16 October 2017

II

Implementation of Decision WHA70(10) 8(b)

Evidence for *"Scoping Paper on approaches to seasonal influenza and genetic sequence data under the PIP Framework" ("Scoping Paper")*

Compilation II. GSD Analysis

This document compiles Member State and stakeholder input to the 2016 PIP Framework Review[1] and the PIP Advisory Group work on the handling of GSD under the PIP Framework[2]. Excerpts from PIP Advisory Group documents are found in Compilation III. This document contains the following:

NB: Each transcript or submission found in this compilation is associated to a reference number. These reference numbers are used in Part B of the Scoping paper[3].

= = =

Every effort has been made to be as inclusive, balanced and thorough as possible while recognizing the need to maintain a focus on the topics relevant to the Scoping Paper. Member States and stakeholders are encouraged to contact the Secretariat at pipanalysis@who.int regarding issues or comments on the content of this document which should be considered to be a living document.

[1] Information on the 2016 PIP Framework Review can be found at http://www.who.int/influenza/pip/2016-review/en/.
[2] For information on the PIP Advisory Group's work, see http://www.who.int/influenza/pip/advisory_group/gsd/en/
[3] http://www.who.int/influenza/pip/scopingpaper.pdf

Table II.1 Excerpts from transcripts of interviews with the 2016 PIP Framework Review Group

This table contains excerpts from transcripts of interviews conducted by the 2016 PIP Framework Review Group, with representatives from GISRS laboratories, and other stakeholders. The excerpts focus on two questions asked of each interviewee:

1) How should GSD be treated under the PIP Framework?
2) What are the advantages and disadvantages as well as the opportunities and challenges of including GSD more fully under the Framework?[4]

Excerpts have been anonymized to preserve confidentiality.

Reference No.	Source	Excerpts from Interview Transcripts
G1.	Transcript of GISRS interview for 2016 PIP Review Group	I think the problem is that sequence data deposits in the database is not clearly stratified with our non-PIP materials. If sequenced information is certified as PIP materials, nobody may deposit it into a public open access database to protect the IP of the sequence. That will result in delayed sharing of genetic information and delayed risk assessment based on genetic data. Fortunately semi-open accessible database, GISAID, worked well so far to share data timely and more sequenced data of seasonal and pandemic potential viruses are deposited into GISAID. As I mentioned, at the time of 2013, when outbreaks with H7N9 occurred in China, China Collaborating Centre timely shared sequence data with other CCs through GISAID so we were able to conduct risk assessments by using the sequenced data prior to sharing H7N9 virus from China. So, this is a crucially important responsibility by WHO CCs at the initial process of pandemic, during waiting time of virus sharing. […] Yes, I think this is a whole series of questions we would have on GISAID in general. I don't know if you can comment on your perspective about the use of GISAID and the how and the clarification on the use of GISAID within the framework, but one of the questions was more about biosecurity essentially and people's ability to access it and sequence viruses or genetic materials that may be dangerous to release, for example. So, it's more a biosecurity question. Well, I think there is already an awful lot of data out there on H7, H10, H5 viruses. It's hard to see how you can make the situation worse, I would think. I think if people have got a mind for that, there is already enough information out there. So, there's a lot of work going on within the PIP Framework Advisory Group and working groups on GISAID in terms of how we execute the sharing of sequences and benefit-sharing. Do you have any specific comment on whether the process is addressing the how and what is needed to put in place robust systems enable that sharing of data and benefits? I think GISAID is working well and keeps going on the system. From my point of view, I think it is very difficult to try and protect sequence data in the same way that you may be able to protect virus isolates. How many changes to a sequence makes that virus different from another virus or from another sequence? I think it's a very difficult element to try and draw into the whole PIP programme and to maintain some way of getting a benefit from that and I think the danger of trying to control the sequence data is that people might be... have some reluctance to enter sequence data in a timely fashion. So, I think there's got to be a balancing of the outcomes for any changes to the way things are being done currently. I think GISAID is achieving the aims and countries are... at the moment seem quite happy to put information into to GISAID and I'm not sure whether there is further incentives needed for countries to improve the way they're doing things currently. Following up on that, how do we ensure that... Since there's quite a lot of data out there and anyone can access it, how do we ensure that there is equity in the sharing of benefits from all those that may access and use the information that they are getting from the open access databases? Personally, I'm not sure it's worth the effort. Maybe the other more specific question is if you had GISAID or a database with those characteristics versus the entire public access, no sign-in required databases, based on your current read/experience, which one would you actually go for essentially? I think [name of interviewee] said that GISAID is working and is... has the potential at least to try and protect IP. So, do I hear a preference for that kind of system versus GenBank? I think GISAID does more than try to protect the intellectual property. It also gives some recognition to the laboratories where that data is coming from. And even if they're not generating that data, it gives some credit for the country which is providing those isolates and that's not always so evident within the GenBank. So, I'm not sure it's always down to dollars

[4] See Report of the 2016 PIP Review Group, "Method of work, pp. 28-29, and Appendix II: Detailed Methods of work, pp. 103-106, available at http://apps.who.int/gb/ebwha/pdf_files/EB140/B140_16-en.pdf?ua=1.

Reference No.	Source	Excerpts from Interview Transcripts
		and intellectual property rights and things like that. I think some basic recognition of the work the laboratories are doing who are collecting these samples and possibly sequencing these samples to me is as important as trying to protect intellectual property and raise funds from that area. Yes, I agree with [name of interviewee] 's comment. But then how do you balance that? It's important to acknowledge the work that centres are doing but, on the other hand, we are having countries, actually, who are trying to hold on to information because of intellectual property rights. How do we balance the two? I know as scientists you really want to move forward and acknowledge the work that is being done but how do we balance the political aspect of it to make sure that countries... everyone is happy? Yes, that's a difficult question. I think the whole idea of owning a virus is something we want to try and move away from, if possible. I mentioned at this moment the sequence data deposited in the database is not clearly classified with our non-PIP materials. So, this is a problem. Look, I think companies usually like to do the right thing. So, maybe something like [name of interviewee] is mentioning, that at least it raises the whole PIP issue and PIP programme so that they may embrace it voluntarily rather than trying to use a big stick and going after these companies. At the end of the day there's a lot of information on GenBank, so even if you said to them we're restricting information or restricting access to GISAID, a lot of this information is also available on GenBank without any strings attached at all. Interesting you've referred to GenBank and it has no restrictions. Do you think we need to be working with GenBank to see whether they can also start tracing the... to have some regression of some kind or some traceability mechanism of some kind to see who is accessing the information and what they are doing with the information? I don't think they will agree with that whatsoever. They're a very open public accessible system and they'll never change that and I doubt whether they would ever track anything for you. So, I think that's a wasted effort. I think, again, having something in the line of the virus description to indicate that it is a PIP biological and there is some contact point or something to at least flag it with companies, so if they do want to do the right thing, that there is a mechanism and some easy way of doing that. So, moving things around. Rather than you trying to track down the companies, you have the companies that all come to you, hopefully.
G2.	Transcript of GISRS interview for 2016 PIP Review Group	Well, I think that we are all going to be relying increasingly on genetic sequence data in many cases, I think, in the future to the exclusion of PIP biological material and therefore if the spirit of the PIP Framework is to be maintained and the system of virus and benefit-sharing is to be maintained, then we absolutely must somehow cover the use of genetic sequence data under the PIP Framework. Otherwise I just can't see how there's any alternative but that the Framework will become irrelevant down the road and I think the time is approaching very quickly because we see in many cases that labs are turning increasingly to sequencing and in some cases to the exclusion of virus isolation and antigenic characterization. So, how would the increased use of GSD impact the work of GISRS? I don't think it would impact our work in any significant way. GISRS is moving toward greater genetic sequence data just because it's a sign of the times and that's the way the technology is going. […] Do we need to be mindful of the danger of inappropriate use of data contained in open access GSD databases? I don't know what you're trying to get at there. Inappropriate use of data, is that for nefarious purposes? Is that what you had in mind? I think that there could be inappropriate use of data in open access GSD databases. I think flu is certainly not an ideal pathogen for nefarious purposes but, nevertheless, one is aware that there is potential for it to be used or for people to think that it would be a good medium for whatever it is they're trying to achieve. And other inappropriate use, I'm not sure.
G3.	Transcript of GISRS interview for 2016 PIP Review Group	Thank you. I know you are not... you said this yourself, that you are not doing the sequencing but what do you think, in your opinion, how should GSD fit under the PIP Framework? I think it's a good idea for everybody because one day I think we are also making sequencing, of course, and we'd like to share it and compare it. It's not only the idea of sharing it, I think this is the most important idea, of comparing it and to see some similarities or something strange in each season. So, I think it's always a good idea and it must go further. If we are to think of any disadvantages of including GSD in the PIP Framework, what do you think of? Maybe for people who are making more scientific work there must be some agreements for using some type of or some parts of this data only for practical work, not for scientific papers or

Reference No.	Source	Excerpts from Interview Transcripts
		something. And do you think there is a need to be worried of any inappropriate use of data contained in open access GSD databases? So, those days when we are talking about terrorism maybe you think it's always a danger but my opinion is that biological war has no chance because you are not selling anything. There is no money in this work, so you must have standard precautions but not something special and there is not a big danger of this data-sharing.
G4.	Transcript of GISRS interview for 2016 PIP Review Group	In general this is a very tricky and difficult question. I think one of the things that we shouldn't forget is that at the current time most labs still use both. They use actual materials, viruses, and sequences, so that in reality, for most of the players, this is not really an issue that is that important because most of the manufacturers would be captured through the provision of materials to them. But clearly there is the potential that some manufacturers could make the product only based on sequences and those, I guess, would then have to be covered in another way for the PIP Framework when they used genetic sequence data. Generally I would think that the sharing of sequence data is extremely important and the sharing as quickly as possible for public health purposes. And therefore the baseline view... My baseline view would be that sharing sequences in as open a format as possible, such as in databases like GenBank would be an ideal situation. But clearly there are countries or laboratories that do not want to share in that way and then I think there is a balance between the open sharing as an ideal and the willingness of countries in other ways and depending on where that balance lies, you may want to have a system where the sharing is not quite as open, so for instance with sign-ups like in GISAID and where you agreed to a user agreement and have to log in to download the sequences. If all countries were to share freely on GenBank, then you wouldn't need GISAID but since not all countries do, then GISAID can be a very, very useful database and it is actually currently used in GISRS very widely. I suppose the monitoring and the benefits of the use of genetic sequence data should be proportionate. If you monitor the downloading of sequences alone and not the provision of materials, there will be lots of people who download sequences and nothing much will come out of that. There could be academics, it could even be undergraduate students working on a project. It really would make no sense to get them caught up in SMTA II negotiations or things like that. So, I think you have to look at what can be gained from capturing the use and the download of sequences versus the effort that goes into it. So, I think more tight monitoring and the capturing of use of sequences for the PIP Framework really should focus on relevant products where products are generated and money is being made from those products. The next sub-question, do we need to be mindful of the danger of inappropriate use of data, I was, again, not quite sure what the inappropriate use in that question refers to. Is that inappropriate in the sense of the PIP Framework, so that people would circumvent the PIP Framework or is it inappropriate in terms of biosafety or biosecurity? If it's about biosafety and biosecurity, then while this is a general concern, I think as a PIP Framework we don't need to be very concerned about it because it's a bigger question and also in a way the horse has bolted. There is so much sequence information about highly pathogenic viruses out there already that if you start restricting, say, H5N1 sequences now, you're not hiding too much. In general I would think influenza is not a good bioterrorism agent but people may have different views.
G5.	Transcript of GISRS interview for 2016 PIP Review Group	Yes, that's a very good question and is part of my answer to one of the subsequent questions that yes; the genetic sequence is extremely useful and can be used for initial risk assessment. But as we heard in this meeting, there are certain things that we like to do such as transmission in animal models, and we really can't do that without the virus itself. We can't do antigenic characterization of the virus to know whether we've got a new antigenic or different virus that has emerged and that may not match some stockpiled vaccines for example. So there are some things we simply can't do without the virus.
G6.	Transcript of GISRS interview for 2016 PIP Review Group	How do you think that GSD should be treated under the PIP Framework? It's impossible to track the sharing of genetic information, so the tracking part isn't able to be drawn upon. But I suppose that ultimately, if the data are used to create a product and it can be demonstrated that the data have been used to make a product, then the PIP Framework can apply. And it's, of course, critical that the PIP Framework is able to take into account GSD because that is the future; that will be necessary to make a vaccine or a diagnostic in most cases. Yes, I agree with that. And I think the sequence database that's being developed under GISAID should be supported within the PIP Framework. It seems to be a very sensible database. It protects the rights of the donors, ensures acknowledgement, and it has restricted access. So I think rather than having multiple databases, maybe better use should be made of the GISAID one within the PIP Framework; I don't know if it is already. Okay. Very interesting, because I think GISAID is mentioned in the PIP Framework, but so are other open databases as well. Yes. The problem with having an open database – in an ideal world, that would be the best thing to have, but it doesn't ensure that the donor of the sequence information is suitably

Reference No.	Source	Excerpts from Interview Transcripts
		acknowledged, as far as I can see. Okay, acknowledgement. But do you think there might also be a danger of inappropriate use of data in open access databases? You mean in terms of bio-terrorism? Yes. I don't think that is an issue for the GISAID databases, because H5 sequences are already on open databases, so they could be used inappropriately already. And so we have no control over that. Okay, fine. Do you have any concerns around GSD, and would you say that the work currently being taken by the PIP Advisory Group is addressing the concerns? I don't know all the details of the approach that's currently being taken, I'm not concerned about potential terrorist uses of sequence. I think that sequences are widely available; they get published very quickly through normal open scientific means, and we shouldn't do anything to impede that. The last time I read the WHO's position on the use of GSD, I thought it sounded very reasonable for the position we were in at that time.
G7.	Transcript of GISRS interview for 2016 PIP Review Group	How should gene sequence data be treated under the PIP Framework? I've made comments and written to the review group about this. It does seem to me that with the change in technology now, that gene sequence data is equivalent to sharing the virus or CDNA and so it makes no sense to have gene sequence data exempt from the PIP Framework and so that the review committee should be considering how this would be done. It needs to be a trackable mechanism if we are still to track virus and I would not believe databases that told you that this is not doable because we can see from evidence from GISAID that you can, in fact, register who is accessing data. How you deal with it downstream, how any haul upon that activity is dealt with downstream would be quite complicated but it's very complicated anyway but upstream it's doable, GISAID does it, the others don't. Should we be mindful of inappropriate use of data contained in an open access gene sequence database? I think you mean by this... I presume that it's things like bioterrorist activity. One should be aware of it. I would have to speak to those involved in proper risk assessments in terms of counterterrorism to see how that fell along their risk register and I would therefore have to... I can't really assess that and I'm not sure how many people really are in a position to assess that in a real... in a significant way and I think a lot of would be speculation rather than seeing what... One should be in an environment in which one talks to counterterrorism about these kinds of activities. Of course, we at [GISRS laboratory], because we have highly pathogenic avian influenza viruses, we have to make our risk assessments, unlike most other, I think, centres within GISRS with risks of bioterrorism in place. We would not get permission to work on genetically modified organisms or highly pathogenic avian influenza viruses if we did not bear in mind the counterterrorism activities and therefore we need to think about that.
G8.	Transcript of GISRS interview for 2016 PIP Review Group	I agree that genetic sequence is very important to be involved in PIP framework. It's now one of the most important analyses for the gene of influenza virus to detect early mutation or, especial in H5N1, in which they expect in [country name] that it may transfer from person to person and [inaudible] about this point so genetic sequencing is very important for each case, especially from [country name].
G9.	Transcript of GISRS interview for 2016 PIP Review Group	Sharing of viruses and genetic sequences are all part of the same continuum because, after all, nowadays I mean you can synthesize a virus from the genetic sequence. So in one sense, by sharing the genetic sequence you are essentially potentially giving away the virus in theory. So I think, you know, so if virus sharing is under the PIP Framework then I, logically I would expect, well logically I cannot see why the genetic sequence information is not regarded in the same way. But however, I mean again, bringing it under MTAs and all this would again further complicate matters. I think at the moment there is more free sharing of genetic information than there is sharing of viruses so I mean logically, as I said, it's difficult to make a distinction between the two. But if you bring sequence sharing, you know, under additional, what you want to call it, MTAs of whatever issues, I think it probably is going to slow things down, it's going to, well slow down the information sharing. So I don't know how to answer that particular question. But certainly I would like to leave it as it is at the moment but if people argue that it should be under the PIP Framework I would find it difficult to logically disagree.
G10.	Transcript of GISRS interview for 2016 PIP Review Group	Okay, so my point is that we should include GSD into the PIP framework as the PIP material. The reasons including are, you know, currently the technology developed - we can synthesize the pathogens only based on the sequences so this technology - you know, you say [unclear] so the advantage for us is that we can produce vaccines, can produce diagnostic kits only based on the sequences, so that's very good. But if we do not include the GSD in the PIP framework, a disadvantage is that, so for some laboratories, some countries, they will worry about, you know, the sequence and they do not

Reference No.	Source	Excerpts from Interview Transcripts
		want to share the sequence so that's because of - as we all know, that the companies can produce vaccines, diagnostic only based on the sequence. So my suggestion is that we should include the GSD into the PIP framework and to set up mechanism to encourage all parties to share the sequence timely and not only benefit, I mean, for the manufacturers but also for the global health system, including WHO, to do the risk assessment. So I think it's very, very important, we should include the GSD to be treated under the PIP framework.
G11.	Transcript of GISRS interview for 2016 PIP Review Group	I think definitely it should be included and also, you know, sometimes if - I think we are probably, some of the [unclear] don't do it as routinely as we should do it but we're trying to get there. Because, you know, some actually might pick up something that's inside their data that somebody might have ignored and you can follow up on it. Because sometimes people, more genetic data generated, people don't always have the resources to really spend time looking at it and, you know, it's a resource that WHO can use as part of pandemic preparedness; I would definitely say yes. We left something out as well so CDC's input service, for instance, it also states on there that when you share samples - we had an experience - that you must certify that the specimen, if it's of animal origin, is Newcastle disease-free for instance. And I know for instance last year I was in a situation where we wanted to share some isolates and actually our veterinary department declined sharing of that isolates because they say they can't guarantee or certify that it is Newcastle disease-free for instance so... And I think in those instances it might be easier to share genetic material. Okay, thank you for bringing that up. Whether they are regulatory issues [overtalking] included in the framework it would enhance virus sharing because we could send [?] the sequence data. That's a very good suggestion. This is the first time I've heard such a brilliant idea.
G12.	Transcript of GISRS interview for 2016 PIP Review Group	It makes sense to do the GSD and from my perspective I guess I'm not convinced that the amount of effort that would have to go into this to do it properly is actually commensurate with the central benefits that would come out of it I think. Perhaps I'm wrong but in my own mind I'm thinking the people, meaning the potential contributors you would capture that are using just GSD that you don't capture through essentially the biological materials that we have now, I'm not sure that's really a huge amount. So I'd be very interested in seeing a little bit more of the justification for this effort of tracking GSD. There are going to be some issues I think with the GSD in terms of even identifying if someone actually used GSD and if it's changed, and where it derived from, did it actually start from PIP biological material or not? I think particularly with the fact that, as I said, PIP biological material is fairly narrowly defined right now as those viruses coming from humans, we have almost identical viruses in the animal reservoir. So I think that's going to be very, very challenging to do that; certainly at the level of product, searching for product and use of genetic sequence data and products that may be the challenge. Identifying what is PIP GSD at the other end in terms of download from databases; we think that's going to have some issue. Obviously there are databases that are a little bit more set up for that right now to do it, [unclear] being the obvious but they're not the only players that could do it. So again to me the whole GSD thing is, I'm not sure. I think the effort to go, and probably the cost to go and to track and trace it, I'm not sure. And if you look at what potential benefits will derive from it, I'm not sure that really makes sense.
G13.	Transcript of stakeholder interview for 2016 PIP Review Group	…everybody had a reason why they didn't want to share; some were out of pure economic perspective. So ownership, intellectual property came up, yes, but others were also fearing there are other economic repercussions: they couldn't be selling their chickens or poultry, or, as I remember in one particular case, not so far away from Europe, they were concerned about tourism being affected if they admit and they share this and they see that there is an outbreak, and at the time that I don't believe there were human cases. At that point, it was very clear, in order to do this, when you say, how do you address the issue of a peer review journal, if I give you this example - so there is a scientific etiquette that scientists, I guess, learn in university that says you must properly acknowledge, do not plagiarise One of the downsides, I think, from the traditional archives, call them public domain, call them whatever you want, the reality is these archives have a much broader role. They are not in that discipline specifically to influenza; they were not set up for that. Most of them are not even what you call relational databases. They store and archive genetic sequence data. Some of you have a background in science and you certainly know that just having a genetic sequence data is sometimes not enough to really understand the problem; you need the epi-data, you need the clinical data, also known as the metadata, to have a better picture. Even more important, we thought you need to be enabled to connect with the source of this information, the submitter, because the submitter may have much more information, and also to start collaboration. […] So this is all about the mechanism per se, not so much a data repository just where you choose what is the nicest repository; it's the mechanism that's behind it. There's transparency and the confidence… […]

Reference No.	Source	Excerpts from Interview Transcripts
		We found giving your identity a very low hurdle. […] Some of these things were shared, of course, with many of those following this process, and I think it's very, very simple: keep your framework together; do not become overambitious, is the advice. Stay true to influenza, of what it is as a special pathogen. Don't let PIP become a regulator and telling countries how to share and how to deposit.
G14.	Transcript of stakeholder interview for 2016 PIP Review Group	Well, so I think that actually having the process that we, realising the process was necessary that we've just been through was very good and I think the big technological change that we're responding to there is that the, or the big driver for sequence data, genetic sequence data becoming more important is that one can do, technically one can do much more with sequence data alone than one could a few years ago and the need for, you know, physical samples is diminishing in many areas.
G15.	Transcript of stakeholder interview for 2016 PIP Review Group	The WHO PIP framework clearly envisions the influenza virus genetic sequence data should be either in the public domain or publicly accessible to the benefit of all and that is actually outlined in the framework. As this is currently implemented, the framework does allow for near instantaneous dissemination of the genetic sequence data from both public domain, for example GenBank, or publicly accessible, for example GISAID databases, within the scientific community and industry and we believe that this is critical for continued influenza and non-influenza R&D efforts. We recognise that some of the databases already have well-established traceability mechanisms in place as they require users to register prior to being given access to the information on the database, and there's also been a suggestion that all the databases have some sort of pop-up window so an individual on entering the database understands that there may be some PIP framework obligations linked to them accessing the data. We believe that the impact of using the pop-up windows and using these traceability systems to identify users of GSD for influenza and attaching PIP framework obligations to the general use and sharing of this publicly accessible GSD should be fully considered as this could seriously discourage individuals and research groups from accessing the data and therefore inhibit R&D on influenza. [Name of stakeholder] support the benefit sharing for commercial pandemic influenza vaccines developed and manufactured using GSD. However, we would like to see that basic R&D is excluded from this and that we use other ways of identifying products developed using GSD, such as a trigger point in the development process for influenza vaccines. If we could identify such a trigger point after which PIP framework obligations are activated, we feel that that would be a preferred approach. We also feel that the anticipated benefits of including influenza GSD in the framework should be rigorously evaluated as we can see that there could be much time and effort applied to this area with limited benefits as many of the companies using influenza GSD already contribute to the PIP framework. [Name of stakeholder] believe that GSD should continue to be shared in public domain or public access databases for the public benefit and in the interest of public health as originally envisioned by the PIP framework. The framework was drafted and adopted without any mechanism in place for placing such special restrictions or obligations to contribute to the framework on the access of such data and we feel that placing such restrictions or obligations on the access could delay the dissemination of this important information, both within industry and within the scientific community, and potentially erect barriers to the public and private researchers alike. Most importantly, we feel that this could hinder rather than bolster progress towards pandemic preparedness. Thank you. […] So I think our concern is that people may be accessing the GSD data for basic R&D in influenza and even research in other areas, and we would not want to inhibit that in any way. We recognise that if companies are using solely GSD to develop products that there should be some benefit sharing if it goes on to become a product, but I guess our request would be that the stage at which you consider the use of GSD as being able to generate that benefit, so I think it's more about the stage in the process where the GSD would be considered to have been used for a benefit. So maybe more attention could be given to that so that we don't inhibit this basic R&D on influenza.
G16.	Transcript of stakeholder interview for 2016 PIP Review Group	I think one huge challenge or threat is not dealing properly with the issue of genetic sequence data. So, there is some discussion whether genetic sequence data is in the Framework or it's not in the Framework and we feel that actually that's not a very good point to raise in the sense that once the information is there, because we want to synthesise biological material from the information, that biological material itself falls within the context of the PIP Framework. So, whether sequence information is mentioned specifically/not mentioned specifically for us is a secondary issue because the main thing is that once there is material synthesised from the

Reference No.	Source	Excerpts from Interview Transcripts
		sequence information, in any case that is covered by the PIP Framework. So, for us whether sequence is mentioned or not is rather a moot point and it's quite clear that because biological materials are covered, it would be rather ridiculous to say that sequence information somehow are not covered and the biological materials are derived from the sequence information. So, I think one huge concern we have is how the GSD is being dealt with and we are concerned that... Because as we see, we have given a number of options on how to deal with GSD. The two main options was for WHO to have its own database and to implement user access agreements so you can identify who is accessing the data. At the same time we know that data can be made available to others to use as well. So, one was for WHO to hold a database. Second was to make it clear that only databases that have user agreements, such as those being implemented... And we know it's not difficult to implement databases because GISAID is doing... Sorry, it's not difficult to implement user agreements because GISAID is implementing such user agreements and it does not hinder the sharing of data. So, following from what GISAID is doing, we feel that all databases that want to host GSD should actually implement the user agreements. So, we think that these are the two options available and not dealing with GSD properly will definitely undermine the PIP Framework and especially the component on benefit-sharing. So, we are no longer treating virus-sharing and benefit-sharing on an equal footing. […] Well, I would double down and doubly emphasise the question of sequence data. I think that the Framework absolutely has to get this right and there's cause for concern that it's not going well so far. You can look at the papers that have been produced and they will clearly indicate in some cases, for example, what the best option is for implementing the framework but then you will see other options that aren't as good that continue to be listed. So, if we don't get the sequence data issue right, I think there's a very serious long-term threat to the Framework and I think maybe later on a larger-scale solution emerges elsewhere but for right now it seems like the data access and user agreements are critical and I think that it's just really, really important that the Framework get this right or else you're going to see synthesis and digital transfer of materials, really undermine the balance in the Framework.
G17.	Transcript of stakeholder interview for 2016 PIP Review Group	Well, so there are two things. The first one is clearly the PIP framework should actually get what I call the techies or the geeks in the discussion, because the tendency to think that they're going to be only certain sequencing centres or big data warehouses that will be the sole and single place where things will happen is probably something that will diminish over time, mainly because sequencing can be done close to the sample and close to the place where it's sampled. So I'm thinking of embedded systems like the one from Oxford Nanopore which are a sequencing device that looks like a USB key that allows you to sequence and type viruses and bacteria while you are close with a very limited amount of material or experimental material, but also very limited lab facilities. We even did that in a hotel room and did the sequencing of the bacterium that were on the toilet of our hotel. Fundamentally you can do that even in a hotel. So that means I think the PIP framework should really look into some of the disruptive technologies that might actually get the sequencing out of the traditional flow which is a country's sequencings, they send that to the coordination centre, whatever country, and if they are in the US that's the CDC [unclear] send all these viruses down to the WHO CCs and they have the vaccine decision, etc. I think it's going to be strongly distributed in the long run. That means people will be contributing directly to the sequences at a faster rate, and the technology that is changing is really getting… last week I was in London for a meeting where basically these sequencers are attached to our iPhone's so you can actually sequence with your iPhone your DNA or genomic sample. That is completely disruptive if you think about places in Southeast Asia, or Africa for that matter, where for the time being it's extremely hard to have access to the sample and also have access to the sample in a good condition. So I think that's the type of thing you might want to keep an eye on and this is actually the way to get these in there, is by getting this conference where you take some of this cutting-edge or sometimes bleeding edge I would say technology. I think those are the type of things which need to be changed. The other one is the people that are from the biohacking part, which are the people that design viruses from a sequence, from DNA and use a synthetic virus. I think this one is by definition something that exists already now and the digital form of it is one form that in any case will leak out and be used for creating viruses. Now, I know that homeland security would not like that because it's a threat. I still believe in mankind and not in the crazy guy. We have to protect the 99.9% from the 0.1% of crazy people. So I think this is the other part which, I guess getting a bit of an understanding of how the PIP framework should adapt to this, this is actually one of the reasons as well why the synthesis of viruses was seen as, we need to protect the DNA sequences and have a block over it. My concern as a virologist is that in any case nature tries a lot of viruses and in any case we'll cover much more of the space of, what are the viruses circulating in the world, by just distributing and sharing that information rather than blocking it behind firewalls, [unclear] wall or data access agreements that nobody understands the ins and outs. But that's my own way of seeing the future of mankind, I guess. You've talked about protecting the GSD. Do you think it's a threat, the breach of security in open access? No, I don't think it's a threat for tourism. The first one, producing a vaccine still takes a lot of resource. Not everybody can be a vaccine producer. Second, at the time people describe what the vaccine is they have to describe full-fledged, and that's an EMA and FDA requirement, what is in their composition, which means you'll get access to the virus sequence that originated from this vaccine. So rather than hitting the people that want to use the data before, the people that are producing the vaccine, at the end of the day they'll have in any case to show their books and show which of the viruses has been used. So now if it's Thailand, or a Vietnamese, or an Indonesian, for that matter, virus that has been used in the composition, then this will

Reference No.	Source	Excerpts from Interview Transcripts
		actually be identified easily at the time the vaccine is produced. That's where actually the PIP framework should get in place because most of the vaccine composition is taking one of the Indonesian ones, hence 20% of the cost, of the benefits go back to Indonesia because they have been… so I think this whole, and that was a bit the issue that we discussed earlier in the TWG, in the advisory group, was that it's trying to set up a gas industry process for something which at the end of the day legally people that describe and produce a vaccine have to demonstrate what they have in there, and that's where actually most of the public health could actually also assess this when they get the vaccine produced. So that means you can have even third parties that are public health that investigate the vaccine composition and in that manner you control the whole process, and anything that comes before, because patenting sequences, and we have that from the American ruling, is not something that is of value; it's nature that does that. Even the virus, I could claim in front of any advisor [?] the virus itself is not owned by any country because it's a natural process of selection, a virus in a given individual, a bird, a pig or a human, so even that is not owned. Merrimack Pharmaceutical that tries to patent the human genome actually got in big trouble with that because the laws in the US have been overturned, and basically this is actually what nature produced. Viruses, yes, they are a bit of a pain – that's what nature produced. It's only when you use these viruses for doing a vaccine, which at the time it's the producer that will know that, it's only at that time that you need to identify who's who and actually then get the cash back to each and every country that contributes to this virus. The advantage of the sequence, if I can add, say the virus doesn't know the boundaries of nationalities so you have that same virus that is circulating in Thailand, Vietnam, Laos – for example, an H5N1, or a Chinese one, say. If you are a manufacturer and you want to put this one in, then basically you can actually then split the cost by the number of sequences that are exactly identical to the ones that are in the vaccine and you do a clé de répartition amongst the different countries, depending on the number on vials of vaccine they are using. That's easy; everybody could do that math. So I think that's a bit always the issue I see with this trying to hunt the access, [sound slip] accessing what rather than going after the… when the vaccine is produced, that's where the value is. Anything before [sound slip] have a value. It's just [sound slip] that is actually what public health means, which means identifying in real-time or as close to real-time the propagation of a virus and the circulation of a virus in the human population, but if that…
G18.	Transcript of stakeholder interview for 2016 PIP Review Group	But to me I think the biggest issue with the recommendations that the group came up with is the sharing of sequence data between databases and so a key aspect of the whole PIP Framework and how genetic sequence data is managed is that there's... the default is that people would be submitting the sequence data to a controlled access database and currently that is the EpiFlu database that's being maintained by GISAID. One of the problems with the current approach is that other databases, like GenBank and the Influenza Research Database don't have access to the sequence data that are being deposited into GISAID and I think that has some unintended consequences because, for example, in the Influenza Research Database the primary sequence data are taken and then additional data is derived by analysing that and I think a lot of that is very useful for pandemic preparedness. So, for example, a whole series of genetic markers for important virus characteristics like host range restriction and transmissibility, anti-viral drug resistance, markers of increased disease severity and virulence have been identified and all the sequences that come into the Influenza Research Database are routinely screened for the presence of these markers of important viral characteristics and then we provide those as features within our comparative genomics analysis tools to see, if you're looking at a particular lineage of viruses, which of these characteristics are found within that lineage. Well, we can't provide any of that for the sequences that are directly deposited in GISAID because we're not able to access those and pull them out and add that kind of value-enhanced data to those sequence records.

Table II.2 Excerpts from Member State and stakeholder written submissions to the 2016 PIP Framework Review

This table contains excerpts from Member State and stakeholder written submissions to the 2016 PIP Framework Review.[5] All submissions were reviewed and content relevant to GSD was excerpted. In some cases, the excerpts have been anonymized to maintain confidentiality.

Reference No.	Source	Excerpts from written submissions
G19.	Australia submission for the 2016 PIP Review	The implications of the availability and use of genetic sequencing data is a growing concern. Handling genetic sequencing data in the context of the PIP Framework has the potential to undermine the Framework's operation through circumventing benefit sharing. This review provides a good opportunity to ensure genetic sequencing data can be sufficiently managed under the Framework. Australia welcomes further consideration of arrangements for sharing of gene sequence data, as well as managing sharing of associated benefits into the future.
G20.	Australia submissions for the 2016 PIP Review	It is crucial that the PIP Framework adapts to evolving aspects of technology – particularly genetic sequencing . The implications of the availability and use of genetic sequence data is a growing concern and should be handled with caution. Genetic sequence data should be sufficiently managed under the PIP Framework as not circumvent benefit sharing. Consideration should be given to arrangements for sharing of sequence data and associated benefits in the future. Australia notes the IHR Review Committee recommendation on sharing of genetic sequence data, and is grateful for the work being undertaken by the PIP Advisory Group – which offers a way forward on this issue in the context of the PIP Framework.
G21.	Australia submissions for the 2016 PIP Review	Australia supports further clarity around the handling of genetic sequencing data (GSD) to ensure consistency with the principles of sharing other PIP materials, and considers benefit sharing arrangements should be addressed as a matter of priority. We note the Review Group's finding that monitoring all access to GSD is unlikely to be feasible due to the many public and private ways of sharing and accessing GSD and associated workload implications, however monitoring GSD use in commercial end products and tracking commercial products would be achievable. It's acknowledged that WHO's leverage in this space will centre on the unwillingness for companies to buy products unless WHO can guarantee their structure and composition. Australia shares concerns that changing the definition of PIP biological materials would require significant amendments and likely complex and timely negotiations. Further detail on the viability of developing an Annex for Article 6 to include GSD is warranted, while this approach appears reasonable in principle, it's acknowledged it may be very difficult to cover all possible sequence variations that could be produced. It's acknowledged that the storage and sharing of digital information, including the uploading online of genetic sequences, is being considered in a number of international contexts, including for example, synthetic biology discussions through the Convention on Biological Diversity. Any approach should be informed by, and align with, developments in other international fora.
G22.	Germany submission for the 2016 PIP Review	We agree, for this, the handling and sharing of genetic sequence data (GSD) for influenza viruses with pandemic potential and benefiting from doing so are crucial issue for the future. The technological development, allowing GSD being used to manufacture vaccines and other influenza-related products, underlines the need to address regulatory and intellectual property issues, including benefit sharing, but also monitoring and tracing methods, as well as biosecurity and biosafety issues. The Technical Working Group (TWG) on the "sharing of influenza genetic sequence data under the PIP Framework" has provided a well appreciated draft paper on "Optimal Characteristics of an influenza genetic sequence data sharing system under the PIP Framework". Germany considers the GISAID EpiFluTM Database addressing most if not all characteristics for the optimal sharing of GSDs as described by the Technical Working Group, in particular allowing the sharing of genetic sequence data in the spirit of PIP Framework, ensuring fair, timely and transparent access to GSD in a sustainable way. GISAID has also the potential through its Data Access Agreement to monitor the use of GSDs, and therefore to ensure that benefit sharing obligations are met by the users. To foster GSD sharing, Germany will therefore continue its support for the PIP-Framework and in particular continue to host GISAID. Germany submitted comments to the TWG paper. In our comments we have applied a "one health" approach, compiling input of our experts form the human and animal health sector. In our comments we have also raised questions on missing justifications for some options and on the rationale why other potential options were not included in the paper. We are convinced and are grateful that the PIP Review Committee will consider Germany's comments on the paper, as well as the Review Committee will do with the comments received by other member states and entities.
G23.	Germany submission for the 2016 PIP Review	Germany welcomes the work of the PIP-Secretariat and the initiated review process. We agree, that the handling of genetic sequence data for influenza viruses with pandemic potential and benefiting from doing so are crucial issues. The technological development, allowing genetic sequence data being used to manufacture vaccines and other influenza-related products, underlines the need to address regulatory and intellectual property issues, including benefit sharing, but also monitoring and tracing methods, and biosecurity and biosafety issues. Germany will continue its support for the Pip-Framework. In particular we will continue to host GISAID as an important and forward looking tool for sharing of genetic sequence data in the

[5] See Report of the 2016 PIP Review Group, "Method of work, pp. 28-29, and Appendix II: Detailed Methods of work, pp. 103-106, available at http://apps.who.int/gb/ebwha/pdf_files/EB140/B140_16-en.pdf?ua=1.

Reference No.	Source	Excerpts from written submissions
		Spirit of PIP Framework, ensuring fair, timely and transparent access to GSD in a sustainable way."
G24.	Germany submission for the 2016 PIP Review	Finding 3 In our point of view publication (at least) IVPP data should be performed preferably in databases having a mechanism to trace the source laboratory as well as the further usage of the GSD (f. e. GISAID offers these possibilities). This procedure should be adequate to the traceability sourcing from end-products (which is not established yet). The idea of handling GSD in a special section seems to be favorably. The complex of traceability should be clearly straightened in this section.
G25.	Germany submission for the 2016 PIP Review	Moreover, genetic sequence data (GSD) are of great importance with regard to pandemic preparedness and has to be included in the PIP sharing and benefit systems. Most GISRS laboratories are providing GSD via GISAID, a database that addresses the principles underpinning the PIP framework. Thus, fair and transparent sharing of GSD is basically ensured and contributors are provided with the necessary incentive to rapidly share GSD.
G26.	Member State submission for the 2016 PIP Review	The lessons learned for the sharing of other pathogens however must be viewed in the context of current highly dynamic developments, among which we underpin, inter alia: • The necessity of fast, real time sharing of pathogens, • The inevitability of sharing materials and data in the public domain, for the sake of transparency and inclusion of 3rd world systems and scientists, • The fast evolving transformation from the sharing of samples to the global sharing of genomic (WGS) data out of them, • The great chances and benefits for public health and science, arising from WGS-technics, on combined, identical repositories for all pathogens from all domains instead of single pathogen, single domain systems. • The ongoing transformation from single domain systems to a One Health approach • New barriers to collaboration: the public health necessity of *multilateral* agreements as a *conditio sine qua non* for sharing of pathogens and genomic data, while for many countries at the same time there is now an obligation to comply to the CBD Nagoya Protocol, based on a diversity of bilateral agreements between Parties.
G27.	United States submission for the 2016 PIP Review	The PIP Framework should continue to facilitate the rapid and timely collection, curation and sharing of genetic sequence data (GSD) of influenza viruses with human pandemic potential (IVPP). Open, rapid sharing of and access to GSD from currently circulating influenza strains to the WHO's Global Influenza Surveillance and Response System (GISRS) and the broader research community engenders critical benefits for the international community, including: the ability of WHO to monitor the emergence and evolution of variant seasonal viruses and make timely decisions and recommendations on the composition of influenza vaccines, antiviral drug use and molecular diagnostic methods; collaboration opportunities and sharing of research and development results, specifically in the development of novel predictive modeling algorithms that can facilitate WHO vaccine virus selection; improved risk assessment in candidate virus selection to avoid mismatch; integration of these data sets with other publicly accessible data sets; and a publicly accessible archive for the valuable data generated by GISRS. Benefits of sharing IVPP GSD include use for pandemic risk assessment and pandemic preparedness efforts including the development of sensitive virus detection methods and the generation of candidate vaccine viruses. Additionally, open and timely access to GSD from currently circulating influenza strains provides information directly used to produce licensed recombinant influenza vaccine and could potentially also allow egg- and cell-based vaccine manufacturers to use synthetic genomics to construct candidate vaccine viruses. Ensuring that the global system has the best technologies at its disposal to analyze viruses and contribute to the production of influenza vaccine remains critical. This includes the data management systems that the GISRS relies upon. The United States encourages the Review Group to explore various avenues of ensuring that the databases upon which GISRS relies are sustainable and resilient, including but not limited to developing standard operating procedures for such databases or potential terms of reference for GISRS-affiliated databases. The United States encourages the Review Group to consider the use of partnership contribution resources for ensuring the sustainability and resilience of such databases, which is an important component of improving global pandemic influenza preparedness and response. A successful data repository supporting the efforts of the PIP Framework and GISRS requires stable funding and structure and a forward thinking strategy to meet the future needs of GISRS beyond current needs in terms of increases in data volume and complexity and advanced computational tools. To meet GISRS needs for rapid and transparent sharing of GSD, databases would need to support real-time availability of GSD with a comprehensive set of associated and searchable epidemiologic and virologic data. It is also important to make sure that GSD from these data management systems are sufficiently curated and archived so that important historical data is not lost.
G28.	United States submission for the 2016 PIP Review	The United States encourages the Secretariat to organize continued thoughtful discussion of how genetic sequence data (GSD) should be handled under the Framework. We note that future discussions would benefit from robust consideration of the meaningful differences between GSD and biological material. These discussions would also benefit from a more complete understanding of the benefits engendered by the sharing of GSD. Access to GSD may call for different thinking about what type of benefits should be triggered, and a simple mapping of the current benefits tied to biological material may not be appropriate. We encourage Member States to focus on other benefits such as attribution, and inclusion of scientists from GSD-originating countries in research, among others.

Reference No.	Source	Excerpts from written submissions
G29.	Stakeholder submission for the 2016 PIP Review	• Technological advances in vaccine development make it possible to start with the published sequence only to develop the vaccine without having to transfer the physical materials. ▪ While these advances fulfils the need for accelerating vaccine development in possible pandemic scenario, but render the physical material sharing irrelevant. o **Sharing and use of GSD should be discussed more comprehensively in view of these advancements and Nagoya protocol provisions.** ▪ **Link the genetic sequence database to IVTM to monitor and track the utilization at any stage of the value-chain: research, development, innovation, pre-commercialization or commercialization.**
G30.	Stakeholder submission for the 2016 PIP Review	We support benefit sharing for commercial pandemic influenza vaccines developed/manufactured using GSD. GSD of Influenza Viruses with Pandemic Potential (IVPP) are not WHO PIP biological materials (PIPBM), per the definition of PIPBM, and the WHO PIP Framework clearly envisions that GSD should be in the public domain for the benefit of all (see 5.2.2.) i. As currently implemented, the WHO PIP Framework allows for near-instantaneous dissemination of publically available GSD within the scientific community and industry, which is critical for continued influenza and non-influenza R&D efforts. As part of the review, industry is willing to consider an appropriate revision to the PIPBM definition to reflect anticipated technological advances for the specific purpose of the WHO PIP Framework. i. Not all influenza IVPP and IVPP GSD, however, should be included in the definition and subject to WHO PIP Framework obligations, rather only GSD used to directly develop/manufacture commercial IVPP products. ii. Attaching obligations to general use and sharing of publically available GSD would inhibit influenza R&D and is therefore strongly discouraged. iii. In some cases, physical samples and/or GSD are used for R&D of non-influenza products and should ***not*** be subject to the WHO PIP Framework. [Name of stakeholder] recognize that some databases (e.g. GISAID) already have mechanisms in place to monitor and trace the general use of GSD provided by those databases, such as database use agreements and user registrations. However we strongly oppose the use of this information to identify individuals/entities as potential contributors to the PIP Framework either through the Partnership Contribution or the SMTA2s. [Name of stakeholder] believe that GSD should continue to be shared in open databases for the public benefit and in the interest of public health, as originally envisioned by the PIP Framework. The Framework was drafted and adopted without any mechanism for placing special restrictions and conditions on the access to such data. In fact, placing restrictions on access and use of GSD would delay the dissemination of this important information within industry and the scientific community and erect systematic barriers to access by public and private researchers alike. Most importantly, it would hinder rather than bolster progress toward pandemic preparedness.

Advisory Group work on handling GSD under the Framework

As part of its work on the handling of GSD under the Framework, the Advisory Group established two expert groups and held several consultations.[6] The table below contains submissions provided by Member States and stakeholders to the first expert group – the Technical Expert Working Group (TEWG).

Table II.3 Technical Expert Working Group (TEWG) on genetic sequence data - Excerpts from stakeholder submissions

The TEWG was asked to provide an assessment of the scientific, technical, operational and intellectual property implications of using IVPP GSD data rather than physical materials for influenza research and vaccine production, including how the transfer of such data could be monitored.[7]

Stakeholders were asked to provide their views on the following questions[8]:

- What are the current known uses of GSD in relation to influenza related technologies, products, inventions and patents?
- What additional uses are anticipated in the future?
- What are the prospects for any of these uses to result in commercial products?
- What are the current issues (including regulatory pathways, IP issues, etc.) that must be resolved before commercialization or the licensing of products?
- Considering methods to generate, store, retrieve and share GSD, is it possible or feasible to monitor or trace the sharing of GSD? What are the relevant technical, legal or other issues?
- What, in general, are the potential biosecurity issues related to the use of GSD?
- What are the prospects for influenza GSD to be used to develop non-influenza related products?

This table contains submissions from stakeholders to the Technical Expert Working Group on genetic sequence data (TEWG). All submissions were reviewed and content relevant to Part B of the Scoping Paper was excerpted. Excerpts have been anonymized to preserve confidentiality.

Reference No.	Source	Excerpts from written submission to Advisory Group TEWG
G31.	Stakeholder submission	Some proposed edits to the TEWG Preliminary Report nevertheless propose that (a) bilateral material transfer agreements be required for the use of influenza GSD originating from the GISRS, (b) that restrictions be imposed on the sequencing and use of GSD obtained from PIP biological material, (c) that native influenza gene sequences (regardless of evidence of the origin of the sequences) be considered unique identifying data that would tag any products containing those sequences as PIP biological materials, and (d) that mechanisms be put in place to monitor the use of influenza GSD. Such measures are impractical, lack a sound legal basis, diametrically oppose the findings and conclusions of the working group, and would harm the public by severely hampering the wide range of activities using influenza GSD that have been identified by the Working Group as serving the interests of public health. In pandemic and outbreak response, speed is essential to prevent widespread loss of life. Creating an incentive to delay the start of pre-pandemic, pandemic, or seasonal synthetic vaccine virus generation until relevant influenza GSD is widely available from sources other than the GISRS would slow the development of essential vaccines. Some of the proposed measures threaten to undermine the GISRS itself. As routine clinical diagnosis of respiratory infections shifts from the current antigenic tests, which must be particular to each pathogen, to more general sequence-based diagnostics that do not require pathogen isolation, the generation of influenza GSD is likely to shift increasingly from the GISRS to a more dispersed and heterogeneous set of laboratories. Creating disincentives to the use of influenza GSD from the GISRS but not from other sources is likely to diminish the role of the GISRS and the level of support that it currently enjoys. The PIP framework's goal of increasing global access to the benefits of influenza immunization is essential to promote the health of much of the world's population. Some of the proposed measures to designate native influenza GSD as PIP biological materials, although intended to further this important goal, are likely to be counterproductive and harm public health.
G32.	Stakeholder submission	[…] [name of stakeholder] strongly supports the Member States' appreciation of the value of open access to the data. […] As you are probably aware, [name of stakeholder] has a long and productive history of supporting open access to the results of the research it funds. That approach, I believe, recognizes the fact that science is cumulative and that the greatest benefit to public health will be achieved if scientists can rapidly and easily access the research that has come before them. As [stakeholder]'s Data Sharing Policy states, "data sharing is essential for expedited translation of research results into knowledge, products, and procedures to improve human health." [name of stakeholder] has repeatedly translated this policy position into concrete objectives, most recently with the [stakeholder] Public Access Policy, which requires all research papers resulting from [stakeholder] funding to be made publicly available as soon as possible after publication and no longer than 12 months after. In conclusion, we believe that the GenBank model of openness and accessibility provides the best model for advancing science. Scientists the world over have a working assumption that as a community resource, new sequence data belongs in GenBank because having a common public archive is the most effective and efficient way for science to progress. Likewise, to address global pandemics and other

[6] See GSD Timeline for Process for handling Genetic Sequence Data under the PIP Framework, http://www.who.int/influenza/pip/advisory_group/GSD_timeline.pdf?ua=1.
[7] See Terms of Reference, Technical Expert Working Group on genetic sequences under the PIP Framework, available at http://www.who.int/influenza/pip/advisory_group/TEWG_ToRs.pdf?ua=1.
[8] See Technical Expert Working Group on genetic sequences under the PIP Framework, Report to the Advisory Group, Annex 3: Methodology for the development of the Report, October 2014, available at http://www.who.int/influenza/pip/advisory_group/PIP_AG_Rev_Final_TEWG_Report_10_Oct_2014.pdf?ua=1

		critical incidents it is vital to eliminate barriers to data access in order to engage the broadest community of scientists to bring their collective expertise to the incident at hand. I believe that the PIP Framework would most effectively serve the community of influenza researchers, the community of genomic scientists at large, and most importantly, global public health needs, by endorsing free, unencumbered deposition and access to influenza virus data.
G33.	Stakeholder submission	(i) The definition of PIP biological material (PIPBM) in Section 4.1 is non-exhaustive as it uses the word "includes". The definition of PIPBM was a controversial issue during the PIP negotiations, and thus there is a requirement on the Director-General to inform the World Health Assembly on a biennial basis "the experience arising from the use of the definition of PIP biological material in Section 4.1"2. It is our contention that the definition of PIPBM includes genetic sequence data since any distinction between the physical material and the information contained therein would be an artificial distinction. As argued above material accessed directly from GISRS and GSD (and materials generated there from) should be considered as equivalent under the PIP Framework. This is regardless of the source of the GSD, provided that the same sequence is found in GISRS strain(s). Some may argue that GSD-generated vaccines would need to be characterized, and because that process might require direct use of GISRS materials, that it might then be assumed that the Framework's benefit sharing provisions would still apply. This argument takes an overly simplified view of the set of entities potentially involved in the research and development process. It would be possible for a company developing a GSD-generated vaccine to use contractors or other collaborators to perform characterization (or, for that matter, other steps in the process potentially involving physical access to GISRS materials), and thereby not engage in benefit sharing.5 Accordingly, the WHO should consider all entities asserting intellectual property that claim GSD and/or GSD-generated materials to be subject to the PIP Framework's full complement of benefit-sharing requirements, including both the partnership contribution and the benefit sharing obligations incurred in the SMTA. This should include entities asserting IP who then license rights to others. The same holds true for all products developed utilizing GSD, even if intellectual property is not directly claimed over GSD and/or GSD-generated materials. Specifically on GSD, what this means is that all users of GISRS that benefit from the use of GSD (which falls within the definition of "using GISRS") should make a partnership contribution. In this respect, companies, universities, and other "non-producing entities" that produce and acquire intellectual property on the basis of research utilizing GISRS materials must be required to make a partnership contribution to reflect the benefits that they derive from the WHO system. It stands to reason, however, that contributions will be scaled, and that the assessment for major vaccine, antiviral, and diagnostic producers will be rather larger than that of smaller users of WHO GISRS. Accordingly it seems justifiable that WHO to either invest in developing a database for the purpose of depositing the sequence data or to form a formal partnership with existing databases. The aim would be to allow access to sequence information subject to certain terms and conditions in the format of a click-wrap agreement. Important conditions would include that all data be available to all persons, and that accessing entities must agree not to sell or make any type of proprietary claim over the sequences accessed and an agreement to share benefits if the GSD is used in a manner that financially benefits the accessing entity. If the GSD were transferred to a third party, similar conditions would also bind the third party. These types of click-wrap agreements are frequently found for software and other media and can be adapted for the abovementioned purpose. WHO GISRS laboratories would have to be bound to use this database. The abovementioned database would allow WHO to monitor the sharing of GSD for purposes of implementation of the PIP Framework.
G34.	Stakeholder submission	The disclosure of GSD in publically-available databases, as envisioned by the Framework, does not raise any particular IP issues that do not routinely arise in other scientific settings. However, commercialization would be hampered if untraceable conditions and legal encumbrances are attached to the mere use of publically-available GSD. Physical viral materials are not part of GSD, and unlike PIPBM, which are physically transferred pursuant to negotiated MTAs, GSD is intangible and capable of rapid and widespread dissemination. A virus isolate, for example, is capable of having its provenance and chain of custody well-documented. Thus, moving forward, it will normally be possible to trace the originators and recipients of such material, to ascertain whether it originates from the GISRS network (and thereby qualifies as PIPBM) or whether it was shared by a member state outside the Framework, and which legal obligations attach to the use of such material by its recipients. Moreover, if contractual conditions for use or sharing such material are breached, it will normally be possible to identify where in the chain of recipients such breach occurred. The same cannot be said of GSD. Viral nucleic acid sequences in public databases may not be identified as originating from the GISRS. Substantially identical sequences may have been deposited by GISRS and non-GISRS laboratories. The identity of individuals who accessed the sequence information may not be recorded or ascertainable. If such sequence data are further distributed, reposted, or published, members of the public would not know whether they incur legal obligations for using this information in further research. Such legal uncertainty would be highly detrimental for vaccine and diagnostic companies. [Name of stakeholder] oppose the idea of having WHO monitor or trace the sharing of IVPP GSD. To date, GSD has been publically available and provided to the scientific community and industry through the Global Initiative on Sharing All Influenza Data (GISAID), GenBank and similar databases. In fact, the PIP Framework directs GISRS laboratories to submit GSD to these databases in a timely manner. The PIP Framework does not require GISRS laboratories to impose any particular conditions on the users of these databases, and these databases do not allow submitters to impose special PIP-like conditions on the use of the submitted data. Should WHO pursue efforts to monitor or trace the sharing of GSD, a number of serious issues would need to be addressed related to the feasibility and legality of such efforts. Would WHO establish a proprietary database for IVPP GSD and direct GISRS laboratories to no longer deposit sequence data in public domain databases? How would WHO ascertain the identity of database users, and the legal capacity of such users to agree to and bind their employers to any conditions of access? Would the conditions of access, and any user obligations, be subject to perfunctory clickwrap licenses or would lengthy contract negotiations be required? How would WHO trace downstream sharing of GSD once it has been downloaded by a user? How can conditions of access in a database use agreement be enforced against

		downstream users of the information, if such downstream users themselves neither agreed to any use restrictions nor had reason to believe that any conditions attach? If substantially identical sequence information is accessible through alternative databases or from sources outside the Framework, will users be required to document the provenance of any sequence information in their possession to establish its status as "PIP-GSD" vs. "non-PIP-GSD?" The use of IVPP GSD from the GISRS may already trigger obligations under Section 6.14.3 of the Framework, which pertains to the annual Partnership Contribution (PC). According to the final text of the PC methodology dated May 8, 2013 and titled "Pandemic Influenza Preparedness Framework: Distribution of Partnership Contribution among Companies," the term "use of GISRS" is defined as "receipt of physical materials or use of data and/or information, some of which may not be routinely provided to the general public."7 The document further states that "the definition is intended to encompass more entities than those receiving physical PIP biological materials (i.e., those that must conclude Standard Material Transfer Agreement 2s)." Companies self-identify by answering "yes" to Question 3 of the annual PC Questionnaire regarding use of GISRS,8 rather than having the WHO construct an apparatus to track and monitor the sharing of GSD which would have broad-reaching negative consequences. We support the continued reliance on independent analysis and decision-making by companies to determine whether or not they have "used the GISRS." Under such an analysis, accessing the GISRS and retrieving GSD and other information from the GISRS could well constitute "use of GISRS," depending on the circumstances. The receipt/download of IVPP GSD does not and should not constitute receipt of PIPBM for the purpose of triggering an obligation to conclude an SMTA2 under Annex 2 of the PIP Framework, as such obligations attached to the use and sharing of publically available GSD would surely inhibit further research and development in the field of influenza. Currently, the receipt of PIP BM, meaning clinical specimens, isolates, extracts, or other physical samples as defined in Section 4.1 of the PIP Framework, obligates the recipient to certain conditions determined under the SMTA. The SMTAs govern the agreed-upon rights and obligations as a condition for the transfer and use of these physical materials. The mere exchange of data without an accompanying physical specimen, for example, is not subject to the SMTA (1 or 2). Thus, GSD can be rapidly exchanged through open access databases such as Genbank and GISAID as well as through the GISRS. The PIP Framework clearly envisions that GSD would be placed in the public domain for the benefit of the public rather than be subjected to proprietary rights under which such data could only be accessed in exchange for valuable consideration. [Name of stakeholder] believe the Framework is unambiguous on this point: it discourages GISRS laboratories from acquiring IP rights in PIP biological materials9 and directs these laboratories to promptly disclose sequence information derived from these materials in public-access and public domain databases. It does not require GISRS laboratories to impose any particular conditions on the users of these databases, nor do these databases allow submitters to impose special PIP-like conditions on the use of the submitted data. Accordingly, [name of stakeholder] believe that GSD should continue to be shared in open databases for the public benefit and in the interest of public health, as envisioned by the PIP Framework. The Framework was drafted and adopted without any mechanism for placing special restrictions and conditions on the access to such data. In fact, placing restrictions on access and use of GSD would delay the dissemination of this important information within the scientific community and erect systematic barriers to access by public and private researchers alike.
G35.	Stakeholder submission	This statement encourages the reader to understand 'public-domain such as GISAID' and 'publicly-accessible such as GenBank'. It is inconsistent with the more precise statements in the PIP-FW, referred to in reference 4. This characterization is potentially damaging, especially as GISAID was created in 2008 foremost as an alternative to the public domain sharing mechanism (e.g. GenBank), in an effort to break the sharing deadlock that existed prior to GISAID's emergence. If the intention was to distinguish between two choices for sharing genetic sequence Data with the public (and it seems an attempt was made by explicitly referring to two methods that are legally entirely different mechanisms) the TEWG report could simply state: *"Those terms of reference include an obligation to deposit the genetic sequence data ('GSD') of the IVPPs in publicly-accessible databases, such as GISAID or GenBank, for pandemic influenza risk assessment and response."* […] The first statement misleads the reader into understanding that sequence information must be uploaded to both GISAID and GenBank, albeit it adds the option of "or similar databases". It is, to say the least, confusing and poor guidance, as to what one should expect, on such an important topic of global health policy. While this statement is included in the Guiding Principles (reference 4), Core Terms of Reference in the PIP-FW (Annex 5), in contrast, state simply and clearly that WHO Collaborating Centres (clause B5) and WHO H5 Reference Laboratories (clause B6) upload sequences "to a publicly accessible database in a timely manner". […] The statement directly challenges the fact that GISAID's database is publicly accessible, when it refers to databases with access agreements as closed databases. The public can and are accessing the GISAID database, despite a requirement for its users to identify themselves before accessing the database and consciously agreeing to its (database) access agreement. Either a database is accessible by the public or it is not. Whether a database is (operated by a) public or private (institution) is entirely irrelevant, as access by the public is the essential criterion of a publicly accessible database. The GISAID database is hosted and maintained by a public institution of the German Government. […] This statement further misleads and implies that only from databases like GenBank, ENA and DDBJ do users have the ability to alter the sequences, compile data into aggregated sets of sequences and further distribute same. Users of the GISAID database can do all the above, provided the data they exchange are shared with individuals who are also registered users of GISAID.[…] Furthermore, contrary to the GenBank, ENA and DDJB databases, GISAID knows the identity of those accessing and downloading data from its publicly accessible database, enabling its data contributors to respond should any irregularity with the use of their data occur. These characteristics do not reduce the openness of the GISAID database. On the contrary, GISAID's modus operandi has fostered the submission of significantly more influenza virus data, especially of recent viruses, making it by far the world's most complete collection of influenza virus data. […] This statement is ambiguous and imprecise. All data in GISAID are openly accessible to the public, regardless of GISAID's request for a one-time identification of those accessing its database, as part of its mission that calls for transparency in access to the data (incl. GSD). […] The report notably omits a major legal point, that once anyone submits sequences to any of the three public domain databases (GenBank, ENA and DDBJ), the interests and intellectual property (IP) rights of

		the submitter are not retained, but removed. GISAID, on the other hand, underlines that the interests and intellectual property (inherent) rights of the data submitter are not affected, hence not removed. Contrary to data that have been deposited in databases that declare GSD as data of the public domain, data deposited in GISAID cannot be subject to the attachment of any (IP) rights (by users), since the interests and intellectual property (IP) rights of its data submitters are not removed. […] The TEWG Report does not include 'public-domain' in the Glossary of Terms, although it uses the term, as does the PIP Framework document. 'Public domain' and its legal significance is clearly defined in the legal field dealing with copyright matters. The definition makes it clear that contributions to the public domain are those of which intellectual property rights have expired, have been forfeited, or are inapplicable. In other words: the submitter of data (GSD) into the public domain relinquishes all of his/her interests and intellectual property rights to the data. […]

Table II.4 Technical Working Group on the sharing of genetic sequence data (TWG) – Excerpts from Member State and stakeholder submissions on "*Draft optimal characteristics of an influenza genetic sequence data sharing system under the PIP Framework*"

Established in April 2015 by the PIP AG, the Technical Working Group (TWG) was requested to identify the optimal characteristics and best practices of a GSD sharing system that best meets the objectives of the Framework.[9] Its report to the Advisory Group "*Optimal Characteristics of an influenza genetic sequence data sharing system under the PIP Framework*" was submitted in June 2016. Member States, GISRS and relevant stakeholders were invited to provide comments on the draft report.[10]

This table contains excerpts from Member State and stakeholder submissions on the TWG draft report "*Optimal Characteristics of an influenza genetic sequence data sharing system under the PIP Framework*". All submissions were reviewed and content relevant to Part B of the Scoping Paper was excerpted.

Reference No.	Source	Excerpt from written submission to Advisory Group TWG
G36.	Submission from Germany	4. One important point is to ensure clarity when addressing the different types of databases. As there are fundamental differences between GISAID EpiFlu and GenBank with respect to the inherent rights on GSD, the wording in the Annex 5, Guiding principles ….., Nr. 9 of PIP-Framework should be corrected in order to avoid the current ambiguity of the language. If a contributor choose to preserve his/her inherent rights while sharing GSD, he or she cannot upload the GSD in GenBank, a public domain database. This doesn't prevent a contributor to upload the GSD in both types of databases, if he or she chooses to give up these inherent rights. 5. The overall objective of the timely sharing of IVPP and access to vaccines and sharing of other benefits (PIP-Framework 2.) does apply for GSD and its products as well. Given the technological development of the recent years, adequate handling of GSD is even more important for pandemic (influenza) preparedness and response, including issues of biosecurity. 6. While describing optimal characteristics of a sharing system it is equally important not to jeopardise the freedom of choice of contributors when sharing their IVPP GSDs in databases. This aspect is of utmost importance for public health authorities as it contributes to the motivation to share GSDs in a timely manner. 7. Reliable and sustainable Databases hosting GSD are crucial for the appropriate handling of GSD. GISAID EpiFlu has proven to be a well-functioning, widely accepted and sustainable funded database, enabling the private, public and scientific sectors, including GIRS, to share GSDs in a responsible way. It is unfortunate that the paper has not addressed this existing sharing mechanism and how it already contributes to the PIP Framework. From our experience, most of the optimal characteristics and best practices, described in this document, are already implemented by GISAID. It would be helpful to acknowledge this when considering sharing systems. First paragraph: the term "publicly accessible database" is not precise enough. We would like to maintain the clear language of the PIP Framework as already mentioned, speaking of "public domain or public access databases", with reference to GenBank and GISAID respectively. Furthermore, it is unclear what is meant and which databases are subsumed under the term "complementary databases with different conditions of access". Number 2: Options 1 and 2: We do not understand the usefulness or practicality of 'sharing' all GSD after 6 months, or any other stated period (a virus may be classified as IVPP at some later date) with INSDC. This would mean that contributors of IVPP GSD who have chosen to protect their inherent rights to their data by submission to GISAID are forced to forego this option and renounce their rights! Since it is not permissible to transfer data from GISAID's database to another database, the data provider would have to upload the data separately to INSDC, thus negating the initial choice. This is, however, unnecessary as the GISAID's EpiFlu database is publicly accessible and has a high quality annotation and curation system for collecting all IVPP GSD, providing a very complete and qualified collection of all relevant IVPP GSD, meeting most, if not all optimum criteria, listed in this document, for sharing data under the PIP FW.
G37.	GISAID submission	This draft document appears to presume that use of IVPP GSD under the PIP Framework (PIP-FW) must be handled very differently from the way that use of PIP Biological Material (PIP BM) is handled. The Influenza Virus Tracking Mechanism (IVTM) was set up to make open and transparent the movement of PIP BM, both within and outside GISRS, whereby each shipment of a PIP BM triggers an SMTA1 or SMTA2, as appropriate. In contrast, it appears that there is a preference by the TWG to ensure anonymity for users of IVPP GSD. While this might be preferred by some, it offers a pathway to avoid Partnership Contributions under the PIP-FW. Would it not be preferable to have equivalent approaches and safeguards for monitoring the use of both PIP BM and IVPP GSD under the PIP-FW? The draft document gives the reader the overall impression that no "IVPP GSD Sharing System" exists for sharing GSD of IVPP in relation to the PIP Framework (PIP-FW), but this is not accurate. While the TWG clearly understands that different databases were established for different purposes at different times, they don't acknowledge the efforts of many to establish not only a database that serves the purposes of the PIP-FW, but also a mechanism that promotes transparent and rapid sharing of GSD in a way that contributed to overcoming some concerns of Member States about sharing

[9] See Technical Working Group (TWG) on the sharing of influenza genetic sequence data, Terms of reference, available at http://www.who.int/influenza/pip/advisory_group/twg_tors.pdf?ua=1.
[10] The Draft report, submissions received and the Final report can be found at http://www.who.int/influenza/pip/advisory_group/gsd/en/.

Reference No.	Source	Excerpt from written submission to Advisory Group TWG
		such data. In essence, GISAID and its data sharing mechanism have been focused on ways to facilitate agreement regarding sharing of influenza GSD that would be acceptable to those in developing and developed countries alike. Essential components of this sharing mechanism/system are that data submitters and data users alike identify themselves and that data users acknowledge those who supplied the samples and those who provided the data. It is critically important for this document to acknowledge that GISAID would not have been created without the H5N1 sample sharing crisis that resulted in the PIP-FW and that members of WHO's GISRS worked together with GISAID to establish a mechanism/system that helped to solve the GSD sharing part of the problem that existed previously. By including a historical narrative for when and why the various databases were created along with their similarities and differences, the TWG will improve immensely on the usefulness of this draft document. […] Finally, It is well documented that mechanisms to share IVPP GSD have been in routine use even prior to the publication of the PIP-FW document in 2011, and the rapid communication, e.g. by Chinese authorities of the full genome sequences of the H7N9 viruses (IVPP GSD) in March 2013, attested to, not only the functionality, but also the acceptance of that GISAID mechanism for sharing data under the PIP-FW. Specific comment 2: Sensitivities about sharing GSD contributed to the emergence of the Global Initiative on Sharing All Influenza Data (GISAID) and its novel sharing mechanism at the heart of the EpiFlu™ database in May 2008. This occurred after extensive consultations with many Member States and most of the interested parties, including components of GISRS, and led to a data sharing system that is entirely consistent with the objectives of the PIP-FW in promoting data sharing and access to benefits. Specific comment 3: It should be recognised that the functionality of the GISAID database in the limited sphere of IVPP GSD is underpinned by its adoption by GISRS (and other influenza networks) as an essential integral component of its year-round surveillance of human influenza including both seasonal and zoonotic infections. […] Specific comment 5: Likewise, it will be difficult if not impossible for most readers to understand the practicalities of Options 1-3 under Best practice 2 (page 7) without a clear explanation of the terms of use relating to GISAID and INSDC databases. Readers should be advised that databases that form the INSDC are Public Domain databases, where access to data by users takes place anonymously, and that such anonymity is contrary to the agreement whereby PIP BM are shared under the Framework. Readers should be advised that databases that form the INSDC are Public Domain databases, where access to data by users takes place anonymously, and that such anonymity is contrary to the agreement whereby PIP BM are shared under the Framework. The PIP-FW's recognition of fundamentally different sharing mechanisms, here Public Domain e.g. GenBank vs. Public Access e.g. GISAID, are of paramount importance to Member States given that it distinguishes for good reason, between Public Domain (US law) where submitters' (IP) Rights are relinquished, and the public access database GISAID, where submitters' (IP) Rights are retained By contrast, the TWG chose to use the ambiguous passage "GISAID and GenBank or similar databases" from a section on guiding principles. Clearly, Core Terms of Reference will supersede guiding principles, especially if it is obvious that language in the guiding principles is ambiguous and invites misinterpretation.
G38.	OFFLU submission	A recent search by this expert of genetic databases where avian influenza sequence data could be posted indicates that currently few isolates are being reported in the public domain. It is not sure that this proposed operating procedure will alter such behaviour. There are suggestions in the document on ways to prevent others from using the data before the team who did the sequencing publishing their data in peer reviewed journals but there may still be concern that once information is released in, say, a risk assessment the power of the publication diminishes given the information is already available. Unless stakeholders are convinced that they won't be disadvantaged they will possibly continue to keep the data in house or only share informally. This expert has a concern that the operating procedure may be largely ignored if voluntary. Perhaps donors could put pressure on institutes conducting work on influenza for early release of GSD they are involved in funding the work. However this would not cover all situations. This expert feels strongly that genetic sequences should be freely available in the public domain so that everyone has the opportunity to access the data for their research and other analyses. Such an approach bests supports the public and veterinary health goals of controlling disease. The draft document provides different options for the recommended best practices. Preferences for each option are included in the summary of specific responses. This expert comments that through his/her communications with some relevant laboratories two major reasons why researchers are not willing to share sequences can be identified. One is that the sequences may be socially sensitive, and thus the submission should be permitted by relevant government officials who are usually not willing to take the risk to issue the permission. Another aspect is that there is no established habit of submitting such data to a database, and hence there is no peer pressure or other competitive pressure to do so. A large part of this effort is to assure that if a vaccine is produced from an isolate, that the originating laboratory receives credit and potentially compensation for making the information available. I am supportive of the originating laboratories getting the recognition they deserve for making the original isolation and their sequencing effort. However, as the database becomes increasingly larger and larger, more analysis of sequences involves hundreds if not thousands of isolates. As the analysis uses larger datasets, it gets harder to acknowledge the source of all the submitters of the sequence. At some point this genetic framework has to consider when the analysis using pooled data needs to be acknowledged in a different way.

Reference No.	Source	Excerpt from written submission to Advisory Group TWG
		I believe 6 months is an acceptable period to allow the original submitters a chance to publish or further protect their intellectual property rights. I believe that it is impractical to require GISAID to get permission from all submitters to allow release of the data. Option 1 provides the ability to extend the time to release to INSDC, but I don't believe sequences should be permanently excluded from INSDC. I would recommend that an additional extension of six months be the limit that sequences are withheld from INSDC. Concern was expressed that it is unlikely that INSDC or the IRD databases will require users to agree to any statement to allow access because of how they are funded and the underlying principles that they operate on. Hence is the comment on "At a minimum, databases should publish…" part of Option 2 or should be a separate option? If that statement is a third option, it would be the preferred option. However, the PIP agreement does not directly involve most veterinary influenza isolates, and it will be impossible to separate out veterinary and human influenza viruses and determine when it is appropriate to include this statement or not. It should not be required of existing databases to do something that they probably can't do Sequences would need to be clearly identified as PIP Framework IVPP GSD at the time of upload so that users are aware of their status. Additional comments: OFFLU guidance on animal influenza sequence databases: Genetic sequences provide key information about the evolutionary, clinical and epidemiological characteristics of influenza viruses; therefore timely deposition of sequence data is a crucial component in protecting animal and human health. OFFLU urges laboratories and scientific institutions to deposit sequence data into publicly available databases. Currently, OFFLU recommends that data are deposited in any 'publicly available database' of the submitter's choice. OFFLU does not specifically endorse or recommend any particular database.
G39.	Third World Network submission	And yet instead of recommending a common data access and use agreement as a standard that has to be complied with by the databases wishing to host IVPP GSD, the conclusions endorses all approaches, even if such approaches undermine achievement of the objectives of the Framework. An alternative proposed by the paper is the inclusion of a general statement, which as explained below is entirely inadequate for implementation of the Framework. Such a general statement is not useful in identifying users or ensuring their adherence to the Framework and operationalizing its benefit sharing aspects. These elements include: an obligation to disclose the use of IVPP GSD in intellectual property applications, scientific publications, in regulatory filings etc, an obligation to use the accession numbers in identifying IVPP GSD, an obligation to share analyses arising from that data with the originating laboratory (Section 5.2.1 of the PIP Framework); obligation to acknowledge the contribution of data providers. In reviewing the draft optimal characteristics, and considering the highest importance of the Framework's balance, on an equal footing, of virus sharing and benefit sharing, it is apparent that a WHO-operated database for IVPP GSD may be the best approach. XXXX has previously raised this as a desirable option in comments to the Advisory Group. A key argument in favor of a WHO-operated database is that it will ensure that the data is promptly uploaded and accessible, and at the same time would allow WHO to easily utilize a data access and use agreement that identifies those that access the data and sets out the expectations of the PIP Framework, particularly its benefit sharing component. Other publicly-accessible databases interested in hosting GSD could access data from the WHO database, provided those other databases also utilize the WHO's data access and use agreement for PIP GSD. This in turn would incentivize databases to utilize a common access and use agreement, an approach the TWG considers "ideal" and best suited to the objectives of the Framework. In such a system, data providers that upload to the WHO database could have highest confidence that the GSD they provide will be treated in accordance with the Framework. This would incentivize sharing and facilitate rapid dissemination of new GSD among databases in the system. Development of technical analytical tools at the WHO database would not be necessary, as these may be offered by other databases in the system. Further, by hosting a database where GSD is initially provided, WHO can ensure the long term storage and integrity of the data (an issue raised in Section VI). A WHO database can also assign accession numbers and ensure consistent use of those identifiers (an issue raised in Section V) through the data access and use agreement. By modern standards the amount of GSD necessary to host, measured as data, is quite small and should not present any major technical obstacles. Collection and curation of associated metadata to optimal standards may involve effort, but given the relatively low number of sequences added annually, does not appear to be an excessively demanding task for WHO either. Further it may be possible to share the burden of this latter task with other publicly accessible databases in the system that are also committed to implementing the PIP Framework. Considering that the PIP Framework is viewed as a success in terms of the benchmark it sets, it is important that WHO adopt a more ambitious long term approach. A WHO-hosted database will enable prompt sharing of data while upholding of the principles of the Framework, and by implementing the common user agreement, it sets a good precedent for how public health GSD should be dealt with equitably. 3. It is also worth recalling that the entire issue of handling GSD was particularly uncertain during the final stages of negotiations of the PIP Framework due to the complexity of the negotiations and uncertainties surrounding the handling of GSD. And as such only some general elements were tentatively agreed with the understanding this matter would be revisited as mentioned in Section 5.2.4. It is thus important for WHO to adopt the best approach for implementing the Framework and achieving its objectives. The TWG has clearly identified the utilization of a common data access and use agreement as the best approach for meeting the objectives of the Framework. […] Approaches that do not lead to the effective implementation of the benefit sharing provisions of the Framework are not coherent with the Framework and do not complement databases that do implement those provisions. Nor is cooperation between noncompliant databases and databases that do implement effective mechanisms to operationalize benefit sharing provisions, an inherently

Reference No.	Source	Excerpt from written submission to Advisory Group TWG
		desirable option to "optimize" the system. In such cooperation, the databases unwilling or unable to implement the benefit sharing provisions may actively undermine the Framework. […] As we noted with respect to optimal characteristic I.2, data providers should upload their data not to any database, but to those publicly accessible databases that implement user agreements that conform to WHO standards and operationalize benefit sharing provisions of the Framework. […] if databases do not implement a common access and use agreement conforming to WHO standards, this optimal characteristic is merely an unenforceable suggestion. Users of GSD could fail to make proper acknowledgement without consequence, whereas acknowledgement could be made mandatory if a proper access and use agreement is implemented by databases. Acknowledgment should extend to the data provider and originating laboratories. […] As mentioned at the outset, the PIP Framework is built on the understanding that virus sharing and benefit sharing will be dealt with on an equal footing. If viruses are shared with databases that are not willing to implement a common data access and use agreement, implementation of benefit sharing will be undermined, making a mockery of the equal footing principle on which PIP Framework developed. […] As not all databases may be willing to carry such a specific agreement, they should at least mention or provide a link to the Framework...", abdicating WHO's benefit sharing responsibility in the Framework and incorrectly concluding that WHO and Member States must allow databases that refuse to cooperate in implementing the Framework to host GSD. In truth, if databases want to host IVPP GSD, they should comply with the requirements and expectations of the Framework. If they do not, they are actively enabling the perpetuation of inequity that existed before the PIP Framework. This is not compatible or coherent with the Framework. […] The Framework's common data access and use standards should not merely "outline the expectations" of the Framework but rather thoroughly describe its requirements and expectations for GSD use, and operationalize its benefit sharing provisions, particularly requirements in the event of synthesis of nucleic acids from GSD. […] With respect to questions of data duplication and storage, we earlier noted that the amount of GSD, as a sheer volume of data, is quite small by modern standards, and questions regarding sustainability seem overstated. Duplication of storage is reasonable, and as mentioned above under "General Comments" it is entirely possible that WHO could itself undertake to house data with its own database. Additional backups are of course possible at other databases that utilize common data access and use agreement consistent with the Framework as discussed above. […] As with a number of other issues confronting the data sharing system, initial deposit into a WHO-operated database (where WHO, in consultation with the data provider, can identify PIP IVPP GSD) and use of a common data use and access agreement (to ensure consistent identification and clarify responsibilities) would obviously be advantageous. […] Provisions in a standard data access agreement, for instance, a requirement to divulge use of GSD in intellectual property applications, would certainly be the most effective way for the GSD sharing system to ensure compliance with the Framework in instances of commercial use. […] An optimal characteristic that all users be required, through the common data access and use agreement, to disclose use of GSD including pertinent accession number(s) in scientific publications and regulatory filings, and product inserts, should be added. This stipulation should form part of the common data access and use agreement.

Table II.5 Excerpts from Member State and stakeholder submissions on "*Options to monitor the use of genetic sequence data from influenza viruses with human pandemic potential (IVPP GSD) in end-products*"

This document was developed by the Secretariat to identify different, potential options to monitor the use IVPP GSD to develop end-products through patent applications, publications, regulatory files, clinical trial files. In November 2016, stakeholders were invited to provide comments on the draft document "Options to monitor the use of genetic sequence data from influenza viruses with human pandemic potential (IVPP GSD) in end-products". The table below contains excerpts from submissions received relevant to part B of the Scoping Paper.[11]

Reference No.	Source	Excerpts from written submissions
G40.	Germany submission	Germany shares the view that "sharing of benefits generated using IVPP GSD" cannot be achieved by one mechanism alone, but only by a combination of different options. Implementation of the benefit sharing mechanism requires the cooperation of many players, as research institutions, international organisations, industry, laboratories etc., and not least databases like that of GISAID. The options listed in the paper give only a rough overview about the components which are required for a broader application of the benefit sharing mechanism in relation to GSD. These are differentiated in upstream options, i.e. tracing and monitoring the downloading of GSD from databases, and in downstream options, looking for the use of GSD in end-products. […] Importantly, the necessary steps to design a benefit sharing mechanism by interlinking the various components of the different options have to be defined and performed in consultation with the players mentioned above. The search engine mentioned in the paper seems to be of special interest; by this particular mechanism, the finding of GSD in publications, patents, clinical trials and regulatory files could be facilitated and thus better contribute to benefit sharing. However there are still many questions regarding the efficiency of such a comprehensive search engine. […] Transparent data sharing and public accessibility of Influenza genetic sequence data are of particular importance for response actions, development of diagnostics or vaccines and basic research. Thus, publishing GSD of IVPP must be made mandatory, while the choice of the use of public-domain or public-access databases for the sharing of sequences of Influenza viruses with pandemic potential and the public accessibility of both must be made clear. As part of the review process this should be implemented in the text of the PIP Framework in an unambiguous way. […] As stated in the Draft Paper, genetic sequence data from IVPP GSD will be generated not only by GISRS laboratories but also by non-GISRS members. These laboratories are often specialized in sequencing and may not necessarily have training in particular aspects of Influenza. These non-GISRS laboratories will generate, publish and use sequences of IVPP strains, but not apply any restriction, tagging or any other upstream monitoring options of IVPP GSD. Furthermore, some non-GISRS laboratories, generating and publishing potential IVPP GSD, report to other initiatives for monitoring Influenza on behalf of global institutions like OIE (World Organization of Animal Health) and FAO (Food and Agriculture Organization of the United Nations). Due to the zoonotic characteristics of H5N1 Influenza and other IVPP strains concerted action by human and animal health authorities is required. […] As there is no reliable definition of Influenza virus with pandemic potential and no consistent reporting system for all GSD submitting laboratories in place, the use of upstream mechanisms to monitor the IVPP GSD by identification, tagging, tracing, monitoring submission or download will not be applicable for these laboratories nor for all institutions that host sequence databases. However, the downloading of GSD from 'public databases with data access agreement', such as GISAID, may well be monitored and traced. […] The usefulness of a general notification statement of obligations under the PIP Framework that are not connected to the use of a particular data entry, is questionable. The example of third-party intellectual property rights is not applicable since public domain databases explicitly do not protect individual property rights. To date, there are no notification statements in public domain databases about the use of data on special groups of Influenza viruses. Thus there are no precedents for this. […] A general statement on a website on the use of IVPP GSD would be easy to implement. In addition it should be recognized that it is already possible to monitor and trace in databases like GISAID the use of individual IVPP GSD entries, in relation to benefit sharing under the PIP Framework.

[11] The draft document, submissions received and the final document can be found at http://www.who.int/influenza/pip/advisory_group/gsd/en/.

Reference No.	Source	Excerpts from written submissions
		[…] A disclosure of all GSD used in a specific research project, independent of a potential invention, is not applicable. Research projects that are unregulated, are conducted over undefined periods of time and may include numerous types of GSD. Therefore only IVPP GSD that are central to the development of an end-product should be included under Downstream Options for linking the use of IVPP GSD to benefit sharing under the PIP Framework. […] The scope of the monitoring system should be limited to uses to manufacture an end-product and not refer to any use of IVPP GSD. Any restriction, monitoring, or measures for disclosure of the origin of GSD for research and public diagnostic procedures would be counterproductive and should be avoided. […] Despite the fact that it is not stated which authority should develop and host such a comprehensive search engine, the success of this system will be dependent on the correct disclosure of IVPP GSD and identification of their origin. This can only be done with consistent use of unique identifiers, i.e. accession numbers from databases. (See also comment below). [...] Best practice guidance for meeting reproducibility requirements should be implemented for any use of IVPP GSD. […] As stated above the scope of such Downstream Options should be limited to the IVPP GSD directly used for the development of end-products. Monitoring of all uses of Influenza GSD e.g. for testing and general animal trials in research approaches is neither necessary, nor applicable. Therefore any obligations to reference sequences should be restricted to relevant GSD. […] Success of WHO guidance will be dependent on a clear description and definition of IVPP Influenza strains. A collaborative approach of stakeholders and institutions is required. The usefulness of the virus name as unique identifier is questionable. Names are subject to changes and can therefore be changed by submitters after submission. In addition databases like GISAID use name/passage combination as identifier that could result in different sequences from different passages for one strain name. Moreover sequence information of a virus with a particular name might be submitted by different laboratories e.g. for frequently used reference strains. All these facts hamper or even circumvent a correct identification of the origin of a sequence by the name only. The accession number is the only unique identifier.
G41.	**GISAID submission**	This is very unfortunate because it is likely that methods to monitor both GSD downloads and end products will need to be put in place in order to accurately monitor use of IVPP GSD to meet the objectives of the PIP Framework. In order to evaluate the merits of all available options for monitoring use of GSD objectively, we must have a clearer description of the features of existing publicly accessible databases, with and without "access agreements and registration requirements", in relation to conditions for use of data. This is essential to understanding how conditions may be 'imposed' on the use of data and the feasibility of attaching any obligations on the use of data for end-products, under the PIP Framework, particularly after the data have been provided to a public domain database. While on the one hand, all databases do have Terms of Use, placing GSD into an archive like GenBank (operated in the United States and therefore subject to the US legal definition for information placed in the Public Domain) relinquishes any inherent (Intellectual Property) rights. The DNA Data Bank of Japan (DDBJ) also makes clear that "The INSDC will not attach statements to records that restrict access to the data, limit the use of information in these records, or prohibit certain types of publications based on these records."2 The draft document therefore does not fully acknowledge the difficulties faced by the PIP-FW in attaching obligations (e.g. Partnership Contributions) to the use of such data. […] While stated that "IVPP are part of a broader set of materials called --- 'PIP BM'", IVPP GSD are not defined in that context within the PIP Framework, since IVPP GSD include only viruses "found to infect humans" (section 4.2), but not, for example, viruses from animals that have been shared with GISRS and used to develop candidate vaccine viruses (included in PIP BM). While GSD of such

Reference No.	Source	Excerpts from written submissions
		viruses would need to be included under a revised definition of IVPP GSD under the PIP Framework, most of these would be identified only retrospectively, which adds an additional layer of complexity. […] In addition to recent developments in synthetic biology technologies, the ease and speed with which sequence data can now be made available (and the amount of data) has also increased the importance of resolving the handling (and sharing) of IVPP GSD as part of the PIP FW These advances also have biosecurity implications (which are for example addressed in the modus operandi of GISAID and its database). […] How would merely "informing entities and individuals accessing GSD … give more legal certainty to users ... and facilitate identification of users", without entering into a licensing agreement (with identified users) as for a "public database with data access agreement"', such as GISAID? The obligations of PIP BM are also not merely communicated through (a notice) informing users, but through bona fide material transfer agreements, e.g. SMTAs Upstream options should therefore also consider the ability to monitor use of GSD directly through mechanisms (already) provided, for example, by the GISAID database. […] GSD can be used in many ways to make products. There can be a one to one relationship between a given sequence and a product, or at the other extreme a product can result from knowledge gained from use of many different genetic sequences, which may be provided by GISRS and non-GISRS labs. The latter will be extremely difficult to monitor and verify. Have there been discussions regarding this scenario? […] It should be noted that every database has some form of 'conditions' i.e. 'Terms of Use' that are agreed to by the user when accessing the database (website). For example, GenBank provides such a notice on the 'Copyright Status' advising its users that 'Information … on this site is within the public domain'. Submitters releasing information into the Public Domain in the United States relinquish all inherent (e.g. Intellectual Property) rights to the data they submit. Incidentally, "Open Access" as described in the Berlin Declaration, does not preclude databases that require user identification. […] While it is acknowledged (below) that "it would be difficult to regulate access to data available in such databases in order to implement benefit-sharing under the Framework", would the attachment of conditions to the use of data not infringe unrestricted rights to use data obtained through 'public domain', for example as defined by US legal definition? On what basis could IVPP GSD be singled out for a special concession in attaching restrictions/obligations - is it not inconsistent with the unrestricted/anonymous use of data that has been placed in the 'public domain'? […] Once data has been (lawfully) submitted into a public domain database (e.g. GenBank) the owner of the data has relinquished any and all (IP) rights to the data without recourse. Therefore any attempt to link conditions to such data would be invalid, and merely posting a notice on a website provides certainly no legal protection. Removing any and all 'third-party intellectual property rights' is at the heart of Public Domain. […] Since public domain databases by default do not conduct positive identification of its users that access GSD, merely posting 'notification statements' somewhere on the website is not helpful, as it would fail to produce any verifiable and auditable acknowledgement of the 'notification' by users accessing IVPP GSD, one that could also be audited. Many users are still unfamiliar with the PIP FW and those who are will not be thinking about the FW when they are doing their research. Unless the user is specifically reminded that use of IVPP GSD may have some requirements when actually accessing a given sequence they will not be aware. In addition, access to such GSD can never be audited. GISAID, on the other hand, has a mechanism that not only provides the ability to tag each IVPP GSD, advising users of the special significance of these data, but also gives member states and their authorities the ability to audit the use of the IVPP GSD they submit to GISAID. […] Although omitted from this section (1.1.2), GISAID also provides the facility to monitor access to and downloading of data, as acknowledged in the report of the TEWG (Oct 2014). As acknowledged

Reference No.	Source	Excerpts from written submissions
		on page 7 of this document, monitoring the use of GSD in end products will capture a more limited number of users than monitoring access to IVPP GSD. ("compared to options to monitor access to IVPP GSD, monitoring the use of GSD in end-products will capture a more limited number of users") […] The potential solutions are quite different. Databases with data access agreements can impose conditions, such as obligations/expectations of the PIP FW, for sharing IVPP GSD and may also afford the capacity to monitor and trace the use of GSD (of which there is little mention in this document) […] Would it be not more practical to monitor all GSD rather than specifically monitor IVPP GSD? […] The ability to trace accession numbers yields significantly more information on how IVPP GSD was obtained, and specifically under what terms and conditions it was obtained, e.g. through public domain or through a license issued by a database with data access agreement. Merely using the 'name of the virus' would reduce the transparency and hamper the ability to effectively implement certain key objectives of the PIP FW.
G42.	**Third World Network submission**	All users of GSD should be subject to a PIP Data Access Agreement implementing the Framework's benefit sharing provisions. This Agreement cannot be implemented with a "notice" type approach, but rather must be explicitly applied before data access is granted. This Agreement should include provisions requiring the user to sign an SMTA2 and, as applicable, make a Partnership Contribution if the user synthesizes or intends to synthesize materials from PIP sequence data. It should also require disclosure of use of PIP GSD in intellectual property claims and publications. While downstream monitoring may be part of the Framework's approach to genetic sequence data (GSD), there are numerous substantial weaknesses in the available methods. These mean that it will not be possible to reliably track use of PIP GSD using downstream methods alone. Moreover, downstream methods may be difficult and complex to implement, and will require significant human inputs, e.g. to manually review and assess patent claims and sequences. On the other hand, more efficient and effective downstream monitoring will be facilitated by a strong PIP Data Access Agreement. For example, by requiring a user to disclose use of GSD in intellectual property applications and publications, measures that are otherwise decidedly impractical or impossible to implement. 1. All users of GSD should be subject to a PIP Data Access Agreement implementing the Framework's benefit sharing provisions. This PIP GSD access agreement should be the same at all databases, and its execution digitally recorded and transmitted to WHO. 2. Implementation of the PIP Data Access agreement at databases with existing access agreement procedures, e.g. GISAID, may initially be easiest, however, the same PIP Access Agreement should be implemented at all other databases with PIP GSD. 3. Databases that are unable or unwilling to require their users' agreement with the PIP Data Access Agreement should not host PIP GSD. 4. The PIP Data Access Agreement should include requirements for the user to: A) Sign an SMTA2 and, as applicable, make a Partnership Contribution if the user synthesizes, or intends to synthesize, materials from sequence data; and, B) Disclose use of PIP GSD in any intellectual property claim arising from use of PIP GSD; and, C) Disclose use of PIP GSD in all publications of research that involves use of PIP GSD. 4. The Framework should have rules for the implementation of the Agreement, to ensure it is consistently applied across all databases hosting PIP GSD. 5. Concepts such as a "notice" will not suffice to ensure the user's compliance or enforceability of the Access Agreement. Notice is likely to be inadequate in facilitating identification of users as it very much depends on whether the user contacts the PIP Framework. "Downstream Options" 1. It is sensible for the Framework to build a capacity to monitor downstream use of GSD, but the limitations of such systems need to be fully understood and acknowledged.

Reference No.	Source	Excerpts from written submissions
		2. Many of the serious limitations for downstream monitoring might be addressed by use of a PIP Data Access Agreement that includes the provisions described above. 3. Downstream monitoring may identify cases where an SMTA2 and Partnership Contribution have been required for some years but were not made. The Working Group should give thought to how to assess the PC in cases where years have passed and payments that should have been due were not made. 4. In the absence of a PIP GSD Access Agreement, it can only be suggested that user clearly and consistently identify GSD in publications and patent applications, i.e. it is optional. 5. Patent applications may deliberately or inadvertently claim PIP BM without any indication or acknowledgement. This complicates downstream monitoring and indicates the need for active human evaluation of patent applications. For example: - Patent claims commonly claim a specified sequence, and any other materials with a certain degree of homology, e.g. 90% or 95%. A patent claim that lists the sequence of an animal virus but which also claims homologous sequences may cover quite a bit of GSD without mentioning a PIP BM strain or listing its exact sequence. - Applicants may hide their use of GSD by lack of identification. For example, the recently published WO2015023461 from the University of Pennsylvania and private company Inovio claims a "consensus HA sequence" of H7 (SEQ ID: 40), indicated as synthetic, but which in fact is the exact sequence of a 2013 H7 isolate from a human infection in China. The patent application does not divulge the origin of the GSD. An international GSD disclosure of origin requirement under patent law would obviously be desirable but is not presently realistic. While creation of such a requirement for GSD under national law, in some countries, may be theoretically possible, the disclosure of information under such a requirement would have to be public (i.e. not only to national authorities), and WHO would need access to its declarations, and need to consistently search all national filings in those countries, in their respective languages and formats, and analyze results. This does not seem practical (however, see below).