Research on highly pathogenic organisms is crucial for medicine and public health, and we strongly support it. This work creates a foundation of new knowledge that provides critical insights around the world’s most deadly infectious diseases, and it can lay groundwork for the future development of new diagnostics, medicines, and vaccines. Almost all such research can be performed in ways that pose negligible or no risk of epidemic or global spread of a novel pathogen. However, research that aims to create new potential pandemic pathogens (PPP) (1)—novel microbes that combine likely human virulence with likely efficient transmission in humans—is an exception to that rule. While this research represents a tiny portion of the experimental work done in infectious disease research, it poses extraordinary potential risks to the public.

Experiments that create the possibility of initiating a pandemic should be subject to a rigorous quantitative risk assessment and a search for safer alternatives before they are approved or performed. Yet a rigorous and transparent risk assessment process for this work has not yet been established. This is why we support the recently announced moratorium on funding new “gain-of-function” (GOF) experiments that enhance mammalian transmissibility or virulence in severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS), and influenza viruses. This realm of work roughly corresponds with the work we have termed PPP above. Because the term “gain of function” in other contexts can be used to describe techniques of scientific research that have nothing to do with the creation of novel potential pandemic pathogens, we think the term can be too broad and can mislead. Throughout this commentary, we focus on research designed to create PPP strains of influenza virus, the type of research that initially attracted attention, leading to the moratorium and for which the most discussion has already occurred. Other types of gain-of-function research on influenza and studies intended to enhance pathogenicity or transmissibility of MERS and SARS coronaviruses may or may not fit the definition of PPP research and further clarification is needed and ongoing. As we discuss near the end of this article, it will be essential to clarify the different risks and benefits entailed by different types of experiments covered by the funding pause (2).

The purpose of this research funding pause is to complete “a robust and broad deliberative process . . . that results in the adoption of a new [U.S. Government] gain-of-function research policy” (3). The moratorium would stop new funding for the following:

. . . research projects that may be reasonably anticipated to confer attributes to influenza, MERS, or SARS viruses such that the virus would have enhanced pathogenicity and/or transmissibility in mammals via the respiratory route. The research funding pause would not apply to characterization or testing of naturally occurring influenza, MERS, and SARS viruses, unless the tests are reasonably anticipated to increase transmissibility and/or pathogenicity. (3)

The new U.S. Government (USG) policy also encourages the currently funded U.S. Government and nongovernment research community to join in adopting a voluntary pause on research that meets this gain-of-function definition. Some 18 NIH research projects that possibly meet that definition have been identified (2). The moratorium does not apply to the larger infectious disease research portfolio supported by the U.S. Government. In particular, it does not affect disease surveillance or vaccine development programs. During the moratorium, a deliberative process will occur that will be led by the National Science Advisory Board for Biosecurity and the National Academy of Sciences. This process is intended to produce “recommendations for risk mitigation, potential courses of action in light of this assessment, and propose methodologies for the objective and rigorous assessment of risks and potential benefits that might be applied to the approval and conduct of individual experiments or classes of experiments” (3).

In this commentary, we discuss key elements of risk analysis and offer an example of an approach that could be taken. We describe benefit analysis, offering an account of the kinds of benefits that are relevant and our own view of those at this point. We note other factors that are important to consider. And we argue that a moratorium is the right approach until a rigorous, objective, and credible risk assessment process can be established.

**RISK ANALYSIS**

Risk assessment for GOF work should be quantitative, objective, and credible. Extensive qualitative arguments have been made on both sides of this issue, and these arguments have not provided sufficient clarity or evidence to resolve concerns or identify a consensus path forward. Quantitative assessments should now be performed so as to provide specific calculations and information to inform decisions. It is also important for these risk assessments to
be objective. Given the stakes in this process, the risk assessment process should be directed by those without a clear personal stake in the outcome, just as peer review of science is performed by those without a direct interest in the outcome. The credibility of the risk assessment will depend both on the rigor of the quantitative process and the perceived objectivity of the process.

The record of laboratory incidents and accidental infections in biosafety level 3 (BSL3) laboratories provides a starting point for quantifying risk. Concentrating on the generation of transmissible variants of avian influenza, we provide an illustrative calculation of the sort that would be performed in greater detail in a fuller risk analysis. Previous publications have suggested similar approaches to this problem (1, 4).

Insurers and risk analysts define risk as the product of probability times consequence. Data on the probability of a laboratory-associated infection in U.S. BSL3 labs using select agents show that 4 infections have been observed over <2,044 laboratory-years of observation, indicating at least a 0.2% chance of a laboratory-acquired infection (5) per BSL3 laboratory-year. An alternative data source is from the intramural BSL3 labs at the National Institutes of Allergy and Infectious Diseases (NIAID), which report in a slightly different way: 3 accidental infections in 634,500 person-hours of work between 1982 and 2003, or about 1 accidental infection for every 100 full-time person-years (2,000 h) of work (6).

A simulation model of an accidental infection of a laboratory worker with a transmissible influenza virus strain estimated about a 10 to 20% risk that such an infection would escape control and spread widely (7). Alternative estimates from simple models range from about 5% to 60%. Multiplying the probability of an accidental laboratory-acquired infection per lab-year (0.2%) or full-time worker-year (1%) by the probability that the infection leads to global spread (5% to 60%) provides an estimate that work with a novel, transmissible form of influenza virus carries a risk of between 0.01% and 0.1% per laboratory-year of creating a pandemic using the select agent data, or between 0.05% and 0.6% per full-time worker-year using the NIAID data.

Readily transmissible influenza, once widespread, has never before been controlled before it spreads globally, and influenza pandemics historically have infected about 24 to 38% of the world’s population (8, 9). The case-fatality ratio of a novel strain is of course unpredictable. The worst case might be a case-fatality ratio similar to that of avian H5N1 influenza virus in people, which approaches 60% (10). A greatly attenuated version of the same virus might have a case-fatality ratio of “only” 1%.

Again, multiplying the pandemic attack rate (24% to 38%) times the global population (~7 billion) times the case-fatality ratio (1% to 60%) would produce an estimate of between 2 million and 1.4 billion fatalities from a pandemic of a highly virulent influenza virus strain.

Putting all these numbers together, the select agent data suggest that a laboratory-year of experimentation on virulent, transmissible influenza virus might have an 0.01% to 0.1% chance of killing 2 million to 1.4 billion, or an expected death toll of 2,000 to 1.4 million fatalities per BSL3-laboratory-year. From the NIAID data, for each full-time person-year of BSL–3 work, we might expect a toll of between 10,000 and over 10 million.

These numbers should be discussed, challenged, and modified to fit the particularities of specific types of PPP experiments. For creation of novel, transmissible, virulent influenza virus strains, they may overstate the risk for the following reasons: (i) most such work is done in BSL3 labs, which may be safer than BSL3; (ii) control measures, including vaccination and antiviral prophylaxis of laboratory workers, might reduce the risk of infection and of spread, although none of these is perfect; (iii) the human case-fatality ratio of an avian influenza virus strain that gains transmissibility could be below 1%; (iv) transmissibility in laboratory animals does not necessarily indicate transmissibility in humans (11, 12); and (v) novel strategies of molecular biocontainment (13), if employed, might reduce the risk of human transmission of a strain used in transmission experiments in other mammals.

On the other hand, these numbers may underestimate the risk because (i) the select agent calculation includes in its numerator only BSL3 labs, but in the denominator, BSL3 as well as BSL2 and BSL4 “registered entities” as separate figures for BSL3 are not publicly available (5); (ii) the rate of accidents is calculated for U.S. labs, while GOF experiments are performed in many countries; if this work expands to some of the many countries with less stringent standards than those in the United States (14), risks could be higher; and (iii) the costs of an accidental pandemic considered here are deaths only, but additional losses would include scientific credibility, nonfatal health outcomes, economic and educational losses, etc.

The illustrative calculations above show that approximate risk estimates are possible for creation of PPP strains of influenza virus. During the deliberative process initiated with this moratorium, the risk assessment approach that is established should be able to provide calculations that reflect these and other available probability and consequence estimates and take into account the range of modifying factors, including those just described. The risk assessment process should also be able to provide calculations related to PPP experiments where the risks are harder to calculate given more limited data, such as enhancement of coronavirus pathogenicity in small mammals.

**BENEFIT ANALYSIS**

On the surface, analyzing the benefits of PPP experimentation would seem more difficult. In the cumulative process of knowledge acquisition that is science, it is hard to see far ahead where a particular type of research may lead. On the other hand, scientists make judgments about the relative merits of experimental approaches on a daily basis in their roles as investigators and grant reviewers. Doing and funding science constitute a process of severe winnowing (especially severe in today’s tight funding climate) in which we choose to pursue one approach and not to pursue others based on judgments of which approaches are expected to have the lowest cost, highest probability of success, and greatest yield of valuable findings, among other considerations. Implicit in this process is the idea of opportunity cost. In prioritizing the week’s or the year’s research work, we do not judge in isolation whether a particular experiment should be done or not done. We decide how to allocate our time and funding among possible approaches, devoting resources to the portfolio of efforts that seems most promising. Similar prioritizations are made by funders when they decide which kinds of research will be funded and which research will not.

The analysis of benefits of PPP experiments should follow this familiar approach. The choice is not between doing PPP experiments and doing nothing. Rather, the appropriate question is, within a portfolio of scientific and public health activities designed
to understand and combat influenza or a coronavirus (or, perhaps, in our portfolio of infectious disease countermeasures more broadly), what are the benefits of including PPP approaches compared to the benefits of expanding other parts of the portfolio to use the resources in another way? From the perspective of public health and the practical goal of preventing and treating flu, alternative approaches include those which, like PPP experiments, seek to enhance our scientific understanding of biology, pathogenesis, and transmission. Alternatives also include efforts to develop treatments and prevention measures, including surveillance, through means other than improving our basic biological understanding of influenza (4). This approach is shown graphically in Fig. 1, which also depicts the risks of PPP research. Such risks should be weighed against the risks of alternatives, which are typically much smaller or even negligible. Figure 1 embodies the idea that PPP research should be a component of our research portfolio only if devoting resources to PPP studies at the expense of alternatives has net benefits that outweigh the unique risks of PPP studies.

This comparative approach to benefits should be informed by a hard-nosed look at the benefits that are readily achievable by PPP experiments, not hypothetical outcomes that could someday lead to unspecified benefits. We acknowledge the possibility that PPP experiments may lead to benefits we cannot today envision. But so could the experiments that are done in their place if support for PPP is reallocated to other scientific approaches. The possibility of unanticipated benefits is surely a reason to do science, but it is not a reason to favor PPP approaches over others, unless some specific case can be made for the unique yet unanticipated benefits of PPP work. Such a case seems hard to imagine for benefits that are by assumption unanticipated.

For example, it has been suggested that mutations or phenotypes identified through PPP experiments could be used to sort through the massive diversity of nonhuman influenza virus strains to prioritize those that should trigger countermeasures, including prepandemic vaccine manufacturing. While this is possible in principle, there are many practical barriers to achieving public health benefits of this sort from PPP studies (15). Lists of mutations, and even phenotypes, associated with PPP studies can be compiled and compared against isolates of influenza viruses from birds and other nonhuman sources (16). We know that these lists are unreliable and can even be misleading: the mutations in hemagglutinin identified by two prominent PPP experiments with H5N1 do not reliably confer human receptor specificity even for other H5N1 viruses (17). The E627K mutation in the PB2 gene, known as a virulence and transmissibility determinant before GOF experiments (16, 18, 19), found repeatedly in GOF experiments in H5N1 (20, 21), and used for pandemic risk assessment in H7 viruses (16), was found in some isolates of the H1N1pdm strain in 2009, leading to concern about possible increased virulence and transmissibility. Yet it conferred neither trait in this genetic background (22).
At this time, the high levels of epistasis—dependence of phenotype on the genetic background in which a mutation is found—make prediction of pandemic risk for any given strain more of an art than a science. Indeed, the very presumption that we will see human cases of an incipient pandemic before that pandemic occurs has never been met in practice (23); we have never observed zoonotic cases of any flu virus before it caused a pandemic. This is not to deny that PPP experiments provide any useful data for surveillance and prioritization. Rather, it is to say that other approaches can also identify such predictors (as in the case of the PB2 mutation [11, 13, 14]) and that the ability to use markers of putative transmissibility or virulence to make reliable predictions remains far in the future (23). The fact that some analysts consider mutations identified in PPP experiments when assessing threats of viruses found in surveillance does not mean that the use of such mutations improves the predictions, a claim for which we have no evidence because no pandemic strain has ever been identified in advance. The analysis of benefits of PPP creation should reflect this state of science.

According to some proponents, the most valuable scientific finding of experiments to make ferret-transmissible mutants of influenza A/H5N1 is the definitive proof that such variants could be produced with a small number of mutations. This could not be definitively proven without doing the PPP experiment to manufacture a potentially pandemic variant of H5N1 (24). While it is now undeniable that ferret-transmissible mutants of influenza A/H5N1 can be created experimentally, the impact on scientific opinion about the risk of a pandemic from H5N1 has been hard to gauge. Prior to the gain-of-function experiments, there was a wide range of expert opinion on the likelihood of an H5N1 pandemic (25). Some influenza experts questioned whether H5N1 was a major pandemic threat. After the publication of the experiments producing potentially pandemic H5N1, one prominent member of this group, Peter Palese, noted the shortcomings of the ferret model for humans and correctly concluded that the question of whether H5N1 can transmit efficiently in people remains unsettled (11), as it must until the phenomenon is directly observed in nature. From a practical perspective, responsible policy makers and public health leaders should have been planning for the possibility of an H5N1 pandemic before PPP experiments on H5N1 were undertaken. In some countries of the world, they were stockpiling vaccines against H5N1 (26, 27) and making plans for non-pharmaceutical (8) interventions in the event of a pandemic. The same remains true after the experiments. We have observed no discernible influence of the H5N1 PPP experiments on H5N1 policy preparations.

**CALCULATING OTHER FACTORS**

During the moratorium, progress should also be made in calculating the risks associated with potential deliberate misuse of PPP strains and with potential deliberate misuse of the information that is created and published following PPP experimental work. This calculation should take into account the possibility of deliberate theft and dissemination by either persons working within a lab or theft by those outside the lab. While the probability of this is likely to be very low for most scientists and most laboratories, it is not zero. There is a precedent of scientists using pathogens from their own labs to cause harm. And as with potential accidents, while the probability may be very low, the consequences could be very high.

This assessment should also take into account the possibility that scientists may deliberately misuse the knowledge gained and published following the experiments by recreating the novel PPP strains in another laboratory using methods from published papers and then purposefully disseminating it. This possibility is typically dismissed out of hand by many scientists. But before dismissing that possibility, an analysis by an assembly of experts in the best position to make that judgment should be conducted. What is the possibility that individuals or groups who would seek to carry out such an act would develop the capacity and skill to carry it out? Given that once knowledge is published, it will be available forever, these questions are not just about the possibility of this happening in today’s world but also anytime in the future. Despite the inherent uncertainties in trying to answer these questions, they should be answered with the best possible expertise.

Similarly, the moratorium should be used as a time to answer, or at least be addressing, another major issue as well: the international approach to funding, authorizing, and overseeing PPP. An accident or deliberate act involving PPP anywhere in the world could conceivably impact the public around the world. Therefore, the community of nations has an abiding interest to set common rules for how this work will be pursued. However, at this point, few countries have begun any kind of deliberative process on an approach to research with these unique dangers. Country X should have the right to know if this work is going on in country Y, and if so, what is being done to ensure it is done with the greatest safety and security. But currently, the way country X finds out about PPP work being done elsewhere in the world is when it is published in a science journal. Given the prestige that some scientists have received for pursuing PPP research, it would be surprising if scientists from countries around the world did not increasingly pursue it. As comparatively less experienced labs decided to pursue this work, this will increase potential dangers.

**A MORATORIUM IS THE RIGHT STEP**

There are prominent scientists who agree that there are potential serious dangers to this work and agree that a risk assessment process is needed but who are opposed to a moratorium being imposed while such a risk assessment process is undertaken. They believe that a moratorium should be avoided for reasons that include the potential damage it can do to the funding and work of that lab and to the careers of those involved in the work.

We have a different view. A substantial number of scientists agree that there are extraordinary potential consequences of the work (15). There is no rigorous, objective, credible risk assessment process to judge the risks and benefits of proceeding with it. We believe that the responsible course is to take a research pause until such a risk assessment process is established, which creates a stronger basis for decisions and actions. This is not solely a scientific issue. It is a scientific and public health and safety issue, and it is an issue in which the public itself has an abiding interest.

We have no interest in stopping scientists from doing their work or preventing laboratories from receiving funding. The narrow and defined area of GOF research intended to create novel potential pandemic strains should be put on pause until the risk assessment process is completed. The same laboratories and scientists whose work has been stopped by the moratorium are free and able to pursue all other avenues of infectious disease research except for that narrowly defined by the GOF definition in the new policy; to the extent that other activities not meeting the narrow
definition in the pause have been included in letters to principal investigators ordering or requesting work stoppage, the boundaries of the funding pause should be quickly clarified to allow important alternative work on flu to continue. We note that there are more than 250 NIH-funded projects listed as active with titles containing MERS, SARS, coronavirus, or influenza (28), of which 18 have been affected by the funding pause. The number that remain on pause may be further reduced by negotiations between investigators and the NIH, which are now under way, that will define which projects truly are within the scope of the moratorium and which do not meet its terms and can resume.

The character and scope of the risk assessments that are applied are important. To establish methodologies and approaches for risk assessment and risk mitigation for this context, it would be valuable to start with a global assessment of the risks and benefits of this realm of research, identifying the common aspects of risk and benefit within PPP experiments and other approaches covered in the funding pause. For example, any risk assessment should include estimates of the probabilities of accidental infection and extensive spread, as well as estimates of the impacts of these events should they occur. The specific values of these estimated parameters will differ for different types of experiments. It will then be necessary to set standards and expectations for the quality and characteristics of risk-benefit assessments for individual experiments, for example, to distinguish coronavirus research from influenza research, enhancements of pathogenicity from enhancements of transmissibility, and other important distinctions.

Given that the term “risk assessment” is used to mean different things by different people, an agreement on an approach to individual risk assessments would be needed to ensure rigor and credibility. Once this kind of analytic structure is established, individual risk assessments for GOF experiments that meet the definition in the new USG policy (3) should become the norm before such experiments are funded. Crucially, this process should be quantitative, rather than relying on unquantified and unverifiable assurances that particular laboratories are safe.

CONCLUSIONS

The results of this risk assessment process are important not only to the U.S. Government, which had been a major funder of PPP experiments, but also to other funders, regulators, and investigators worldwide who consider such experiments. Our support for the funding pause and associated deliberative process does not indicate that we would support a permanent end to all experiments subject to the pause. There may be research endeavors that are subject to the moratorium that have a risk-benefit profile sufficiently favorable to justify their resumption once risks and benefits have been explicitly set forth. After 2 years of debate, we think the balance is evidently unfavorable for experiments to enhance avian influenza virus transmissibility, but other classes of experiments may be different. In the meantime, the moratorium is an appropriate and responsible step while dedicated and rigorous efforts are made to understand the risks and benefits of this work.

REFERENCES


The views expressed in this article do not necessarily reflect the views of the journal or of ASM.