapid, systematic and timely manner [with] other qualified laboratories, to facilitate public health risk assessment, risk response activities and scientific research”2.

The PIP Framework recognizes “that greater transparency and access concerning influenza virus genetic sequence data is important to public health” and that “there is a movement towards the use of public-domain or public-access databases such as Genbank and GISAID respectively”3. The Framework sets out the following expectations regarding IVPP genetic sequence data (‘GSD’):

- “genetic sequence data, and analyses arising from that data, relating to H5N1 and other influenza viruses with human pandemic potential should be shared in a rapid, timely and systematic manner with the originating laboratory and among WHO GISRS laboratories” (See PIP Framework Section 5.2.1)

- WHO CC “upload available haemagglutinin, neuraminidase and other gene sequences of

---

1 See PIP Framework Section 4.1
2 See e.g. PIP Framework Annex 4, paragraph 8.
3 See PIP Framework Section 5.2.2.
Best Process to handle IVPP GSD under the PIP Framework

29 September 2015; Revised 6 May 2016

A(H5) and other influenza viruses with pandemic potential to a publicly accessible database in a timely manner but no later than three months after sequencing is completed, unless otherwise instructed by the laboratory or country providing the clinical specimens and/or viruses (Guiding Principle 9). (See PIP Framework Annex 5 Terms of Reference for WHO Collaborating Centres, Core Term of Reference B.5)

- “WHO GISRS laboratories will submit genetic sequences data to GISAID and Genbank or similar databases in a timely manner consistent with the Standard Material Transfer Agreement.” (see PIP Framework Annex 4, Guiding Principles for the development of Terms of Reference for current and potential future WHO global influenza surveillance and response system (GISRS) laboratories for H5N1 and other human pandemic influenza viruses, Principle 9).

Influenza Virus Traceability Mechanism (IVTM)

The PIP Framework requires that all transfers of PIP BM between GISRS laboratories and to entities outside GISRS be recorded in the IVTM which enables real-time tracking of the movement of PIP BM into, within and out of the WHO GISRS. The IVTM, established by WHO in 2008, allows users to trace shipments of PIP BM and to access information about the biological material shipped, including the nature of the material, its source and its origin.

The IVTM however does not provide any information on GSD generated from PIP BM or its use to develop end-products, such as scientific publications, patents and/or commercial products.

Benefit Sharing

The sharing of PIP BM gives rise to tangible and intangible benefits, which include, for example, pandemic risk assessment and pandemic influenza vaccines, both of which are essential for pandemic preparedness and response. Access to benefits is secured by WHO through 2 key mechanisms:

1) Legally binding contracts – known as ‘Standard Material Transfer Agreements 2’ or ‘SMTA2’ – concluded with all non-GISRS entities that receive PIP BM from GISRS; and

2) The Partnership Contribution, an annual payment made to WHO by influenza vaccine, diagnostic and pharmaceutical manufacturers that use the WHO GISRS.

Genetic sequence data and the PIP Framework

During PIP Framework negotiations, Member States recognized the importance of genetic sequence data for pandemic preparedness and response and requested that the Director-General seek advice from the PIP Advisory Group (PIP AG) on the “best process for further discussion and resolution of issues relating to the handling of genetic sequence data from H5N1 and other [IVPPs] as part of the Pandemic Influenza Preparedness Framework.”

---

4 See PIP Framework Section 5.3. To access the IVTM see: https://extranet.who.int/ivtm/Default.aspx
5 See Annex 2 of the PIP Framework.
6 See PIP Framework Section 6.14.3.
7 Under the PIP Framework, the Advisory Group is a group of 18 international experts that provides “evidence-based reporting, assessment and recommendations regarding the functioning of Framework” to the Director-General. (See PIP Framework Section 7.1.2 (iii)).
Best Process to handle IVPP GSD under the PIP Framework

29 September 2015; Revised 6 May 2016

The matter has gained importance given the recent development of synthetic biology technologies which allow the production of influenza virus proteins, antibodies and influenza candidate vaccine viruses using only genetic sequence data. These developments raise questions about the broader implications of sharing and using IVPP GSD, notably with respect to benefit sharing under the PIP Framework.

Technical Expert Working Group (TEWG) on genetic sequence data

In light of the foregoing, in mid-2013, the PIP AG decided to begin its examination of the issues relating to the handling of GSD under the PIP Framework. Given the Advisory Group’s limited expertise in the subject-matter, it established a Technical Expert Working Group (‘TEWG’) to provide it with background and technical information.

The TEWG presented its final report to the PIP AG in October 2014. On the question of monitoring and tracing the sharing of GSD, the TEWG wrote:

“The objective of benefit-sharing may be met by monitoring use of GSD and/or tracing GSD or by other mechanisms related to influenza-related products. While monitoring and tracing the use of GSD is limited by the medium used to share it, technical mechanisms to trace or monitor downloading of GSD from databases may be implemented. GSD of PIP biological material can also be generated by non-GISRS laboratories. In that case, WHO will likely not know of this, and the sharing of such will be more difficult to monitor. **Notwithstanding, there are other potential mechanisms that could be developed to monitor the use of GSD, such as processes related to influenza-related products (e.g. regulatory approval files and patent applications).**”

Advisory Group guidance on the best process for further discussion and resolution of the issues relating to the handling of GSD

During its October 2014 meeting, the PIP AG held a technical consultation with six database representatives to gather information on electronic databases that house IVPP GSD. In its meeting report to the Director-General, the PIP AG observed that:

“a. Laboratories should continue to share [IVPP GSD] as soon as it becomes available because it is necessary for timely and comprehensive pandemic risk assessment and response.

b. In accordance with Section 6.3.2, laboratories using GSD will meet appropriate biosafety guidelines (WHO Laboratory Biosafety Manual, 3rd edition) and employ laboratory protection best practices.

c. The objective of benefit-sharing may be met by mechanisms related to monitoring products generated using influenza GSD, rather than by monitoring use of GSD and/or tracing GSD, noting that source identification is critical.

---

8 See PIP Framework Section 5.2.4.
10 Throughout this document, the term “database” refers to any institute, collaboration, initiative, organization or other entity that houses genetic sequence data.
Best Process to handle IVPP GSD under the PIP Framework

29 September 2015; Revised 6 May 2016

d. Closer collaboration regarding open sharing of influenza GSD among the many different databases is desirable.” [emphasis added]

Thus, in its guidance to the Director-General, the PIP Advisory Group recommended a process to identify “the optimal characteristics of a system for the handling of IVPP GSD under the Framework, including consideration of [...] systems to monitor use of IVPP GSD in end-products”.

This paper presents possible options for monitoring the use of IVPP GSD to develop end-products, such as vaccines, antivirals and diagnostics, in relation to the PIP Framework. It does not however address how to operationalize these options or which entity would be responsible for implementing them.

II) DISCUSSION

1. Promoting the sharing of benefits generated using IVPP GSD

Because no mechanism is fool-proof, promoting the sharing of benefits generated using IVPP GSD will likely require a combination of both upstream and downstream options. Upstream options should focus on informing entities and individuals accessing GSD of potential obligations and/or expectations under the PIP Framework. This would give more legal certainty to users of the data and could facilitate identification of users of IVPP GSD for benefit-sharing purposes.

In contrast, downstream options would involve implementing mechanisms to monitor use of IVPP GSD to develop end-products, such as vaccines, antivirals and diagnostics. This would allow WHO to identify entities that have used IVPP GSD to generate benefits and result in more transparency for providers of the data.

1.1. Upstream Options

Upstream options are options which are implemented at the point at which IVPP GSD is distributed and accessed. In accordance with the Framework, WHO GISRS laboratories are expected to share their data in publicly-accessible databases (“GISAID and GenBank or similar databases”). “Publicly accessible” is generally understood to mean accessible to the public and not limited to a certain category of users. There are two general categories of publicly-accessible databases: publicly-accessible database without registered user access and publicly-accessible database with registered user access. Upstream options should therefore focus on such databases.

1.1.1. Publicly-accessible databases without registered user access

The term “publicly-accessible database without registered user access” refers to “a database in which

---

11 Upstream refers to the transfer and receipt of biological materials or GSD and downstream refers to end-products produced using GSD such as vaccines, antivirals and diagnostics.
12 In connection with this, a technical working group tasked with identifying the optimal characteristics of an IVPP GSD sharing system was established in April 2015 by the PIP AG. As part of its work, the TWG examined different upstream options that would promote the objectives of the Framework, notably how the data sharing system, including databases, can play a role in facilitating benefit sharing under the Framework.
access to the data is provided to the database user without having to explicitly accept a formal data access and use agreement, and without registration or log-in required”. The use of a publicly-accessible database without registered user access for the sharing of IVPP GSD is consistent with the objective of the Framework to improve virus sharing because it allows easy public access to the data. For benefit sharing, the main disadvantage of these databases is inherent in the fact that they do not identify data users. This means that there is no direct tracking of the downloading, distribution, or use of data and, thus, it would be difficult to regulate access to data available in such databases in order to implement benefit-sharing under the Framework.

A solution would be to request that publicly-accessible databases without registered user access notify their users about PIP Framework obligations and/or expectations. This would be consistent with the process that is in place for the sharing of PIP biological materials (“PIP BM”). Indeed, when a PIP BM is shipped to a laboratory outside of GISRS, a notice is included which informs the recipient that “acceptance of the biological materials in this shipment will entail certain obligations under the Pandemic Influenza Preparedness Framework”.

a) Notifying Users of PIP Framework Obligations and/or Expectations

Publicly-accessible databases without registered user access may notify users of potential third-party claims over the data. For example, a database may include a general statement that the data may be subject to third-party intellectual property rights, “biodiversity-related access and benefit-sharing rights” and/or data access agreements. Such notification statements are generally easily accessible, viewable to all users, and not tied to a particular data entry.

Existing notification statements in publicly-accessible databases without registered user access could therefore be expanded – or new statements could be added – to inform IVPP GSD users of PIP Framework obligations and/or expectations and facilitate their identification. An example of such text is provided:

Notifying Users of PIP Framework Obligations and/or Expectations

Publicly-accessible databases without registered user access may notify users of potential third-party claims over the data. For example, a database may include a general statement that the data may be subject to third-party intellectual property rights, “biodiversity-related access and benefit-sharing rights” and/or data access agreements. Such notification statements are generally easily accessible, viewable to all users, and not tied to a particular data entry.

Existing notification statements in publicly-accessible databases without registered user access could therefore be expanded – or new statements could be added – to inform IVPP GSD users of PIP Framework obligations and/or expectations and facilitate their identification. An example of such text is provided:

17 For example, see National Center for Biotechnology Information, GenBank Data Usage, available at http://www.ncbi.nlm.nih.gov/genbank/
18 For another example, see EMBL-EBI Terms of Use:

Data Services
1. The online data services and databases of EMBL-EBI are generated in part from data contributed by the community who remain the data owners.
2. EMBL-EBI itself places no additional restrictions on the use or redistribution of the data available via its online services other than those provided by the original data owners.
3. EMBL-EBI does not guarantee the accuracy of any provided data, generated database, software or online service nor the suitability of databases, software and online services for any purpose.
4. The original data may be subject to rights claimed by third parties, including but not limited to, patent, copyright, other intellectual property rights, biodiversity-related access and benefit-sharing rights. For the specific case of the EGA database and human data consented for biomedical research, these rights may be formalised in Data Access Agreements. It is the responsibility of users of EMBL-EBI services to ensure that their exploitation of the data does not infringe any of the rights of such third parties.” (EMBL-EBI, “Terms of Use for EMBL-EBI Services”, 21 October 2015, available at http://www.ebi.ac.uk/about/terms-of-use.)
19 Ibid. at par. 4
The use of genetic sequence data of influenza viruses with pandemic potential may give rise to obligations and/or expectations under the “Pandemic Influenza Preparedness Framework for the sharing of influenza viruses and access to vaccines and other benefits” (the PIP Framework) adopted by the World Health Assembly in May 2011. It is the responsibility of users of [name of database] to ensure that their use of the data is consistent with PIP Framework obligations and/or expectations. For further information on the PIP Framework, visit the PIP Framework webpage at http://www.who.int/influenza/pip/en/. For questions about the PIP Framework, contact pipframework@who.int.

Adding a statement on the PIP Framework to a database would have the benefit of being easy and inexpensive to implement. In addition, the fact that notification statements are already in use in many databases provides a useful precedent for implementing this solution. Finally, the inclusion of a PIP Framework statement would make the use of publicly-accessible databases without registered user access more consistent with the objectives of the Framework. Alerting users to possible PIP Framework obligations and/or expectations would encourage participation in the benefit-sharing objective.

b) Identifying IVPP GSD

In addition to the inclusion of a statement, laboratories that upload data in publicly-accessible databases without registered user access could identify each IVPP sequence as PIP Framework IVPP GSD at the time of upload. This could be done through the inclusion of a tag linked to a particular data entry or a specific metadata field. This would greatly facilitate awareness of the PIP Framework by data users.

1.1.2. Publicly-accessible databases with registered user access

The term “publicly-accessible database with registered user access” refers to “a database in which access to the data is granted after the database user registers and explicitly accepts a formal data access and use agreement. After this, access requires using a log-in procedure.” Terms of use may include a requirement that users acknowledge data contributors, including the laboratory where the original virus sample was obtained as well as the laboratory that generated and submitted the sequence data to the database. Such databases also generally limit further distribution of GSD only to other users that have also agreed to the terms of access and use. For example, the GISAID EpiFlu™ database requires that users positively identify themselves during registration and accept the terms of the GISAID EpiFlu™ Database Access Agreement. This agreement contains terms regarding, inter alia, acknowledgement/co-authorship, collaboration with originating laboratories, distribution of the data to third-parties, intellectual property, and suspension of access.

Similarly to the solution proposed above for publicly-accessible database without registered user access, publicly-accessible database with registered user access could be asked to include a statement on the PIP Framework on their website or include PIP Framework obligations and/or expectations as part of their terms of use.

20TWG, supra note 14.
a) Including PIP Framework Obligations and/or Expectations in Terms of Use

The terms of use for publicly-accessible database with registered user access could also be modified to include a mention of PIP Framework obligations and/or expectations. In particular, the terms of use could be expanded to include the following:

(§) PIP Framework. Use of genetic sequence data of influenza viruses with human pandemic potential, may be subject to obligations and/or expectations under the “Pandemic Influenza Preparedness Framework for the sharing of influenza viruses and access to vaccines and other benefits” (the PIP Framework) adopted by the World Health Assembly in May 2011. By using these data, you recognize that you will be contacted by WHO regarding potential benefit sharing under the PIP Framework. For further information on the PIP Framework, visit the PIP Framework webpage at http://www.who.int/influenza/pip/en/. For questions about the PIP Framework, contact pipframework@who.int.

b) Source identification

In addition, where appropriate, the terms of use could be modified, for instance, to include a requirement for source identification using accession numbers. As discussed in greater detail below in section 1.2.1, identifying the IVPP GSD with accession numbers would greatly facilitate monitoring its use. For example, the terms of use could include the following requirement:

(§) Origin of Data. Published results should provide the following information: the laboratory where the clinical specimen(s) and/or virus isolate(s) were first obtained, the laboratory where Data have been generated from the virus isolate(s) and/or the original specimen(s) received and submitted to the Database if applicable.

(§) Use of Data. Use of Data from influenza viruses with human pandemic potential should be identified using the accession number for the Data in scientific publications, intellectual property applications, clinical trial registration information, product inserts and regulatory filings.

c) Information about data access

Because publicly-accessible database with registered user access require registration and log in, they could provide information to WHO about institutions and companies that access IVPP GSD. In combination with downstream monitoring options, this would facilitate identification of end-products.

1.2. Downstream Options

Downstream options include methods for monitoring the use of IVPP GSD after it has been shared and used to research and develop end-products, such as diagnostics, antivirals and vaccines. Downstream options can monitor the development of a product at different stages. For example, patent applications and research trials occur before regulatory applications are filed or end-products are brought to market. As a result, downstream options include not only the monitoring of end-products themselves, but also steps that occur earlier in the vaccine, diagnostic, or treatment research and development process.

1.2.1. Preliminary remarks
a) Disclosure

Monitoring the use of IVPP GSD in end-products relies on users clearly and consistently identifying the data in publications, patent applications, or regulatory approval files, with a unique identifier, such as the name of the virus or a database accession number. Therefore, the strength of any downstream monitoring system will depend on whether users of IVPP GSD have clearly and consistently disclosed their use of the data.

The disclosure of data is a common requirement for scientific publications, patent applications, clinical trial files and regulatory approval files. For example, an overwhelming majority of peer-reviewed scientific journals require that “all data necessary to understand, assess, and extend the conclusions of the manuscript […] be available to any reader”, including GSD. To facilitate this disclosure, many scientific journals require that GSD be deposited in a publicly-accessible database and that published papers include accession numbers. Similarly, most jurisdictions require sufficient disclosure of data in a patent application so as to enable a person skilled in the art to reproduce the invention. In addition, regulatory agencies will typically require product data, manufacturing information and details of both pre-clinical and clinical trials in order to assess the efficacy and safety of end-products before granting marketing authorisation. These different sources of information could potentially be used to monitor the use of IVPP GSD in end-products. However, there is currently no consistent practice of identifying GSD using a database accession number in patents or regulatory files.

In each of these potential sources, the disclosure requirement relates to information that is central to the research claims or necessary to carry out the invention. For example, information on the GSD used to test a product, but that are not essential to reproduce the invention, will not necessarily be included in a patent application. Thus, indirect uses of the data will not always be captured. As a result, compared to options to monitor access to IVPP GSD, monitoring the use of GSD in end-products will capture a more limited number of entities.

Potential solutions could include broadening disclosure requirements to all GSD used in the course of developing an invention or conducting a specific research project. Alternatively, benefit sharing under the PIP Framework could be linked solely to uses of GSD that are central to the research and development of an end-product.

For the time being, a practical – albeit non-comprehensive – solution involves retrieving information on the use of GSD from various existing downstream sources, thus allowing for downstream monitoring of IVPP GSD at the various stages of the end-product research and development process.

### Accession Numbers

As described above, monitoring the use of IVPP GSD in end-products relies on users clearly and consistently identifying the data in publications, patent applications, or regulatory approval files, with a unique identifier, such as the name of the virus or an accession number.

---

23 Further discussion on the potential use of common accession numbers can be found below.

24 Science, General Information for Authors, accessed on available at http://www.sciencemag.org/site/feature/contribinfo/prep/gen_info.xhtml
All databases assign numbers to genetic data sequences typically following an established format; these are called “accession numbers”.

Accession numbers are assigned upon submission of a sequence to a database. For example, accession numbers assigned to nucleotide sequences submitted directly to GenBank will typically take the form of a single letter (e.g. J, K, L, M) + 5 numerals, or 2 letters (e.g. FJ) + 6 numerals. For example, for influenza A/California/04/2009 (H1N1):

- PB2 gene accession number is: FJ966079
- PB1 gene accession number is: FJ966080
- PA gene accession number is: FJ966081
- HA gene accession number is: FJ966082
- NP gene accession number is: FJ966083
- NA gene accession number is: FJ966084
- M gene accession number is: FJ966085
- NS gene accession number is: FJ966086

Sequences can have more than one accession number. This is the case, for instance, when a database extracts sequences from another database to include it in its own database. This is often referred to as “data mining”. When this happens, the database that has extracted and taken a sequence from another database will generally include two accession numbers for a sequence: the original accession number, and its own accession number in its own format. As a result, a single gene sequence can have several different accession numbers. Thus, for example:

- In GISAID’s EpiFlu™ database, the hemagglutinin gene for A/Shanghai/MH01/2013 (H7N9) has the following accession numbers associated: the EpiFlu™ number EPI542311 and the INSDC number KF609503.
- In the OpenFlu database, the listing for this same sequence includes the following accession numbers: the OpenFlu database number OFL371034 and the INSDC number KF609503.

Considering that an influenza virus has 8 RNA segments, one virus could potentially have more than 24 accession numbers.

b) Defining the scope of the monitoring system

In setting up a downstream monitoring system, it is essential to first consider which uses of IVPP GSD will be monitored. For instance, under such system, “use” could refer to any uses of IVPP GSD; alternatively, it could be limited to uses to manufacture an influenza end-product, such as a vaccine or a diagnostic. This, however, would limit the number of entities captured by excluding those that conduct research or develop technologies or non-influenza products using the data. A potential

---

26 The International Nucleotide Sequence Database Collaboration, which includes GenBank, the European Nucleotide Archive and the DNA Data Bank of Japan.
27 On the one hand, having a broader definition of “use” would allow identifying more users for benefit-sharing. On the other hand however, this might require negotiating and entering into agreements with a significant number of entities, which could be very time-consuming and resource intensive.
solution would be to consider technologies to be “end-products”, allowing WHO to access potentially important benefits.

1.2.2. Downstream centralized search engine
In order to develop a comprehensive monitoring system, a search engine could be established to retrieve information on the use of IVPP GSD from publicly available downstream documents, such as scientific publications, patents, research trial documentation, and regulatory files. This information could be compiled into a central and accessible database for monitoring the use of IVPP GSD. Such a tool would enable early awareness of potential end-products giving time to put in place appropriate benefit sharing agreements. In addition, this would allow monitoring the use of IVPP GSD where a product is developed but full commercialization does not occur.

There are a number of sources or mechanisms that can be used to monitor use of IVPP GSD. These include:

- Patents
- Research Trial Documentation
- Regulatory Files
- Published Results
- Voluntary disclosure

1.2.3. Patents
The specifics of patent laws vary greatly between domestic jurisdictions. In an attempt to establish global minimum standards for the protection of intellectual property, Members of the World Trade Organization (WTO) negotiated the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) in 1994. Any new member (with the exception of least developed countries that benefit from a transition period until 2021) must incorporate TRIPS into domestic law as part of the overall obligations as a WTO member.

TRIPS specifies minimum standards for patent protection, including for patentable subject matter, minimum patent terms, and conditions for patent applications. WTO Members must require that patent applications disclose inventions in “a manner sufficiently clear and complete for the invention to be carried out by a person skilled in the art”. This is the only stipulation under TRIPS as to what information must be included in a patent application. Members can stipulate further details in their national laws.

The World Intellectual Property Organization also administers several treaties that contain provisions relevant to disclosure requirements, most importantly the Patent Cooperation Treaty (“PCT”). The disclosure requirement under article 5 of the PCT mirrors the TRIPS requirement, but the Regulations under the PCT contain additional specifications, notably in regards to nucleotide and/or

---

28 Art 27, TRIPS
29 Art 29, TRIPS
30 Ibid.
amino acid sequence disclosure. Thus, Standard for the Presentation of Nucleotide and Amino Acid Sequence Listings in International Patent Applications Under the Patent Cooperation Treaty (PCT) (“Standard ST. 25”) provides basic requirements for the inclusion of genetic sequences in international patent applications. These requirements include symbols and formats for enumeration, number of bases per line, and where in the patent sequence data is to be included. Accession numbers are also listed under ST.25, but only as optional identifiers. Almost all patent offices use ST.25, ensuring consistency in how genetic sequence data is presented in patents.

For inventions reliant on microorganisms and biological materials, written descriptions are often insufficient, and access to the biological material itself is required. Under the Budapest Treaty, disclosure is done by depositing the microorganisms and biological materials in an International Depositary Authority. Rule 13bis.3 of Regulations under the PCT specify that reference to deposited biological material in patent applications shall indicate the name and address of the depositary institution with which the deposit was made and the accession number given to the deposit by that institution.

As a result, patent applications for end-products derived from biological or genetic materials in WTO Members and PCT Contracting parties will typically include the names of viruses or the nucleotide sequences used to develop the end-product, but not necessarily the source of the materials or GSD. For example, in patent application WO2012072788A1 “Vaccine against influenza h5n1 viruses, medicament and treatment of h5n1 viral infections”, the following information is included:

**MATERIALS AND METHODS**

Sequence analysis.

[037] The amino acid sequences of the hemagglutinin (HA) protein of H5N1, H1N1, H2 and H6 viruses were retrieved from Influenza Virus Resource, a public database that gather influenza genome sequences from the National Institute of Allergy and Infectious diseases (NIAID) and the Genbank of the National Center for Biotechnology Information (NCBI). Sequences were aligned using MAFFT (Katoh et al., 2005) then viewed and edited using BioEdit version 7.0.5.3 (Hall, 1999).

This invention relates to an antigenic determinant of hemagglutinin, which is conserved in H5N1 lineages. The claims in the patent include antibodies for passive immunization, an H5N1 specific vaccine and a method of diagnosis. The invention utilizes both biological materials and sequence information as its direct source of viruses. Although sequence information is included, the inventors

---


34 Ibid, at paragraph 31.

Best Process to handle IVPP GSD under the PIP Framework

29 September 2015; Revised 6 May 2016

do not include source identification using accession numbers or disclosure of origin\textsuperscript{36} and thus it is very difficult to know the origin of the sequences used to generate the invention.

Ideally, in order to achieve optimal benefit sharing under the PIP Framework, patent applications should include source identification using accession numbers. While it may be difficult to legally mandate this, the practice could be included as part of a best practice guidance for meeting reproducibility requirements.

For instance, the statement on a patent application could be as follows:

Nucleotide Sequences.

[xx] The nucleotide sequence utilized in this invention was retrieved from [name of database] and is as follows: A/Anhui/1/2013 (H7N9) HA segment ([database] accession number: XXX123456). This sequence was submitted by [name of laboratory].

1.2.4. Research Trial Documentation

Research trials may include both pre-clinical animal model trials, as well as clinical trials in any phase. Typically, public access to research trial information does not occur until – and if – the data is published. There is therefore limited to no public access to pre-clinical research trial data.

In contrast, clinical trial data is generally publicly available, as clinical trials must be registered within the relevant domestic jurisdiction for quality and safety purposes. The scope of information available from registered clinical trials may vary. Generally, publicly-available clinical trial documentation does not contain information about manufacturing processes and methods and, therefore, may currently only provide incomplete information about the use of IVPP GSD.

However, regulatory bodies that wish to participate in the WHO International Clinical Trial Registry must comply with a set of specific criteria and minimum data collection requirements established by WHO. As such, there may be scope to include IVPP GSD as part of the information available in the WHO International Clinical Trials Registry Platform, which could make the platform an invaluable resource to directly retrieve information on the use of IVPP GSD.

\textsuperscript{36} The term “disclosure of origin” typically refers to the requirement that patent applications disclose the country of origin of genetic resources used in inventions. The World Intellectual Property Organization (WIPO) has been the primary forum for disclosure of origin discussions. In October 2000, the WIPO General Assembly established an Intergovernmental Committee (IGC) to address, inter alia, intellectual property issues in relation to the Convention on Biological Diversity. Discussions have focused on a proposal to require patent applications to disclose the origin of genetic resources and associated traditional knowledge in inventions, as well as to provide evidence of prior informed consent and benefit-sharing. Since 2009, the mandate of the WIPO IGC has been to undertake negotiations with the objective of reaching agreement on an “international legal instrument to ensure the effective protection of traditional knowledge, traditional cultural expressions and genetic resources”. The issue of disclosure requirements has also been discussed at the World Trade Organization (WTO) where several WTO Member States have proposed the inclusion of a disclosure of origin requirement in article 29 of the TRIPS agreement on the conditions on patent applicants.

Many countries already have this requirement in their national legislation. For example, section 8b. of Norway’s Patent Act states that “if an invention concerns or uses biological material, the patent application shall include information on the country from which the inventor collected or received the material (the providing country). […] if the providing country is not the same as the country of origin of the biological material, the application shall also state the country of origin.” Whereas in China, the Patent Law of the People’s Republic of China states that “with regard to an invention-creation accomplished by relying on genetic resources, the applicant shall, in the patent application documents, indicate the direct and original source of the genetic resources.”
a) **The WHO International Clinical Trials Registry Platform**

The WHO International Clinical Trials Registry Platform (ICTRP) provides a single point of access to specific information about ongoing and completed clinical trials.\(^{37}\) The ICTRP has developed best practice criteria that domestic registry bodies should use in the collection and management of data (WHO Registry Criteria and WHO Data Set)\(^ {38,39}\). This ensures consistency in the information that clinical trial bodies submit to the ICTRP.

The ICTRP Search Portal includes data from the following domestic registry bodies:

- Australian New Zealand Clinical Trials Registry (ANZCTR)*
- Brazilian Clinical Trials Registry (ReBec)†
- Chinese Clinical Trial Register (ChiCTR)*
- Clinical Research Information Service (CRiS), Republic of Korea†
- ClinicalTrials.gov*, United States
- Clinical Trials Registry - India (CTRI)†
- Cuban Public Registry of Clinical Trials (RPCEC)†
- EU Clinical Trials Register (EU-CTR)*
- German Clinical Trials Register (DRKS)†
- Iranian Registry of Clinical Trials (IRCT)†
- ISRCTN.org*
- Japan Primary Registries Network (JPRN)†
- Pan African Clinical Trial Registry (PACTR)†
- Sri Lanka Clinical Trials Registry (SLCTR)†
- The Netherlands National Trial Register (NTR)*

The WHO Data Set contains the minimum amount of information that must be included for a clinical trial to be considered fully registered. Currently, the WHO Data Set contains 20 items, including several Secondary Identifying Numbers.\(^ {40}\) The Secondary Identifying Numbers are other identifiers beside the Trial Identifying Number allocated by the Primary Registry. These can include:

- The Universal Trial Number (UTN)
- Identifiers assigned by the sponsor (record Sponsor name and Sponsor-issued trial number (e.g. protocol number))
- Other trial registration numbers issued by other Registries (both Primary and Partner Registries in the WHO Registry Network, and other registries)
- Identifiers issued by funding bodies, collaborative research groups, regulatory authorities, ethics committees / institutional review boards, etc.

---


The ICTRP database is updated on a weekly basis for registries marked (*) and on a monthly basis for those marked (†).

\(^ {40}\) These are: Primary Registry and Trial Identifying Number; Date of Registration in Primary Registry; Secondary Identifying Numbers; Source(s) of Monetary or Material Support; Primary Sponsor; Secondary Sponsor(s); Contact for Public Queries; Contact for Scientific Queries; Public Title; Scientific Title; Countries of Recruitment; Health Condition(s) or Problem(s) Studied; Intervention(s); Key Inclusion and Exclusion Criteria; Study Type; Date of First Enrollment; Target Sample Size; Recruitment Status; Primary Outcome(s); Key Secondary Outcomes.
There is no limit to the number of secondary identifiers that can be provided. Thus, IVPP GSD could potentially be included as a secondary identifier.

The source of the Secondary Identifying Number for IVPP GSD could be its accession number. This could be expressly included in the WHO Data Set by requiring that any clinical trial that uses a product developed from IVPP GSD – such as vaccine trials – include the relevant Accession Number(s) in the Secondary Identifying Number form.

1.2.5. Published Results
As explained in section 2.2.1, most journals require disclosure of data used to conduct an experiment, including GSD. Thus, journal articles may provide a useful source for IVPP GSD monitoring.

Contrary to clinical trial documentation, journal articles have the added benefit of potentially capturing the use of IVPP GSD in animal studies, which are often conducted to test influenza end-products, such as vaccines. The results of animal studies are typically not reported unless published, and so unlike human studies, journal articles may be an important downstream option for tracking the use of IVPP GSD early in the research and development process.

However, published data is inherently selective for research that has been completed, reviewed and selected for publication. Therefore, journal articles may serve as a downstream input source for the monitoring use IVPP GSD to supplement other sources.

1.2.6. Regulatory Files
In many jurisdictions, public access to applications for regulatory approval of vaccines, antivirals, and diagnostics is limited and sometimes prohibited by law. For example, under 21 US Code of Federal Regulations 314.430, no data or information in an application for regulatory approval is available for public disclosure before an approval letter has been sent.

Decisions by regulatory authorities, for example 510(k) decision summaries for diagnostics by the US Food and Drug Administration, assessment reports by the European Medicines Agency or summary basis of decision from Health Canada, may contain information on vaccine, antivirals and diagnostics manufacturing methods and processes, including biological or GSD sources. However, this information may not be available for public disclosure unless it has been previously disclosed to the public or it relates to a product or ingredient that has been abandoned and it does not represent a trade secret.

Therefore, regulatory files are likely to be an incomplete source of information in their current form. However, applicants seeking regulatory approval for an influenza product may be open to updating the information publicly provided in approved regulatory applications where GSD has been used to develop an end-product. This could include adding a reference to any relevant accession numbers or the GISRS origin of the IVPP GSD used in the end-product development.

41 In limited circumstances, the US FDA Commissioner may release limited data on safety and efficacy. 21 US CFR 341.430(d)(1).
Best Process to handle IVPP GSD under the PIP Framework

29 September 2015; Revised 6 May 2016

For example, this short statement could be included:

Note: This [product] has been developed using the following nucleotide sequences of influenza viruses with human pandemic potential submitted by [name of laboratory]: ([database] accession number(s): XXX123456) accessed from [name of database].

1.2.7. Voluntary Disclosure

At the very end of the downstream options, monitoring the use of IVPP GSD may include voluntary disclosure of the use of IVPP GSD to WHO. Under this system, users of IVPP GSD would voluntarily report their use of GSD to research and develop an end-product. WHO could encourage voluntary disclosure through promotion of the objectives of the Framework and the importance of benefit sharing to ensure continued access to IVPP GSD. This options would need to be examined further to determine feasibility and operationality. An option could be to gather information on the use of IVPP GSD either through a questionnaire (for example, using the Partnership Contribution Collection questionnaire) or by using an online registration system. However, because reaching out to all users may be very difficult, such mechanisms would probably need to be used in conjunction with the other options discussed previously.

III) CONCLUSION

WHO has already commenced investigation into a possible search engine for a downstream monitoring system with the World Federation for Culture Collections (WFCC) and the World Data Centre for Microorganisms (WDCM). The tool they have developed searches different databases to identify uses of IVPP GSD in scientific publications, patents, clinical trial files and regulatory approval files.

As discussed in this paper, the feasibility of such a system will depend on several factors.

First, identifying uses of IVPP GSD to generate end-products will likely require a combination of both upstream and downstream options.

Second, comprehensive monitoring will depend on publicly-accessible source identification from multiple sources. As highlighted in this paper, there is currently no consistent practice of identifying use of IVPP GSD in publicly available downstream documents, such as patents, research trial documentation, and regulatory files. To encourage this, WHO could issue guidance to users of IVPP GSD specifying that all uses of IVPP GSD to generate an end-product should be acknowledged using accession numbers.

Third, implementing this system will require cooperation by a number of entities, including databases, GISRS laboratories, industry and other stakeholders. Best practices or guidance could therefore be developed in consultation with these entities.

Lastly, although the objective of this paper is to identify options to monitor the use of IVPP GSD, an
alternative could be to monitor uses of IVPPs in all forms (GSD and physical materials). Thus, instead of tracing accession numbers, WHO could search for mentions of the name of the virus (e.g. A/Anhui/01/2013), which is a unique identifier. This would have the added benefit of providing more information to WHO in order to identify potential contributors to the Partnership Contribution and to better understand the influenza research and development process.