16 December 2015

PANDEMIC INFLUENZA PREPAREDNESS (PIP) FRAMEWORK ADVISORY GROUP

MEETING OF THE TECHNICAL WORKING GROUP ON THE SHARING OF INFLUENZA GENETIC SEQUENCE DATA (TWG)

29-30 SEPTEMBER 2015, GENEVA, SWITZERLAND

Meeting Report

Organization and process of the Meeting

1. The Technical Working Group ("TWG") met at WHO headquarters in Geneva, 29-30 September 2015.

2. There were three members of the PIP Advisory Group ("PIP AG") and 10 experts present. The list of participants in the meeting is provided in Annex 1. Three invited experts were unable to join the meeting; they are also listed in Annex 1.

3. The meeting was Chaired by Didier Houssin.

4. The Chair welcomed participants.

5. The TWG adopted the meeting agenda. The agenda is provided in Annex 2.

6. The WHO Legal Officer reviewed the process for Declarations of Interest and the undertaking on confidentiality. The summary of Declarations of Interest is found at Annex 3.

7. The Chair made introductory remarks, reviewing progress since 2013 in implementing the best process to handle genetic sequence data for influenza viruses with human pandemic potential ("IVPP GSD") under the PIP Framework and the reasons why the PIP AG convened a Technical Working Group ("TWG") in 2015. The Chair explained that the task of the TWG was to develop a draft document ("TWG document"), for the consideration of the PIP AG, defining the optimal characteristics of a GSD sharing system best suited to meet the objectives of the PIP Framework. The TWG document should also include best practices for operationalizing such a system. It will then be shared widely with relevant stakeholders (e.g. GISRS labs, industry associations, databases, academia and civil society organizations) for input. Following finalisation of the TWG document in 2016, the PIP AG will develop its advice for the Director-General.

8. The WHO Secretariat updated the TWG on the progress of related projects. Firstly, a paper has been developed identifying different options for monitoring the use of GSD in end products, including checking patent applications, clinical trial filings and regulatory approvals for evidence of IVPP GSD use. WHO is working with the World Federation for Culture Collections and World Data Centre for Microorganisms to develop a search engine to centralise this information. Secondly, WHO is conducting a survey of existing systems for sharing influenza GSD. The survey has been sent to around 150 stakeholders, including laboratories, industry and academia, of whom 17 responded by 25 September 2015. The report on the survey will be finalised in December 2015.

9. The Chair reviewed the method of work for the TWG that was established during the TWG's 27 July 2015 teleconference. Three sub-groups, each led by a team coordinator,

were set up to prepare some initial draft text on the specific technical characteristics identified by the TWG's Terms of Reference for inclusion in the draft TWG document. The Chair circulated a copy of this draft text to meeting participants and explained that the objective of the two-day Geneva meeting would be to discuss and develop the draft optimal characteristics and best practices, in order to produce a complete draft TWG document for the PIP AG's October meeting.

Section 1: Obligations and Expectations of Data submission

10. The TWG discussed two main approaches to sharing sequence data. Currently some countries prefer to submit IVPP GSD to a publicly-accessible database, such as GISAID EpiFlu™, that requires all users to sign an access agreement that sets conditions on use of the data. This mechanism is appreciated by developing countries as it requires that data providers and originating laboratory be credited for the data. It also provides the ability to track those who access the data, who may subsequently use these sequences to develop commercial products. Some participants argued strongly that an obligation to submit data to a database without these safeguards might result in some countries choosing not to share their IVPP GSD at all. They argued that the GISAID data access agreement was reasonable and complementary to the PIP Framework.

11. Other participants said they preferred that IVPP GSD be submitted to databases with no restrictions on the use of data, such as GenBank. They put forward the view that any data access and use agreement could act as a barrier to the free flow of information for research activities. One participant said more effort could be made to provide incentives to countries to share their data in the public domain databases.

12. Participants with database expertise said that it was relatively straight forward to include information about contributing laboratories and to require that users acknowledge the contribution of laboratories, but it would be much more difficult to be involved meaningfully in policing the use made of IVPP GSD.

13. Participants discussed how monitoring the use of the data could either be done at the database end (upstream) or at the endpoints by monitoring the use of IVPP GSD in commercial product development. The latter approach was considered more feasible. In order to monitor the use of IVPP GSD, it was suggested that one technical solution might be to use the GenBank structured comments field to identify IVPP GSD as being subject to benefit-sharing expectations under the Framework.

14. The Chair said both approaches to data sharing were understandable, reflecting the differing backgrounds of providers. He suggested there was scope for greater cooperation: databases with different conditions of access could complement each other regarding short-term and longer-term access and together make up a more coherent system.

15. The Chair suggested developing an operational practice that would combine and take advantage of the two approaches. The group discussed a proposal for a best practice that would require sharing IVPP GSD with the INSDC following a period of 6 months after submission to GISAID EpiFlu™. It was agreed that this best practice should include an option to "opt out" rather than ask for individual data providers' consent. Therefore, participants could opt out of their data being shared with INSDC.

16. The Secretariat suggested that best practices should be defined from the perspective of implementing the PIP Framework. The data sharing system needed to foster three main points: sharing of IVPP GSD based on trust and the technical aspects of the system; data to be made available quickly for rapid risk assessment; and ease of benefit tracking. Since the introduction of the PIP Framework the overall momentum of virus sharing had grown.

17. Summarizing the discussion, the Chair proposed that the draft TWG document present two forms of best practice regarding the obligations of IVPP GSD submission and data sharing, one in general terms about GSD sharing and one a best practice under the PIP Framework. The latter would suggest that data and metadata submitted to GISAID EpiFlu™ should, after a period of six months, be shared with INSDC unless the data provider exercised the opt out. Participants agreed the suggestions of a six month timescale would be used to survey data providers' reactions to this proposal. The PIP AG could decide between possible options when it formulated its advice to the Director-General. This approach was agreed by participants.

Section 2: Timeliness of Data Submission

18. It was agreed that the data sharing system should encourage submission of IVPP GSD to a database as soon as the data were available (including preliminary or incomplete sequences). Following discussion, participants agreed that IVPP GSD should be submitted to a database within one month of becoming available and that the system should then provide public access to submitted IVPP GSD within 24 hours of data submission.

19. Most participants supported a suggestion that immediate sharing of new IVPP GSD would be encouraged if the newly submitted sequences were covered by a temporary publishing embargo. This would mean data were made immediately publicly accessible but that they could not be submitted for scientific publication within, for example, 60 days except by the original data providers. Such an arrangement would allow rapid access to IVPP GSD for public health authorities to conduct risk assessments and engage in emergency pandemic response activities, while also protecting the original data submitter's first rights of publication. The temporary publishing embargo could be lifted early by the data submitter if a manuscript describing the data was published before expiration of the embargo, or at the data submitter's discretion.

20. The approach would mirror similar efforts in other fields that have proved successful, such as the genomics Fort Lauderdale[1] agreement and the Toronto Statement on prepublication data sharing[2]. Participants said an embargo would enable those on the front line of an emergency public health pandemic influenza response to publish their work in peer-reviewed journals after sharing the IVPP GSD. It would be preferable to the current system whereby data providers who submit data to, for example, GenBank can choose the option not to make the sequence available to other users until after publication.

21. Participants discussed how such an embargo could be implemented. It was suggested that the submitting laboratory could flag their IVPP GSD as being covered by the temporary publishing embargo. The embargo would only cover publication in peer-reviewed scientific journals and would not affect data analyses, risk assessments, vaccine development or other associated research and development activities that used the data. The timescale for the embargo was suggested to be 60 days but it was agreed that this would need to be discussed. The activities covered by the embargo would also have to be carefully defined.

[1] Wellcome Trust, "Sharing Data from Large-scale Biological Research Projects: A System of Tripartite Responsibility: Report of a meeting organized by the Wellcome Trust and held on 14–15 January 2003 at Fort Lauderdale, USA." (Fort Lauderdale Statement), available at https://www.genome.gov/Pages/Research/WellcomeReport0303.pdf

[2] Toronto International Data Release Workshop Authors, "Prepublication data sharing", Nature 461, 168-170 (10 September 2009), http://dx.doi.org/10.1038/461168a

22. One participant argued against an embargo system, saying it would not be a productive approach for the academic community. Another asked about situations where very similar IVPP strains emerged in animals and humans at the same time: the former would not be covered by the embargo but the latter would.

23. It was agreed that the practical aspects of implementing an embargo should be discussed with journal editors by TWG members and the Secretariat, including the question of how to enforce the embargo system. Participants felt that it was unrealistic to expect the reviewers of journal papers to police an embargo system. Participants agreed that self-policing by the scientific community was most likely to play an important role.

24. One participant updated TWG members on recent developments in the field of scientific publication that could offer alternative or complementary mechanisms for immediate release of IVPP GSD. These include "self-referencing" platforms, such as bioRxiv and Nature Scientific Data, which enable authors to make their findings available to the scientific community before submitting papers to academic journals.

25. Another participant, backed by the Chair, suggested developing further incentives to release IVPP GSD quickly. This could be achieved by giving public recognition to scientists and institutes sharing the data, including by establishing best practices for papers and journals to acknowledge the timely manner in which data had been uploaded and the role this played in the wider subsequent research; public health pandemic preparedness organizations should also acknowledge the role of early data release.

Section 3: Quality Assurance of Data

26. Participants recognised that there will always be a trade-off between rapid data submission and data quality. However, it was agreed that quality standards should not block the sharing of early, possibly imperfect or incomplete data that can still be important for public health risk assessment and the development of vaccines, diagnostics and other medical products for influenza pandemic preparedness. There were therefore two issues: first, to decide the essential and ideal characteristics of IVPP GSD and metadata; second, to discuss how a data sharing system can best accommodate the early submission of preliminary data that is incomplete or has not been fully quality assured, accompanied by clear caveats about the quality status. Participants agreed that full submission of an IVPP sequence usually requires several updates and that quality assurance and data control is a dynamic layered process that may require several updates.

27. Participants agreed the draft TWG document should make it explicit that data quality can be improved several times. Tiered quality criteria should capture this dynamic process by specifying mandatory (minimum) standards for IVPP GSD initial submissions of different types and the expectations for subsequently improved data. The data sharing system should enable data providers to come back several times to update their data and remove earlier errors. A historic record of the different versions of the GSD should be maintained.

28. One participant with database expertise said it would be possible to annotate a sequence as draft data (and similarly for metadata) to indicate that it was not fully quality controlled. The quality status could then be revised with each update.

29. The Chair said that responsibility for quality control and quality assurance lay across the data sharing system and everyone involved in the data sharing system should be responsible for quality in their own area of activity. This would start with good laboratory practice, followed by database system tools for GISRS laboratories to support quality

control and detect the presence of potential sequence artefacts. It was agreed that quality evaluation should make a distinction between the quality of IVPP GSD, the completeness of the data, and the quality status of the metadata.

30. One participant explained that the point at which data was uploaded provided the key opportunity for quality improvement and curation of IVPP GSD. Real time quality feedback to data submitters against defined quality criteria would enable uploaders either to make immediate adjustments or to provide annotations about known issues. The data submission interface should provide tools that support compliance with agreed standards. Databases should work together to develop tools and avoid duplication of effort by creating separate tools.

31. Participants discussed how the quality criteria and the quality assurance process needed to respond to the rapid technological changes in the field of GSD. The criteria would need to be fluid over time so their detail could be regularly reviewed and maintained by an appropriate group from across the PIP Framework stakeholder community. It was agreed that this approach to "externalising" the quality criteria would ensure they reflected new technologies and timescales. The draft TWG document should therefore describe the suggested core elements of this tiered structure but the detail should be captured and maintained in separate standards document.

32. One participant suggested it would be useful to develop a score reflecting the quality of the submission, based on the presence of features of poor quality data and quantification of the completeness of the sequence record and metadata.

33. Metrics for accuracy, completeness and quality should be defined so that data submitters could be guided to make the necessary improvement to their GSD, and users of the IVPP GSD could judge the appropriateness of the data for downstream use.

34. GISAID EpiFlu™'s submission system was seen as an example of quality control at the point of submission. One participant suggested that if data were initially routed through GISAID EpiFlu™, and the GISAID EpiFlu™ submission system was enhanced to support the quality standards as they evolved over time, then this would be an improvement on the current system.

Section 4: Upload and Completeness of Data Annotation

35. The participants discussed the need to have a tiered approach to metadata so that there would be: a "core" set of standardized mandatory metadata for all data; additional mandatory metadata according to the different type of source sample; and optional metadata elements that would ideally be provided over time. The core sets should be kept to a minimum so that they do not serve as a barrier to early submission of IVPP GSD.

36. The status of a virus as IVPP (as defined by the PIP Framework) should be prominently indicated in a metadata field that is easily searchable. This would help to inform data users that they have accessed IVPP GSD that are covered by the PIP Framework. There was some discussion about whether an alternative approach could be to indicate PIP Framework status in a virus name.

37. It was agreed that the TWG's draft document should include proposed core and optional metadata requirements for IVPP GSD but that the detailed structure and definitions of core and optional metadata for different types of IVPP GSD should be "externalised" to a separate standards document that would evolve over time, mirroring the approach agreed for the quality standards for the actual sequence data (see Section 3).

38. The participants agreed that the system should also support the curation of metadata and the ability to specify versions and updated metadata as improved information became available. The process of data providers updating submissions over time with improved metadata would resemble what had been discussed for the sequences themselves (see Section 3).

39. Similarly, participants said that quality assurance processes at the time of metadata upload should be incorporated. The system should automatically assess the completeness of the uploaded metadata and its compliance with data standards and then provide real time feedback to the data submitter, detailing any detected metadata anomalies that needed investigation/resolution by the submitting GISRS or other laboratory. This real time feedback would encourage data providers quickly to address any shortcomings in the metadata.

40. The question of confidentiality of patient information and whether this posed difficulties for submitting optimal metadata for IVPP from humans was raised and discussed. It was agreed that submission of some metadata would sometimes need to wait until it could no longer identify an individual patient, for instance when there had been a large enough number of cases.

41. One participant with database expertise suggested that consistency of metadata was important to ensure interoperability across different databases for datasets from different sources. Metadata sharing across different databases should be facilitated by the use of common structured metadata submission templates and data dictionaries, structured data capture online user interfaces, and/or xml database schema.

Section 5: Ease of access to and use of IVPP GSD

42. One participant set out four points to consider on access to and use of IVPP GSD. First, the data sharing system should make it relatively easy to identify IVPP GSD and download it in the standard formats that other analysis tools might use. Second, the system should provide access to the IVPP GSD using third party bioinformatics analysis software tools; two ways should be considered: a) the user would download IVPP GSD and use them locally on their own computer with third party tools; b) the system would provide access to the data through a programmatic interface whereby the tools could directly access the data within the context of the data sharing system. It was agreed both options should be considered. Third, the system should provide for access and use of IVPP GSD by external resources through interfaces. Fourth, the data sharing system should provide users with a common data access and use agreement that outlined the expectations of the PIP Framework and required the person requesting the data to agree to specified terms of use.

43. The Secretariat added that, in its April 2015 Meeting Report to the Director-General, the PIP AG recommended that the Secretariat consider initiating discussions with database managers on the possibility of including a statement about the PIP Framework on their websites. The Chair asked participants if all databases would be likely to accept responsibility to flag information about the PIP Framework to data providers and users.

44. Participants discussed implementation across databases of a common data access and use agreement or statement on the PIP Framework. The majority of participants agreed that a data access and use agreement, or a statement, was important for benefit-sharing under the Framework. One participant said it was challenging to educate people about the PIP Framework and notifying people up front at the point of data access would help in this

regard. However, some participants disagreed on how this would be implemented and enforced, and by whom. There was also debate on the terms of the data access and use agreement and/or the wording of a statement. These could range from users simply being asked to acknowledge an informal statement informing them about the existence of the PIP Framework, to being asked formally to enter into a legally binding agreement containing requirements that users respect PIP Framework expectations.

45. One participant with database expertise said it would be possible to add mention of the PIP Framework into its general data use agreement but stressed that most databases would not want to be involved in enforcing a legal agreement that requested users to comply with the PIP Framework. Participants said this responsibility should lie with whoever enforces the Framework.

46. Another participant explained that, for unrestricted public access databases, the tradition has been not to attach any constraints or licensing requirements but also not to take part in presenting data use agreements. This included data providers making any reference in the structured comments field to restrictions or data use agreements. Presenting any such agreement implied some form of contract, he said, and would open up challenging issues such as how to confirm it had been seen, how to track users, and what level of involvement to have with the entity(ies) that polices the agreement. Currently these databases carry only general statements on use of the data, including that there may be intellectual property claims or that licensing restrictions may exist, but they do not try to police these relationships. He advocated for a clearly stated informal agreement on data use, and that more effort be put into a sophisticated and energised education program to make sure everyone understood their PIP responsibilities. It was agreed that implementation needed to be acceptable across all databases.

47. The Chair asked for clarification from participants on how to request the addition of a statement about the PIP Framework to the GenBank general notice to users. A participant said changes to the statement were rare and were kept very generic. To pursue this, the WHO should contact the three collaborating INSDC partners or explore the idea informally with people involved in the process.

48. The discussion that followed also raised a number of practical issues: an individual agreeing to the data access and use agreement might not have the institutional authority to do so; seeking authority would be cumbersome if a data access and use agreement appeared for each individual download of IVPP GSD; one individual might access IVPP GSD through several different email addresses making them difficult to identify.

49. The question of the automatic sharing of IVPP GSD between databases also highlighted different perspectives on how freely available the sequences should be. While some participants were in favour of open transfer, others again stressed that some Member States had specifically chosen the type of database to which they preferred to submit their sequences and would not want them automatically shared with other databases. Some participants reiterated that they feared some Member States might as a result stop sharing IVPP GSD. The Chair suggested that Member States should be asked by WHO for their views, pointing out that the PIP Framework was specifically designed to protect Member States' interests through the benefit sharing element.

50. Some participants again suggested that sharing of IVPP GSD between databases could be delayed until after a specified period of time following the original submission, for example six months, with the time period to be discussed (see Section 1). Thus one solution could be to modify the relevant part of the data access and use agreement on the GISAID Initiative website so that the data would be released to GenBank after a six month period in the GISAID EpiFlu™ database.

51. The Chair reiterated that the spirit of the PIP Framework was about rapid data sharing but that it also included benefit sharing particularly for the poorest countries and there were concerns that this part of the deal might be lost. Implementation of a common PIP Framework statement across all databases would not only inform users about PIP Framework context and expectations but also open the door to further cooperation and initiatives.

52. The Secretariat said the idea of databases carrying a common statement on the PIP Framework would be interesting and could provide legal certainty for manufacturers, some of whom when contacted by WHO for benefit sharing contributions have said they were unaware of the PIP Framework.

53. The Chair suggested that, given the spectrum of views, the TWG could suggest different options in its draft TWG document regarding the issues of a common database statement on the PIP Framework and the sharing of data between databases. Under one option, a general reference and link to the PIP Framework might be sufficient as a minimum. Under the second option, a fuller statement on the PIP Framework obligations would be appropriate, and the document could suggest some draft text for this.

54. One participant suggested a data access and use agreement that exists as a contract agreed to by data users, and which is mediated by the PIP Framework, but is not implemented in any way by the technical database part of the data sharing system. As with the Toronto Statement, agreements can exist outside of the mechanisms that are used for transferring the data.

55. It was agreed all references/links and statements about the PIP Framework would also have to be visible to people accessing IVPP GSD through a programmatic interface instead of via a web page.

56. For practical reasons, it was agreed that an individual or organisation should only have to sign up to a data access and use agreement for IVPP GSD once, but annual reminders of the PIP Framework expectations should be sent out by databases to make sure that they were not forgotten by data users.

57. One participant envisaged the possibility that GISAID EpiFlu™ could be the primary repository for a sequence, which would therefore be covered by the GISAID access terms and conditions for the agreed period, but that the sequence could then be accessed from another database using a programmatic interface. This would mean that users could analyse GISAID EpiFlu™ data using novel analysis tools of another database. Such an arrangement would be an important advancement in current accessibility of the data, said one participant.

58. In addition to databases carrying the proposed reference/link or statement on the PIP Framework, TWG members said it was important to set up and operate an education programme to ensure the PIP Framework was more widely understood. Journal editors and reviewers should be included in the educational and information activities.

59. Finally, it was agreed that it was also important that the data sharing system should support the linkage between sequence data (and their metadata) with epidemiological data based on WHO reports, OIE reports, FAO reports, national reports of influenza cases or outbreaks or based on surveillance activities. This could be implemented based on interoperability strategies between databases, as is already in place between EMPRES-I/OpenFlu/IRD for animal, environmental and human zoonotic influenza.

Section 6: Sustainability and Security of the System

60. Participants agreed IVPP GSD data and metadata are valuable and need to be available in perpetuity. However, all database systems depend on financial support, which cannot be guaranteed long term. Therefore it was suggested that one approach was to disseminate the data among many sources so that sustainability did not depend on any single database.

61. Another participant argued that consortium funding of groups of databases promotes sustainability: if any one resource loses its financial support, the data will still be available through another resource. Collaborative structures also promote varied feedback and accountability. Another possible solution would be to put in place agreements with governments to preserve IVPP GSD in the event that support for any particular database no longer be available.

62. Another participant said the level of financial support was also important as resources determine the available infrastructure for data sharing. He also argued that greater collaboration would also help leverage the total support provided to databases, with scope for example to share tasks and the development of tools for analysis.

63. Participants agreed that the IVPP GSD system security should ensure that IVPP GSD and metadata cannot be altered by outside parties without permission of the data provider. However, it was agreed that the wording in the draft document should allow for a data provider to cede editorial control over IVPP GSD to designated representatives. Otherwise, the database system would become blocked by data that could not be updated due to the original provider no longer being involved in the work. In some instances it might be appropriate for third parties to provide comments or information by annotating the data; some databases already offer this functionality and it could be introduced more widely through a second layer of annotation.

64. Defences against intentional breaches of the integrity of the data are essential. Asked by the Chair about whether cyber-attacks had taken place on GSD databases, participants with knowledge in this area said the databases had good security systems against such attacks and all data was backed up off-site, ensuring that primary data can always be retrieved.

65. It was agreed that best practices on security should require databases to comply with best technical practices. This should include maintaining historical backups of their data content, which is important for recording modifications to the data and tracking any malicious behaviour.

Section 7: Source Identification

66. Information about the use of IVPP GSD by commercial entities during the development and/or manufacture of pandemic influenza preparedness products is important for benefit sharing under the PIP Framework. Participants continued earlier discussions on two possible approaches to source identification. One option would be to record the identity of the commercial entities that download a specific IVPP GSD, each of which has a unique identifier. However, this approach is not typical practice for most databases (GISAID EpiFlu^{TM} being the main exception) and would require modifications to standard operating procedures. In addition, monitoring access does not in itself provide information on how an IVPP GSD has subsequently been used. The second option would

be to start from each influenza-related end product and trace backwards to identify the IVPP GSD used in its development. This approach is less disruptive to most current models of GSD access and sharing but creates challenges related to the complexities of IVPP GSD itself. For instance, many sequences are similar so it would be necessary to trace the original identifier; companies can introduce deliberate changes to a sequence; it is difficult definitively to link a product to a specific IVPP GSD.

67. All GSD submissions, regardless of the database, are automatically accompanied by a unique accession number. This number could be readily identified and used by IVPP GSD users to identify correctly the use and origin of specific IVPP GSD. As part of this approach, sequences would need to be clearly identified as PIP Framework IVPP GSD at the time of upload so that users were aware of their status (see Section 4). Participants said that such a system, based on identification of IVPP GSD used to develop commercial products, would depend on the good faith of manufacturers of vaccines, anti-virals and relevant diagnostics and would need to build on the relationship between WHO and industry that has strengthened over the past five years. The challenge would be adherence by smaller companies that are unaware of the PIP Framework. It was suggested that these entities would need to be identified and educated about the PIP Framework. One participant raised the additional complication that some products are based on a large number of sequences and that identifying and valuing the contribution of any single one might be difficult.

68. Participants with database expertise reiterated that, from their point of view, it was possible to monitor who was accessing IVPP GSD but that some databases would not want to support the policing of how it was subsequently used.

69. On the status of animal viruses, one participant clarified that routine animal surveillance and diagnostic samples from animals were not considered to be PIP biological materials except when used for the development of pre-pandemic or pandemic candidate vaccine viruses as recommended by the WHO. It was suggested that this could create an incentive for the veterinary sector to submit more sequences and share their data. Another participant said the animal side generally needed to be more informed and encouraged by incentives to contribute by sharing sequences.

70. The meeting decided that the draft TWG document's relevant proposed content for Section 7 was already covered under Section 4, so this section of the document should be deleted and replaced by a reference to the earlier section.

Section 8: Support for the Regulatory Process

71. Participants who had looked into this issue said the designation of a sequence as PIP Framework IVPP GSD in itself appeared unlikely to influence, facilitate or simplify the process of obtaining regulatory approval for pandemic influenza preparedness products derived from these data. Nor did any specific mechanism for controlling IVPP GSD have the potential to streamline the regulatory process. Therefore, the focus should be on ensuring that regulatory authorities have easy access to IVPP GSD if, and when, they require the sequence data.

72. Some regulators consulted by the participants in preparing the TWG document reported having had institutional difficulties in signing data access and use agreements similar to that of the GISAID Initiative (although, according to one participant, it has been flexible in these situations and is also considering changing some language that causes problems for regulators).

73. Some participants suggested that an optimal data access and use agreement would implicitly allow the system and data users to share IVPP GSD with regulators without any additional agreement. However, it was also suggested that there could be separate agreements for regulators. The consensus was that the data sharing system should offer regulatory authorities acceptable terms for access to IVPP GSD. The confidentiality of manufacturer/regulator interactions should make this easier to implement and control.

74. Quality assurance and quality control issues are important for regulatory agencies. Specifically, it was agreed that an important database function is to provide information on the quality cycle, the curation environment and relevant standards in a searchable, easily useable and downloadable form that gives data users and regulators user-friendly access to this information.

Next steps and concluding comments

75. The Chair and participants agreed a mechanism for producing a revised full draft of the document and the subsequent next steps:

    a. The Secretariat would prepare an updated draft, incorporating the amended optimal characteristics and best practices discussed during the meeting, plus supporting text that reflected in a consistent manner the points raised during the two-day meeting.
    b. The new version would first be sent for comment to the three sub-group coordinators who had prepared the draft text for this meeting. It would then be revised and sent to all TWG members for comment.
    c. A revised version would be submitted to the PIP AG in time for discussion at its 15 October 2015 meeting.
    d. Given the short deadline, there would also be an opportunity for further input from the TWG following the PIP AG meeting when there would be time available to refine the document further.
    e. An agreed version would be sent out for the wide consultation with stakeholders.
    f. Following the consultation, the TWG would again be asked for feedback and input into a revised draft.
    g. A final version of the document would be agreed and submitted by the TWG to the AG by April 2016.
    h. The PIP AG could then draw on the document when providing advice to the Director-General on how to manage IVPP GSD under the PIP Framework.

76. Further work on the document would include:
    a. discussion by TWG members and the Secretariat with scientific journal editors on the proposed temporary publishing embargo;
    b. consideration of new platforms such as bioRxiv and Nature Scientific Data as possible alternatives/additions to a temporary publishing embargo;

77. One participant suggested the overall document should incorporate a monitoring and evaluation approach with metrics, where possible. These metrics could be at the activity level (i.e. assessing whether best practices had been implemented) and at the impact level (i.e. assessing the results of data sharing). In addition it would be informative to monitor the amount of PIP material downloaded and the level of interaction between databases.

78. The Chair confirmed that the TWG members were the authors of the document and would finalise the content.

79. The Chair thanked participants for their work so far on the document and closed the meeting.

Annex 1

PANDEMIC INFLUENZA PREPAREDNESS (PIP) FRAMEWORK ADVISORY GROUP

MEETING OF THE TECHNICAL WORKING GROUP ON THE SHARING OF INFLUENZA GENETIC SEQUENCE DATA (TWG)

29-30 SEPTEMBER 2015, GENEVA, SWITZERLAND

List of participants

PIP Advisory Group members	Affiliation
Prof. Didier Houssin (Chair)	French Evaluation Agency for Research and Higher Education, France
Dr Olav Hungnes	Norwegian Institute of Public Health, WHO National Influenza Centre, Norway
Dr Oleg I. Kiselev	Research Institute of Influenza, Ministry of Public Health and Social Development, Russian Federation

Experts participants	Affiliation
Dr Guy Cochrane	European Bioinformatics Institute (EMBL-EBI), European Molecular Biology Laboratory, United Kingdom
Dr Nancy Cox	Formerly Influenza Division, Centers for Disease Control and Prevention and GISAID Scientific Advisory Council, USA
Dr Gwenaëlle Dauphin	Animal Health Service, Food and Agriculture Organization of the United Nations (FAO), Italy
Othmar Engelhardt	WHO Essential Regulatory Laboratory, Division of Virology, National Institute for Biological Standards and Control, UK
Dr Otfried Kistner	Formerly Baxter Innovations; Senior Consultant & Independent Vaccine Expert Austria
Dr Richard Scheuermann	Influenza Research Database and Director of Informatics, J. Craig Venter Institute, USA
Prof. Marietjie Venter	One Health Program, Global Disease Detection, US Centers for Disease Control and Prevention, South Africa
Dr Richard Webby	WHO Collaborating Centre for Studies on the Ecology of Influenza in Animals, St. Jude Children's Research Hospital, Department of Virology and Molecular Biology, USA
Dr David Wentworth	Virology Surveillance and Diagnosis Branch, Influenza Division, Centers for Disease Control and Prevention, USA
Prof. Ioannis Xenarios	Swiss-Prot and Vital-IT Group, Swiss Institute of Bioinformatics, Switzerland
Dr Keith Hamilton (absent)	Formerly World Organization for Animal Health; Kansas State University, USA
Dr Marilda Siqueira (absent)	Instituto Oswaldo Cruz, WHO National Influenza Centre Brazil
Dr Dayan Wang (absent)	China Centre for Disease Control and Prevention, WHO Collaborating Centre China

Annex 2

PANDEMIC INFLUENZA PREPAREDNESS (PIP) FRAMEWORK ADVISORY GROUP

MEETING OF THE TECHNICAL WORKING GROUP ON THE SHARING OF INFLUENZA GENETIC SEQUENCE DATA (TWG)

29-30 SEPTEMBER 2015, GENEVA, SWITZERLAND

Agenda

Tuesday, 29 September 2015

09.00 – 09.30	Registration
09.30 – 10.00	Welcome from the Chair o Introductions o Adoption of Agenda o Declarations of Interest
10.00 – 10.30	Background and objectives of the meeting
10.30 – 11.00	Coffee
11.00 – 12.30	Presentation of Subgroup 1 and discussion
12.30 – 13.30	Lunch
13.30 – 14.30	Discussion of Subgroup 1 (continued)
14.30 – 15.00	Coffee
15.00 – 17.00	Presentation of Subgroup 3 and discussion
17.00 – 17.30	Wrap-up

Wednesday, 30 September 2015

09.00 – 10.30	Presentation of Subgroup 2 and discussion
10.30 – 11.00	Coffee
11.00 – 12.30	Discussion of Subgroup 2 (continued)
12.30 – 13.30	Lunch
13.30 – 15.00	Discussion of Best Practices
14.30 – 15.00	Coffee
15.00 – 17.00	Finalization of TWG draft document
17.00 – 17.30	Wrap-up & Next steps

Annex 3

PANDEMIC INFLUENZA PREPAREDNESS (PIP) FRAMEWORK ADVISORY GROUP

MEETING OF THE TECHNICAL WORKING GROUP ON THE SHARING OF INFLUENZA GENETIC SEQUENCE DATA (TWG)

29-30 SEPTEMBER 2015, GENEVA, SWITZERLAND

Summary of Declarations of Interest by members

- In accordance with WHO policy, in advance of this meeting, all experts were asked to provide a duly completed Declaration of Interests to inform WHO about real, potential, or actual Conflicts of Interest that they might have in relation to the subject matter of this meeting.

- Over the course of this meeting the Technical Working Group developed a document, for the consideration of the PIP Advisory Group, on the optimal characteristics of an influenza genetic sequence data sharing system that is best suited to meet the objectives of the PIP Framework.

TWG experts serve in their individual capacity acting as international experts serving WHO exclusively. The following interests and/or affiliations were disclosed to the Secretariat and were deemed relevant to the subject of the TWG's work:

Name	**Interest declared**
Guy Cochrane	Team lead of the European Nucleotide Archive
Nancy Cox	Formerly affiliated with a GISRS laboratory; co-Chair of the GISAID Scientific Advisory Council
Othmar Engelhardt	Affiliated with a GISRS laboratory, receives funding from IFPMA
Olav Hungnes	Affiliated with a GISRS laboratory
Otfried Kistner	Formerly employed by Baxter Innovations GmbH, currently consulting for Nanotherapeutics and Takeda
Oleg Kiselev	Affiliated with a GISRS laboratory
Richard Scheuermann	Receives funding from the NIH for the development of the Influenza Research Database
Marilda Siquiera	Affiliated with a GISRS laboratory
Richard Webby	Affiliated with a GISRS laboratory; serves on the Scientific Advisory Board of US BARDA
David Wentworth	Affiliated with a GISRS laboratory, formerly employed by the J Craig Venter Institute, where he received funding from the NIH on influenza virus sequencing
Ioannis Xenarios	Employed by the Swiss Institute of Bioinformatics, which hosts OpenFluDB

No other interests declared by members of the TWG were deemed relevant to the work of the group.