Risk-Benefit Assessment of Pandemic Virus Identification

by

Geetha Jeyapragasan

B.Sc McMaster University (2022)

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of

Master of Science

at the
Massachusetts Institute of Technology
September 2024

Authored by: Geetha Jeyapragasan
               Program in Media Arts and Sciences
               August 16 2024

Certified by: Kevin Esvelt
                 Associate Professor, Program in Media Arts and Sciences

Accepted by: Joseph Paradiso
                Academic Head, Program in Media Arts and Sciences

Risk-Benefit Assessment of Pandemic Virus Identification

by

Geetha Jeyapragasan

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, on August 16 2024 in partial fulfillment of the requirements for the degree of

Master of Science

Abstract:

Pandemic Virus Identification (PVI) aims to assess unknown viruses for their pandemic potential in immunologically naive human populations. While proponents argue that PVI could facilitate targeted spillover prevention and accelerate medical countermeasure development, critics raise concerns about biosafety and biosecurity risks. This thesis presents a comprehensive mathematical framework to evaluate the benefits, biosafety risks, and biosecurity risks associated with PVI research.

Using a combination of mathematical modeling and expert elicitation, we developed a structured approach to estimate the potential impacts of PVI. Our framework suggests that identifying a single pandemic-capable virus through PVI could potentially save lives by reducing natural pandemic risks. However, this benefit is substantially outweighed by the estimated anthropogenic risks from potential accidental pandemic events and deliberate misuse scenarios. The overall expected value of identifying a single pandemic-capable pathogen was estimated to be strongly negative.

Significant uncertainty exists in many key parameters estimated through surveys, with wide confidence intervals reflecting the lack of consensus among experts. Expert opinions varied considerably on topics such as the likelihood of funding for medical countermeasures and the potential for deliberate misuse of pandemic agents. This modeling work primarily aims to provide exploratory estimates to guide future work.

Our findings underscore the urgent need for improved governance of research involving potential pandemic pathogens. This study provides a quantitative basis for ongoing discussions about the balance between scientific advancement and public safety in high-risk areas of life sciences research.

Thesis advisor:

Kevin Esvelt

This thesis has been reviewed and approved by the following committee members:

Thesis Advisor: _____

Kevin M. Esvelt, Ph.D.

Associate Professor of Media Arts and Sciences

MIT Media Lab

Thesis Reader: _____

Marc Lipstich, Ph.D.

Professor of Epidemiology

Harvard T.H. Chan School of Public Health

Thesis Reader: _____

Kenneth Oye, Ph.D.

Professor of Political Science and IDSS

Massachusetts Institute of Technology

Acknowledgements

Table of Contents

Glossary

BSAT - Biological Select Agents and Toxins (list of biological agents and toxins managed by the FSAP)

BW - Biological Weapon

BWC - Biological Weapons Convention

cDNA - complementary DNA

CDC - Centers for Disease Control and Prevention

VDi - Virus Discovery

EV - Expected Value

DURC - Dual Use Research of Concern

FSAP - Federal Select Agents Program

GOF - Gain of Function

LAI - Laboratory Acquired Infection

NIAID - National Institutes of Allergy and Infectious Diseases (U.S.)

NIH - National Institutes of Health (U.S.)

NSABB - National Science Advisory Board for Biosecurity (U.S.)

PPP - Potential Pandemic Pathogen/Pathogen with Pandemic Potential

PVI - Pandemic Virus Identification

Chapter 1: Introduction

"Through its technological applications, science has become a dominant element in our lives. It has enormously improved the quality of life. It has also created great perils, threatening the very existence of the human species. Scientists can no longer claim that their work has nothing to do with the welfare of the individual or with state policies...This amoral attitude is in my opinion actually immoral, because it eschews personal responsibility for the likely consequences of one's actions."

Joseph Rotblatt
A Hippocratic Oath for Scientists
1999

The pursuit of scientific research in the life sciences has had tremendous impacts on our society, improving our understanding of ourselves and the world around us. It has drastically reduced morbidity and mortality, facilitating improvements in public health and medicine, and expanded the boundaries of what was possible regarding the types of lives we could live. We see this in the development of the now widely accepted germ theory of disease, revolutionizing how we understand and respond to infectious diseases, saving billions of lives. Advances in microbiology , molecular biology, and genetic engineering have resulted in advances in almost every domain, including the development of disease or pest-resistant genetically modified crops resistant to damage from pests, development and efficient production of vaccines, and advances in various therapies to treat diseases. These benefits did not come alone, with these pursuits and insights also creating risks, where the research itself or information produced inadvertently or deliberately caused harm. Commonly referred to as the "dual-use dilemma", research within and outside of the life sciences has long grappled with the challenge that the products of scientific research can be used for beneficial purposes (e.g. medical advances) and morally undesirable purposes (e.g. development of weaponry)[1,2].

The consequences of some research products with high dual-use potential have been particularly striking, such as the Manhattan Project's use of the discovery of nuclear fission to develop nuclear weapons in the 1940s, resulting in the devastating bombings of Hiroshima and Nagasaki killing tens of

thousands of people. Other examples of this include the accidental discovery of nerve agents such as tabun and sarin by German chemist Gerhard Schrader, who discovered these lethal compounds while aiming to develop more effective insecticides[3]. This accidental discovery was shared with the Nazi regime by the pharmaceutical company Schrader worked for, who quickly began stockpiling the nerve agent in large quantities. Although the Nazis did not ultimately deploy sarin in warfare, sarin was used as a chemical weapon by the apocalyptic cult Aum Shinrikyo, releasing the agent in the Tokyo subway and killing 12, and was used by the Iraqi forces under Saddam Hussein, resulting in the largest chemical weapons attack directed against a civilian-populated region in human history in Halabja, killing thousands and injuring many more [4]. Though the threat of chemical and nuclear weapons still remain today, the dual-use risks of nuclear fission and chemical agents resulted in multilateral treaties such as the Nuclear-Test-Ban Treaty (1996) which prohibits all nuclear explosions, and the Chemical Weapons Convention ratified in 1997, which bans the production and stockpiling of nerve agents such as sarin.

The life sciences have presented unique challenges compared to other forms of dual-use research, where technologies and scientific insights in certain areas of life sciences have been contentiously debated with little consensus achieved. Most notably, research involving potential pandemic pathogens (PPPs) have come under the spotlight time and time again, with disagreements surrounding the value this research provides towards pandemic prevention as well as the risks posed by this work. PPPs are generally defined by the following characteristics: (1) highly transmissible, capable of efficiently spreading widely and uncontrollably amongst human populations, (2) high virulence such that it can cause substantial morbidity and mortality in humans, (3) lack of pre-existing population immunity to the pathogen, and (4) genetically distinct from known pathogens currently circulating, such that existing medical countermeasures are likely to be ineffective [5,6]. There are various forms of research involving PPPs, such as the reconstruction of extinct pandemic pathogens [7], characterization experiments that aim to evaluate a virus's characteristics to assess its pandemic potential, and gain-of-function (GOF) research that either confers pandemic potential to a virus or enhances the characteristics of an existing PPP. Proponents of research with PPPs note this work is crucial for preparing against future natural pandemic events, while critics cite the biosafety and biosecurity risks associated with handling and publishing information regarding high-consequence pathogens, increasing the risk of anthropogenic pandemic events [8–10].

The challenges of navigating the dual-use potential of work with PPPs can be seen throughout history. From a biosafety perspective, well intentioned actors and groups have inadvertently caused harm while working with pandemic agents. In 1955, a man-made polio outbreak occurred when the pharmaceutical company Cutter Laboratories failed to properly inactivate the polio virus in their

vaccine, resulting in over 200,000 children accidentally injected with a live polio virus, killing five and leaving hundreds disabled [11,12]. In 1978, a medical photographer at the University of Birmingham Medical School was accidentally exposed to the smallpox virus due to a laboratory accident. This resulted in the photographer's death, marking the last recorded case of a smallpox-related fatality after the disease had been fully eradicated [13]. Malicious actors such as have also demonstrated their interest in PPP research to develop biological weapons (BWs), The Soviet BW program Biopreparat applied advances in molecular biology and genetic engineering in their efforts to develop more effective BWs, using viral pathogens such as variola virus, Ebola, and the Marburg virus. They modified wild-type pathogens to be resistant to antibiotics and vaccines, developed chimeric viral weapons to combine the characteristics of multiple pathogens, and mass-produced the variola virus using advances in modern cell culture methods [14]. Other groups such as the Aum Shinrikyo cult and Al-Qaeda also expressed interest in use of PPPs such as Ebola and the bubonic plague to cause large-scale harm [15,16].

 Some of these historical events have led to concrete changes, such as updated regulation around vaccine production, improved biosafety standards, and restricted access to certain high-consequence pathogens with stricter oversight. However, there is often a strong lack of consensus regarding how these risks should be managed making governance particularly challenging, and has at times led to inconsistent regulatory decisions. There has been contentious debate for decades surrounding whether the remaining smallpox stocks around the world should be destroyed following the eradication of the disease, with some advocating for the destruction of these stocks to prevent accidental release or misuse , wile those supporting their retention argue preserving them allows for further research to develop medical countermeasures to protect against a potential smallpox bioterrorism attack [17]. This debate mirrors the larger debate surrounding PPP research, with arguments in favor suggesting researching these agents is crucial to better understand these threats to improve surveillance efforts and countermeasure development, with those opposed raising concerns potential of an accidental release of a PPP, as well the dual-use insights that may aid malicious actors seed a pandemic [2,18]. Those opposed to this work also suggest there are safer alternative approaches to achieve the same potential benefits without the risks, and note these should be prioritized over riskier research [6]. Given the complex dual-use nature of PPP research, various efforts have been made over the years to establish governance frameworks and guidelines for responsible conduct in this field.

1.1 History of Dual Use Research with Potential Pandemic Pathogens

One of the earliest developments in the governance of dual-use life sciences research emerged in the mid-20th century with the advent of genetic engineering and molecular biology techniques. Specifically, scientists who had participated in the development of recombinant DNA technology quickly realized

the risks associated with these new capabilities, noting the insertion of foreign DNA into bacteria or viruses could result in the emergence of new pathogens, either by enhancing the virulence of existing organisms or by creating entirely new threats [19]The 1975 Asilomar Conference on Recombinant DNA marked a pivotal moment in addressing these concerns. At this conference, scientists voluntarily agreed on guidelines to safeguard life sciences research, including a temporary moratorium on certain types of recombinant DNA experiments. This self-regulatory approach was a significant step in the responsible management of emerging biotechnologies and served as a complement to existing international agreements, such as the 1925 Geneva Protocol and the 1972 Biological Weapons Convention. While these treaties aimed to prohibit the development and use of biological weapons, they primarily focused on the intent behind the research rather than the specific types of research or potential impacts. The Biological Weapons Convention, for example, prohibits the development, production, and stockpiling of biological agents intended for use as weapons but does not ban research with peaceful purposes, even if it involves dual-use technology (BWC, Article I).

One of the next significant advances in virology was the development of reverse genetics systems, which allowed researchers to generate live, infectious viruses by manipulating and assembling viral genomes in a laboratory setting. Reverse genetics systems provide the capability to generate live, infectious viruses entirely from cloned cDNA, allowing researchers to reconstruct viruses by assembling their genomes in a laboratory setting, rather than needing to obtain live virus samples from the environment. This was first demonstrated in 1981 using the poliovirus, where researchers demonstrated a complete cloned cDNA copy of the poliovirus genome could be transfected into mammalian cells to produce infectious poliovirus [20]. In 1999, researchers demonstrated their ability to synthesize influenza viruses from cDNA through an eight plasmid reverse genetics system [21], allowing researchers to reconstruct viruses and introduce specific mutations and study their effects on the virus. In 2001, a laboratory in Australia introduced the interleukin-4 (IL-4) gene into the ectromelia virus (mousepox virus) using reverse genetics techniques, inadvertently creating a more virulent strain of the virus that was lethal even in immune mice[22], raising significant concerns within the scientific community regarding the dual-use potential of such research. Shortly after, a 2002 study marked another significant advancement in capabilities, where researchers synthesized the full-length poliovirus cDNA from scratch using oligonucleotides based on the virus's genome sequence, which was then transcribed into viral RNA, translated and replicated in a cell-free system, resulting in the creation of an infectious virus solely with the genome sequence, without the need for the original viral specimen [23]. These experiments had sparked concern amongst scientists and the public, noting it was irresponsible to publish research that may serve as a blueprint for bioterrorists. These growing concerns led to the publication of the notable Fink report in 2004, commissioned in response to increasing concerns about the potential misuse of

biotechnological research [24]. This report outlined the dual-use dilemma the life sciences was facing, recommended the establishment of a National Science Advisory Board for Biosecurity (NSABB) to provide guidance on dual-use research, and encouraged scientists to consider the dual-use implications of their work and engage in more responsible research practices. This report also distinguished the varying levels of risk, distinguishing the general concept of dual-use potential with dual-use research of concern (DURC), outlining "experiments of concern" that posed significant dual-use risks such as research increasing virulence or transmissibility of pathogens. One notable aspect about the Fink report is the emphasis on self-governance amongst scientists and institutions to manage the dual-use risks of their own work, rather than formal regulations.

The year after the Fink report was released, two controversial pieces of research were published that re-ignited widespread debate around dual-use virology research. In October 2005, a team at the US Armed Forces Institute of Pathology sequenced the influenza strain responsible for the 1918 Spanish flu through, obtaining a sample by an in situ lung biopsy from a 1918 influenza victim buried in permafrost. The team published the whole genome sequence of the virus, which had claimed approximately 20 to 50 million lives over the course of the pandemic [25]. This sequence was then used by Tumpey et al. at the Centers for Disease Control and Prevention (CDC) to reconstruct the live, infectious virus using reverse genetics, that demonstrated similar properties such as extremely high virulence and replication efficiency as the original wild type. The combination of the genome sequence and reverse genetics protocol raised significant dual use concerns; there was now enough information for malicious actors with a scientific interest in using the virus as a bioweapon, and questioned whether it was necessary to publish all the findings from this work, or conduct this research at all [26]. Some argued this work was crucial to developing a better understanding of why the virus was so lethal, help identify signatures within viral genomes to inform surveillance efforts, and develop countermeasures [7]. This research was published in its entirety after the NSABB deemed it to be suitable for publication, though this decision was met with criticism, noting a proper risk-benefit assessment had not been conducted, and that the tangible benefits of this work were unclear [27]. These controversies resulted in the NSABB releasing a report in 2007 titled "Proposed Framework for the Oversight of Dual Use Life Sciences Research: Strategies for Minimizing the Potential Misuse of Research Information", where they developed a criteria to identify DURC, recommended federal guidelines for oversight, and outline responsibilities and enforcement mechanisms for various stakeholders . Around this time, another report by the National Academies was published, formally titled 'Globalization, Biosecurity, and the Future of the Life Sciences', highlighting the biosecurity risks associated with rapid developments in biotechnology [28]. The report recommended moving beyond a narrow focus on specific pathogens or

experiments to define DURC, instead a more comprehensive evaluation of research based on potential applications and consequences.

The early 2010s saw intense dual-use debates sparked once again by two controversial H5N1 influenza studies. In 2011, Ron Fouchier and Yoshihiro Kawaoka independently conducted experiments that made H5N1 transmissible between ferrets through genetic modifications and serial passaging [29,30]. These studies raised concerns around the world due to the biosafety and biosecurity risks associated with this work, resulting in intense debates and reviews by the NSABB and World Health Organization (WHO) regarding whether this work should be published. Initially, the NSABB recommended publishing the research without complete methodological details to prevent replication. However, after further review, both studies were eventually published. This controversy originally prompted a voluntary 60-day research pause in 2012 by H5N1 researchers. The National Institutes of Health (NIH) had received criticisms from various groups, with some outraged by the restriction on publishing the entirety of the research, while others concerned the NIH had funded this work at all, deeming it exceptionally risky [2,31]. These controversies led the development of various policies, starting with the U.S. government releasing the first policy to govern federal oversight of life sciences DURC in 2012, defining a set list of biological agents and experiments that constituted DURC that would require additional oversight by federal funding agencies, with later policies outlining responsibilities for research institutions receiving federal funding. In 2014 after a series of biosafety incidents involving PPPs, the U.S. government placed a moratorium on federal funding for gain-of-function research involving influenza, SARS, and MERS viruses from 2014 to 2017, noting key uncertainties around the risk-benefit tradeoffs of this research needed to be addressed.

Subsequently, various additional policies have been developed such as the HHS Framework for Guiding Funding Decisions about Proposed Research Involving Enhanced Potential Pandemic Pathogens, and updates to existing policies. Internationally, DURC policies and guidelines have been developed in various countries such as Canada, the Netherlands, China, and Germany. As biotechnology capabilities continue to rapidly increase, policies continue to be updated and introduced. Recent developments include the HHS's 2023 screening framework for synthetic nucleic acid providers and the NIH's 2024 guidelines on research with recombinant or synthetic nucleic acids [32,33]. As the largest funder of life sciences research globally, these policies influence who is able to carry out various experiments with PPPs, what types of research are conducted, the oversight and risk mitigation activities researchers are required to carry out, and how the research products of this work get published and shared.

1.2 Pandemic Virus Identification

Pandemic virus identification (PVI) is a field of research working with PPPs that aims to identify which pathogens pose the greatest capability for causing a pandemic based on their biological properties. Through laboratory characterization experiments, PVI aims to identify a pathogen whether a pathogen *would* cause a pandemic, if an immunologically naive human population was exposed to the pathogen. These experiments aim to determine characteristics such as the host range, cell entry and replication dynamics, immunogenicity, tissue tropism, transmissibility (using animal models), and evasion of existing medical countermeasures. Researchers perform characterization experiments such as measuring the capacity to infect relevant human primary cells, such as airway epithelial cells, replicate to high titers in primary human cells, and transmit between animal models such as ferrets or transgenic mice expressing human receptors. These assessments may also evaluate population immunity, ability to evade pre-existing medical countermeasures, and other measures of pathogenicity. While some experiments aim to solely characterize wild-type viruses, many studies introduce mutations or evaluate potential reassortant viruses to assess whether viruses are within mutational distance of being pandemic-capable.

Examples of PVI include a study Hou et al. who carried out many characterization experiments to evaluate the pandemic potential of a pangolin coronavirus. They evaluated the virus' ability to infect human airway epithelial and nasal epithelial cells, its potential to spread through airborne transmission through animal models, and evaluate whether existing vaccines and antivirals were effective against the virus [34]. Other examples include the evaluation of the H9N2 avian flu viruses pandemic potential through testing their ability to replicate and transmit between ferrets [35]. These researchers evaluated five wild type H9N2 viruses isolated from birds, identified a specific mutation in the hemagglutinin protein necessary for transmission, and created a reassortment virus combining the surface glycoprotein genes from the H9N2 virus and internal genes from a human H3N2 virus. They found this reassortant virus demonstrated enhanced replication and direct transmission though no aerosol transmission was seen. In 2015, Menachery et al. conducted characterization experiments for both wild-type and chimeric versions of the SARS-like WIV1 coronavirus, using reverse genetics to construct the various strains, followed by replication studies in human airway epithelial cells, pathogenicity studies in mice models, and experiments to evaluate the efficacy of vaccines and monoclonal antibodies against the virus [36]. U.S. Government agencies launched programs such as USAID PREDICT and DEEP VZN aiming to discover viruses circulating in zoonotic hotspots, and characterize viruses to evaluate their pandemic potential [37,38].

1.2.1 Risk-Benefit Tradeoffs of PVI

There is a current lack of consensus amongst experts surrounding the risk-benefit tradeoffs of PVI from a pandemic prevention standpoint. Arguments in favor of PVI research suggest it is important to know whether a pathogen is pandemic capable before it spills over into the human population.(Carlson et al. 2021) This would allow for targeted spillover prevention interventions to be deployed and provide academic and industry stakeholders sufficient time to develop pathogen-specific countermeasures earlier, speeding up developments of diagnostics, therapeutics and vaccines . Those carrying out PVI research note it has generally focused on novel unknown pathogens as human populations generally have limited to no pre-existing immunity to novel viruses, and because there are no pathogen-specific diagnostics, medicines or vaccines ready to be deployed in the event of an outbreak. Accordingly, the argument follows that if PVI efforts are successful, we will identify the next pandemic agent before it causes a large-scale outbreak, allowing for more efficient and effective preparedness, prevention and response efforts.

Arguments made in opposition to PVI broadly cover one of two concerns: PVI does little to prevent future natural pandemics or mitigate their impact, and the anthropogenic pandemic risks created by PVI are large enough that they may outweigh the benefits [39–41]. The skepticism surrounding the magnitude of the benefits have been questioned the feasibility of accurately predicting which specific pathogen amongst many will be responsible for the next pandemic event, the regulatory constraints that would need to change for the proposed benefits to be realized once a novel PPP is identified, and how the benefits of PVI compare to other pandemic preparedness interventions. Arguments focused on the dual-use concerns point to the biosafety and biosecurity risks associated with working with potential pandemic pathogens (PPPs) [42]. On the biosafety side, working with PPPs in a laboratory creates opportunities for an laboratory worker or person in the nearby community becoming accidentally infected with the agent being studied, resulting in uncontrolled transmission seeding an accidental pandemic. The biosecurity raised by this work primarily focus on the information risks or hazards posed sharing information about novel pandemic-capable viruses to the public, where malicious actors interested in causing large scale harm may seek synthesize (or obtain through other means) and disseminate the pathogen themselves to deliberately seed a pandemic.

1.3 Project Motivation

The lack of a quantitative assessment of the risks and benefits makes direct comparisons of risks against benefits particularly challenging, and has at times led to inconsistent regulatory decisions. As such, there have been numerous calls for a comprehensive mathematical framework to evaluate the risk-benefit

trade-offs [43]. Developing these may facilitate more informed decision-making, and foster more nuanced and productive discussions regarding not only PVI research, but various types of life sciences research that bear significant dual-use potential. In the following chapters, we construct mathematical models and estimate: (1) the benefits of PVI through its influence on natural pandemic risks, (2) the biosafety risks of PVI influencing the risk of an accidental pandemic event, and (3) the biosecurity risks of PVI through its influence on the likelihood of a deliberate pandemic event.

Chapter 2: Benefits Assessment

2.1 Background Information

Most global pandemics responsible for over a million deaths have originated from zoonotic spillover events. In these cases, a virus circulating in non-human animals jumps into humans, followed by human-to-human transmission. This allows the virus or its chimeric descendant to spread across the world[44]. Many viruses capable of efficient human-to-human transmission might never spill over, while those that frequently spill over from animals to humans do not tend to transmit efficiently between humans [45]. In some cases, a poorly-transmitting animal virus that spills over may acquire key human-adaptive mutations or recombine with an endemic human virus to generate a chimera capable of efficient transmission in humans [46].

Pandemic prevention efforts aim to reduce the likelihood of spillover by identifying geographic hotspots, monitoring the animal-human interface, and empowering local communities to detect and suppress epidemics before they spread. While some efforts are directed at preventing known high-consequence pathogens from seeding another outbreak, many efforts are focused on a potential unknown threat, often referred to as Disease X. The concept of Disease X refers to a serious epidemic or pandemic caused by a currently unidentified pathogen with pandemic potential, commonly termed Pathogen X. Pathogen X has been placed on lists of priority pathogens for research and development by the World Health Organization and the Coalition for Epidemic Preparedness Innovations [47–49]. Given the history of recent pandemics and typical characteristics of pandemic pathogens, Pathogen X is most likely to be an RNA virus [50,51]. Due to this, we use the term "Virus X" to represent the hypothetical virus responsible for the next pandemic event for the remainder of this chapter.

To prepare for Disease X, efforts are being made on various fronts to develop medical countermeasures such as rapid-response platform technologies capable of producing vaccines against unknown pathogens, and bolster non-pharmaceutical interventions such as early warning surveillance systems [47,52]. Some research initiatives focus on broad strategies to identify and prepare for a range of potential threats, while others target specific high-risk virus families or aim to identify viruses that could potentially be Virus X. These approaches can be broadly categorized into two main strategies: virus discovery (VDi) and pandemic virus identification (PVI).

2.1.1 Virus Discovery and Pandemic Virus Identification

Some efforts seek to discover and sequence new viruses in hopes of mapping natural diversity. This mapping could facilitate prioritization of reservoirs, vectors, and pathogen lineages for surveillance and non-pharmaceutical intervention efforts where spillover is most likely to occur[53]. In theory, these virus discovery (VDi) efforts could also increase the likelihood that a broad-spectrum vaccine or therapeutic effective against Virus X will be available at the start of the next pandemic [54].

Other efforts aim to identify specific viruses that may carry pandemic potential through their ability to cause considerable virulence and transmit efficiently between humans, where it could become Virus X. Pandemic Virus Identification (PVI) described above is conducted with this goal in mind, aiming to characterize viruses to predict with any confidence whether a given virus has the potential to spread efficiently in humans. Proponents of PVI note identification of a sufficiently concerning pathogen could unlock funding sufficient to develop targeted vaccines and therapeutics, and also better target anti-spillover interventions [8]. For example, Nipah virus has merited $100m from CEPI to support vaccine development, with three candidates in clinical trials[55], despite infecting fewer than a thousand people in low-income nations. Therefore, it is possible – if as yet unprecedented – that governments and philanthropies might similarly fund interventions for a laboratory-characterized potential pandemic virus that has not yet spilled over. Figure 1 summarizes our understanding of the two general categories of risk assessment research.

Virus Discovery

- Which geographic regions are zoonotic hotspots?
- Which animals are contributing to zoonotic spillover events?
- Which viruses are frequently circulating at animal-human interfaces?
- Which circulating viruses are predicted to bind to human receptors?

Pandemic Virus Identification

- Can Virus X infect relevant human cell lines?
- Can Virus X cause disease in animal models?
- Does Virus X transmit between relevant animal models?

Figure 1: Spillover risks can holistically be assessed across pathogens through virus discovery efforts [left], while onward transmission risks can be evaluated for individual pathogens through pandemic virus identification efforts [right].

2.1.2 Challenges of Cost-Benefit Analysis

Cost-benefit analyses of VDi and PVI are hampered by a lack of empirical data on the efficacy of different interventions. The financial costs of pandemics are relatively well-established, with the World Bank estimating that moderate to severe pandemics could cause 14.2–71.1 million deaths and a global GDP decrease of 2%–4.8%[56]. Due to this, it is generally established that pandemic prevention interventions could generally be quite cost effective, with Dobson et al. estimating that an annual investment of $30.7 billion in prevention would be cost-effective if investments could reduce spillover risks by 26.7% [57]. Bernstein et al. estimated that 3.3 million deaths and $212 billion are lost annually due to viral epidemics [58], and suggested $20 billion as the median estimate for feasible annual primary prevention costs (viral discovery and surveillance, wildlife surveillance and management, and deforestation reduction).

While many studies have examined drivers of zoonotic risks [48,56–61], evaluation of the cost-effectiveness of interventions is severely hampered by the dearth of empirical data or estimates of their efficacy. It is generally accepted that earlier prevention interventions are more cost effective than control measures later into an outbreak, but efficacy estimates of specific interventions are lacking. Madhav et al. note costs associated with pandemic preparedness and response interventions are poorly tracked, and data that is available regarding costs and potential benefits are produced by high income countries, likely leading to biased assessments regarding which interventions are optimal [60]. They note for low-middle income countries (LMICs) in particular, the most cost effective pandemic preparedness interventions involve improving core public health infrastructure such as water and sanitation systems.

There is considerable controversy within the fields of virology and global health over the benefits of VDi and PVI. The proposed benefits, based on the anticipated theory of change put forth by researchers in these fields, remain largely theoretical to date. Some critics of this work assert that attempting to discover and successfully predict the virus responsible for the next pandemic is not feasible, pointing to the very large number of viruses in nature [40,62]. Instead, they advocate for monitoring the animal-human interface to better understand spillover risks and spot epidemics earlier. Proponents of VDi have noted the emerging feasibility of computational evaluation to determine which newly discovered viruses are likely capable of binding to human receptors, which could plausibly identify hotspots based on viral diversity – rather than overall biodiversity – and assist the development of broad-spectrum if not targeted medical countermeasures[8]. Others argue that knowing precisely which viruses can transmit well enough to cause a pandemic, or are within mutational distance of that capability (e.g. PVI), is important for developing targeted vaccines and therapeutics[29].

While the potential benefits of these approaches are debated, their costs have been estimated. Dobson et al. report that spillover reduction programs, including initiatives like USAID PREDICT, EPT programs, DARPA PREEMPT, and the Global Virome project, annually cost between $120 million and $340 million USD [63]. The Global Virome Project aims to sequence 70% of unknown potentially zoonotic viruses for $1.2 billion, or $7 billion for the entire virome, then support characterization of the highest-risk viruses and constructed chimeras that could plausibly exhibit enhanced transmission[63]. Proponents of this work suggest these discoveries could have substantial return on investment through enhancing diagnostics and identifying spillover hosts [57,59]. Unfortunately, there is little quantitative data to evaluate the respective merits of different approaches to understanding pandemic risks and interventions.

Here we develop a mathematical framework to estimate the anticipated benefits of VDi and PVI. We establish a baseline risk from naturally emerging pandemic events and their sources, estimate the potential reduction in risk viral discovery and monitoring efforts can contribute, and then model the reduction in risk from PVI. To estimate key parameters for our model and obtain quantitative outcomes, we employed a combination of estimates found in prior literature, close proxies to the parameter of interest (such as the use of seasonal flu vaccine efficacies to estimate broad spectrum vaccine efficacies), and academic surveys sent to domain experts in One Health and medical countermeasure development.

2.2 Methods

To construct the model, we first surveyed the literature to gather a list of pandemic prevention interventions, which we group into three broad categories: preventing initial spillover from animals to humans, suppressing transmission to end nascent epidemics, and mitigating harms from an epidemic that has spread to become a global pandemic (Table 1) [64]. Through this review, we noted which interventions researchers listed as those VDi and PVI could potentially influence through providing information that could influence prioritization or directly contribute information valuable for countermeasure development.

Table 1: Pandemic mitigation interventions

	Category	Intervention	VDi	PVI
	Research	Map zoonotic hotspots	✓	

Preventing spillover (Primary prevention)		Monitor high-risk animal reservoirs	✓	For Virus X
	Land-use	Prevent deforestation near hotspots	✓	
		Minimize habitat fragmentation	✓	
	Markets	Strictly regulate and monitor the wildlife trade	✓	For Virus X
		Strictly regulate and monitor wet markets	✓	For Virus X
		Improve animal husbandry practices	✓	For Virus X
	Education	Educate communities about risks from animal reservoirs	✓	For Virus X
		Offer training in safe practices for at-risk occupations	✓	For Virus X
Suppressing transmission (Secondary prevention)	Diagnostics	Develop targeted rapid tests for specific viruses		For Virus X
		Develop rapid tests for high-risk virus families	✓	
		Equip communities with sequencing-based diagnostics	✓	
		Develop and administer broad-spectrum vaccines	✓	
	Health workers	Train healthcare providers to use diagnostics and vaccines	✓	
		Provide protective equipment and safety training	✓	
	Containment	Preparing for travel restrictions in the event of an epidemic	✓	
Mitigating harm		Develop and approve targeted vaccines		For Virus X
		Develop and approve broad-spectrum vaccines	✓	
		Develop and approve broad-spectrum therapeutics	✓	
		Develop and approve targeted therapeutics		For Virus X

All interventions would benefit from improved resource allocation towards higher-risk communities, reservoirs, and pathogens. VDi can increase the likelihood that any broad-spectrum vaccines will be effective against Virus X, and may improve hotspot targeting, which currently relies on biodiversity estimates. PVI allows for the development of targeted pathogen-specific vaccines and therapeutics and may direct anti-spillover efforts towards regions with identified pathogens. Theoretically, the availability of vaccines or therapeutics for a virus that has successfully begun transmission between humans could prevent an uncontrolled outbreak through ring vaccination efforts, both reducing the chances of a local outbreak spreading globally, and mitigating the amount of morbidity and mortality caused by outbreaks.

To establish a baseline risk of pandemics over the next decade, we evaluated the expected harm of natural zoonotic pandemics based on the likelihood of occurrence and the consequences of a pandemic (measured by the average number of deaths posed by the pathogen over the next decade). Then using the estimated number of pandemic viruses circulating around the world by survey participants, we estimated the expected harm posed by a single pandemic virus - Virus X. We defined Virus X as a novel virus with pandemic potential that has not yet spilled over in humans, currently unknown to humans. We use risk synonymously with expected harm, where:

$$risk\ =\ likelihood\ \times\ consequences$$

To evaluate the benefits of VDi, we considered the scenario where VDi efforts discover and sequence 3 times as many viruses as we know today through metagenomic sequencing efforts.

We then considered the scenario where Virus X is discovered and identified as a pandemic-capable virus. In this scenario, laboratory experiments are conducted to assess whether primary human cell lines are permissive to efficient infection and amplification of Virus X, and whether Virus X is transmissible in relevant animal models (e.g. ferrets, humanized mice). We assume the results of these experiments, alongside the whole genome sequence of Virus X, are made publicly available to be used in surveillance and potentially used for MCM development efforts. Figure 2 below summarizes the overall structure of the model, noting the interventions that may be influenced through VDi efforts and PVI efforts.
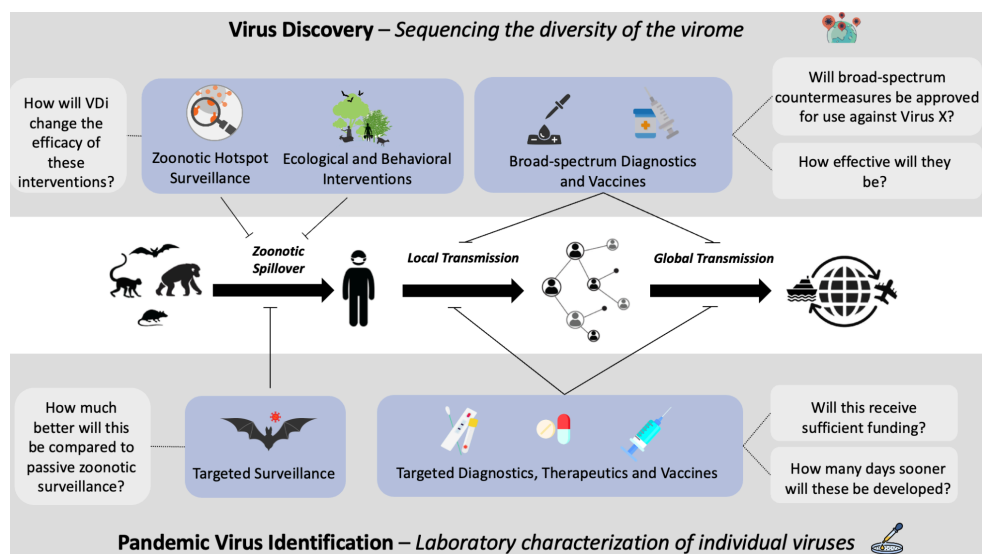
Figure 2: Steps of a global infectious disease pandemic. The top panel outlines pathogen-agnostic and broad spectrum interventions that could be aided by intensified virus discovery efforts. The bottom panel illustrates targeted interventions that could be accelerated through pandemic virus identification efforts.

2.2.1 Model Scenarios and Assumptions

2.2.1.1 Scenarios

This model establishes a framework to estimate the per decade expected pandemic harm from a *Virus X* pandemic in three scenarios: the "baseline" scenario, the virus discovery scenario, and the pandemic virus identification scenario. There are numerous ways in which the world could look with and without VDi or PVI. Below, we establish our assumptions about how each of these hypothetical scenarios would look with respect to research, prevention, and response activities related to Virus X.

Baseline Scenario
- Virus X remains unknown to any humans and only exists in the zoonotic hotspot prior to spillover.
- Virus X's presence is only revealed if X spills over into the human population. The initial group of humans infected with Virus X would result in a cluster of atypical symptoms, alerting health agencies to the novel threat, resulting in X being identified and sequenced. Interventions to prevent zoonotic spillover of Virus X are typical threat-agnostic interventions carried out in all high-risk interfaces without knowledge of a specific threat.
- Development of any Virus X-specific therapeutic or vaccine begins after detection in humans. We assume the development timelines will mirror those of the recent COVID-19 pandemic. Additionally, we anticipate it would take 722 days to develop a targeted antiviral, based on the timeline for Paxlovid's development and approval [65].

Virus discovery (VDi) Scenario
- Samples from animals within a high-risk zoonotic hotspot are collected. These samples are transported back to a lab and sequenced, potentially resulting in the discovery of Virus X (amongst several other pathogens).
- Virus X is potentially amongst the several virus species discovered but is not prioritized or flagged as a pathogen of concern if discovered. If Virus X is in a viral genus/family that contains other viruses that have demonstrated epidemic/pandemic potential (e.g. influenza, coronaviruses, filoviruses, paramyxoviruses), broad spectrum therapeutics and vaccines might be developed for the viral family/genus of Virus X that otherwise would not have covered Virus X.
- If Virus X is circulating in a hotspot deemed to be particularly high priority, either due to its genus or other viruses found in the region, the hotspot may be prioritized for non-pharmaceutical interventions (see table 1 for examples).

Pandemic Virus Identification (PVI) Scenario
- A series of characterization experiments are conducted to estimate the probability Virus X has pandemic potential. The whole genome sequence of Virus X and the results of characterization experiments are publicly published and distributed to relevant stakeholders.
- This information potentially results in targeted non-medical interventions (e.g. Virus X specific surveillance where Virus X was found) and efforts to develop promising therapeutic and vaccine candidates for Virus X prior to spillover.

2.2.2 Academic Surveys

Quantifying the public health benefits of both viral discovery and PVI research efforts can be quite challenging, both due to the inherent challenges associated with evaluating the benefits of any form of scientific research, and the lack of research evaluating the efficacy of various pandemic prevention strategies.

In constructing our mathematical model to assess the benefits of both VDi and PVI, we first used prior literature on historical pandemics as well as data from the recent COVID-19 pandemic to establish a baseline risk from natural pandemics, as well as establish baselines for targeted vaccine and therapeutics efficacies, as well as development and distribution timelines, distribution timelines and efficacies.

We also identified several key parameters that had not been estimated in prior literature. To gather estimates for these parameters, we sent out two surveys to experts in relevant fields. Participants for the survey were selected based on a criteria established around academic journal publications. Specifically, we identified all authors who had published at least twice between Jan 1, 2019 and May 1, 2023 in any of the following journals: *The Lancet Infectious Diseases*, *Emerging Infectious Diseases*, *Immunity*, and *Nature Reviews Immunology*. The publications had to be either in the 'Article' or 'Review' category, and could not have more than 30 authors. The two publications did not necessarily need to be in the same journal. This provided us with a list of 3,557 authors. We sent out two surveys, an initial survey asking participants to estimate various parameters of our model (n=207), and a followup survey to gather clarification for specific parameters where there was ambiguity, and estimate additional parameters (n = 42). Table 2 below outlines some key parameters of the mathematical model, noting which parameters we were able to generate estimates for using prior literature, and which required input from relevant experts.

Table 2: Key Parameters of VDi and PVI Benefits Assessment Model

Parameter	Value	Description	Source
v	65	Effective number of pandemic-capable viruses with equally likely probabilities of seeding a pandemic event.	Survey
$\Delta p_{BSV\,|\,VDi}$	9%	Increase in likelihood of an approved broad-spectrum vaccine effective against Virus X prior to spillover through discovering 3 times as many viruses as today	Survey
m_{BSV}	42%	Relative reduction in harm due to the availability of a broad-spectrum vaccine effective against Virus X prior to spillover	Literature
$\Delta p_{NPI\,|\,VDi}$	14%	Relative increase in likelihood of improved non-pharmaceutical interventions from VDi	Survey, inferred (Methods)
$\Delta r_{NPI\,|\,VDi}$	38%	Relative reduction in harm from better targeting of non-pharmaceutical interventions towards potential Virus X hotspots	Survey
p_{TMCM}	35%	Probability of targeted Virus X medical countermeasures receiving sufficient funds for development prior to spillover and outbreak	Survey
Δm_{TV}	54%	Relative reduction in harm due to earlier release of targeted vaccines due to identification of Virus X	Literature + Survey
$\Delta r_{NPI\,|\,PVI}$	52%	Relative reduction in harm due to targeted non-pharmaceutical interventions from identification of Virus X	Survey

The complete parameter table, survey data and data used to estimate the remaining parameters can be found in Appendix A.

2.2.3 Parameter Estimation

2.2.3.1 Number of Pandemic-Capable Viruses

In our model, we used survey data to estimate (1) the total number of pandemic-capable viruses currently circulating around the world, v_{total}; and (2) the statistical equivalent number of viruses with equally likely probabilities of seeding a pandemic event, v.

To estimate (1) **the total number of viruses,** we first applied a log transformation to the lower and upper bound of each category in question 3 of the initial survey. We then took the log midpoint of each group X_i, using 1 as the lower bound for the "less than 10 viruses" category, and 100,000 as the upper bound for the "more than 30000 viruses" category. We then used these midpoints to calculate the weighted average, using the number of respondents for each category as the weights, w_i. We then exponentiated this weighted average to convert the value back to the normal scale.

$$Weighted\ log\ average\ =\ \frac{\sum\limits_{i=1}^{n} w_i X_i}{\sum\limits_{i=1}^{n} w_i}$$

$$=\ \frac{(0.5 \times 17 + 1.24 \times 39 + 1.74 \times 33 + 2.24 \times 26 + 2.74 \times 22 + 3.24 \times 20 + 3.74 \times 15 + 4.24 \times 3 + 4.74 \times 10)}{185}$$

$$Weighted\ log\ average\ =\ 2.23$$

$$Average\ =\ 10^{2.23}\ =\ 172$$

The median category was 30-100 viruses. Using a similar approach, this would result in a median of $10^{1.74}$ = 55 viruses.

However, PVI efforts are most likely to identify those circulating at high-risk hotspots which are at greatest risk of spilling over in human populations, as viral discovery efforts tend to target key taxa that are most likely to carry zoonotic viruses, such as non-human primates. Participants noted the top 20% of most risky pathogens contribute to 70% of the expected mortality in the followup survey. To address this diversity in likelihood in our calculations, we adjusted the 172 estimate using this risk distribution to get the (2) **effective number of viruses** with equally likely probabilities of seeding a pandemic event.

20% of pathogens contribute to 70% of the risk $\rightarrow 0.2\ \times 172\ =\ 34$
80% of pathogens contribute to the remaining 30% of the risk $\rightarrow 0.8 \times 172\ =\ 138$

Let v be the *effective number of viruses*. We use a weighted approach, using the proportion of risk each group of pathogens contributes to as the weights.

$$0.7 \times \frac{34}{v}\ +\ 0.3\ \times\ \frac{138}{v}\ =\ 1$$

$$v\ =\ 34\ \times\ 0.7\ +\ 138\ \times\ 0.3$$

$$v\ =\ 65.2\ \approx\ 65$$

2.2.3.2 Vaccines

Both VDi and PVI have the potential to influence timelines associated with vaccine development. We established the baseline Virus X scenario as one where there is initially no vaccine available, and a targeted vaccine becomes available 357 days after the outbreak begins.

With VDi, we considered the scenario where a **broad-spectrum** vaccine (either pan-genus or pan-species) is developed prior to the Virus X outbreak and immediately distributed once the outbreak begins. This does not influence the timelines associated with the targeted vaccine, which are developed and distributed in the same way they are in the baseline scenario.

With PVI, we considered the scenario where a **targeted** vaccine begins development prior to the outbreak rather than after the virus has spilled over, resulting in an accelerated vaccine approval and distribution. The additional deaths averted come from this earlier release.

To generate quantitative estimates, we used COVID-19 mortality and vaccine data to both establish a baseline and to evaluate the alternate scenarios.

Broad-Spectrum Vaccines (m_{BSV})

The parameter m_{BSV} in the model represents the reduction in harm due to the release of a broad-spectrum release at the beginning of the pandemic. We estimate this parameter by using data from the first two years of the COVID-19 pandemic to establish the baseline cumulative death toll and vaccine coverage. In this scenario, we make the following assumptions:

- The broad-spectrum vaccine is released 300 days prior to the release of the targeted vaccine, approximately 50 days into the pandemic.
- Vaccine distribution follows the same trajectory as the targeted COVID-19 vaccines in the U.S. and U.K.
- The broad-spectrum vaccines are half as effective as targeted vaccines at preventing mortality.

Więcek et al. estimate 240,715 additional lives could have been saved between the U.S. and U.K. if the COVID-19 vaccine was released 90 days earlier[66]. We extrapolate that if the vaccine had been released 300 days earlier (approximately 50 days into the outbreak), this would have resulted in an additional 802,383 lives saved between the U.S. and U.K. Using our assumption that a BSV would be half as effective as a targeted vaccine, a BSV being released 300 days prior to the release of a targeted vaccine

would result in an additional 401,191 lives saved within the U.S and U.K. Given the U.S. and U.K. reported a total of 964,000 COVID-19 deaths within the first two years, we estimate mBSV to be:

$$m_{BSV} \ = \ \frac{401{,}191}{964{,}000} \ = \ 0.416$$

Accelerated Targeted Vaccines (Δm_{TV})

The parameter Δm_{TV} represents the additional reduction in harm due to the earlier release of a targeted vaccine. In our expert survey, the median response of how much earlier a targeted vaccine would be released was 198 days. In our followup survey, when asked how long it would take to develop a targeted vaccine without PVI, participants provided a median estimate of 382 days, closely matching the 357 days it took to from the beginning of the COVID-19 pandemic to when the Pfizer-BioNTech vaccine received emergency use approval [67]. Similar to above, we estimate m_{TV} by considering the counterfactual scenario where the targeted COVID-19 mRNA vaccines were approved for use 198 days earlier. To estimate the additional deaths prevented, we primarily draw from the 2023 Więcek et al. study which estimated the potential lives saved by earlier COVID-19 vaccination in scenarios where vaccines were available 30, 60, or 90 days earlier than the actual timeline.

They estimate that within the first two years (by Jan 2022), between the U.S. and U.K. 240,715 [117,731; 332,397] additional deaths would have been prevented if targeted vaccines were released 90 days sooner, for an average of 2675 deaths prevented per day between these two countries. Over the course of 182 days, this would result in 524,300 deaths prevented between the two countries. The U.S. and U.K. reported a total of 964,000 COVID-19 deaths by Jan 2022. This suggests roughly 0.54 an additional life could have been saved for every death recorded.

$$\Delta m_{TV} \ = \ \frac{524300}{964000} \ = \ 0.54$$

There were approximately 5.49 million deaths due to COVID-19 recorded by this time. We estimate that globally, this means a targeted vaccine released 198 days earlier would save an additional 2.965 million lives globally for a pandemic similar to COVID-19.

2.2.3.3 Therapeutics

Accelerated Targeted Therapeutics (Δm_{TT})

The parameter Δm_{TT} represents the additional reduction in harm that would come from an earlier release of a targeted Virus X therapeutic due to PVI efforts. To estimate m_{TT}, we considered how much earlier the therapeutic would be released, how effective it would be, and how many of those infected would have access to the therapeutic. During the COVID-19 pandemic, Pfizer developed Paxlovid (nirmatrelvir–ritonavir), an orally administered antiviral therapy. Paxlovid was approved for use 722 days into the outbreak (Fig. 3) and was reported to reduce the risk of hospitalization or death by 88% amongst unvaccinated high-risk patients with COVID-19. We use these values as proxies for the efficacy and baseline timeline of a targeted Virus X therapeutic.
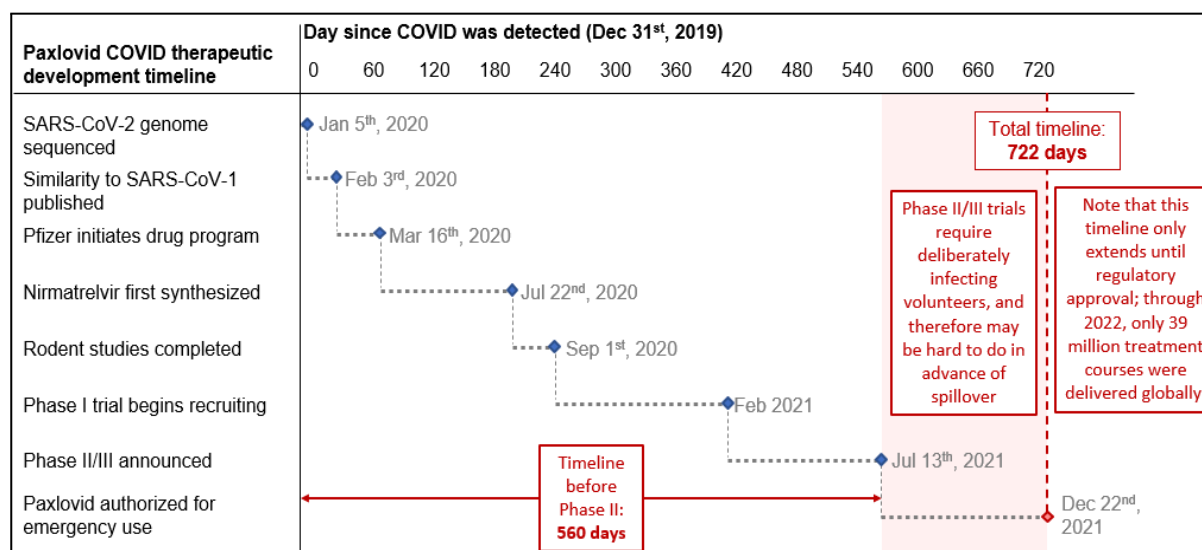
Figure 3: Key milestones in the development of Pfizer's COVID-19 therapeutic from the initial sequencing of the genome to the date the medication received emergency use authorization from the U.S FDA.

According to our survey, the median response to how many days sooner a Virus X therapeutic would be available if sufficient funds were invested and the virus had been characterized and flagged in advance was 300 days. We'll estimate m_{TT} by estimating the additional reduction in harm that could have been achieved if Paxlovid was approved February 24, 2021. On May 5th 2023, the WHO declared the end of the COVID-19 pandemic, at which point 6.93 million deaths were recorded [68]. We use this as the baseline number of Virus X pandemic deaths.

On day 722 of the pandemic, 5.3 million cumulative deaths were recorded, while on day 422, 300 days earlier, 2.62 million cumulative deaths were recorded, such that 2.62 million deaths occurred between these days [68]. The CDC reported a 28.4% adoption rate amongst eligible patients in the U.S. between April and August of 2022 [69]. Using these values, we estimate that 650,157 additional lives could have been saved during that period through access to Paxlovid.

$$Additional\ Lives\ Saved\ =\ deaths\ during\ 300\ days\ \times\ adoption\ rate\ \times\ efficacy$$
$$=\ 2,620,000\ \times\ 0.282\ \times\ 0.88$$
$$=\ 650,179$$

Now taking into consideration the total deaths recorded during the pandemic:

$$\Delta m_{TT}\ =\ \frac{additional\ lives\ saved}{baseline\ pandemic\ deaths}\ =\ \frac{650,179}{6,930,000}\ =\ 0.094$$

Therefore we estimate that if a targeted therapeutic were to be accelerated by 300 days due to PVI efforts, this would result in an additional 9.4% of deaths prevented during a Virus X pandemic.

2.2.4 Model Equations

2.2.4.1 Baseline Risk

Since 1889, five natural pandemics have killed over a million people within a few years of spilling over. Historical data suggests a 3.75% annual likelihood of a natural pandemic event, with an average severity of 18.1 million deaths. These results closely match those of Fan et al., who estimate the overall annual probability of a pandemic to be 3.6% with an average severity of 21.6 million deaths [70]. For simplicity, we use their figures for subsequent calculations. This results in approximately 7.8 million expected deaths per decade from zoonotic pandemics, underscoring the importance of effective interventions.

$$E[Harm_{base}]\ =\ 21,600,000\ \times\ 0.36\ =\ 7,777,600$$

To establish the baseline harm from a Virus X pandemic, we make the simplifying assumption that the pandemic risk over the next decade is from a novel virus rather than an already-known pathogen such as Nipah. Accordingly, the baseline harm from a Virus X pandemic over the next decade is 7.8 million.

2.2.4.2 Virus Discovery (VDi)

We define VDi as the scenario where three times as many viruses are discovered in high-risk hotspots as we know today. Through our model, we aim to answer the question: "How would the risk of a Virus X natural pandemic event decrease if we discovered and sequenced three times as many viruses as we have today through current discovery and monitoring efforts?"

Our model considers two pathways in which VDi could cause downstream changes to reduce the risk of a Virus X pandemic. The first path is through the influence of pathogen-agnostic or broad-spectrum non-pharmaceutical interventions (NPIs), where the Virus X hotspot(s) might be prioritized for some spillover prevention interventions noted in table 1. The second pathway is through influencing the development of a pan-genus or pan-family broad-spectrum vaccine (BSV) that otherwise would either not have been developed and approved, or would not have worked against Virus X. For each pathway, we consider both the probability that VDi will influence the efficacy of the intervention, and the change in efficacy of the interventions themselves.

We consider a few key parameters to quantify the reduction in risk from VDi. First, we consider the change in likelihood that BSV will be effective against Virus X due to VDi, $\Delta p_{BSV|VDi}$. We also consider the magnitude of the reduction in harm a broad-spectrum vaccine would provide if available at the start of the outbreak, m_{BSV}.

For non-pharmaceuticals, we note the key parameters as the difference in likelihood of prioritized non-pharmaceutical interventions due to VDi efforts, $\Delta p_{NPI|VDi}$ and the reduction in harm from prioritized non-pharmaceutical interventions due to VDi, $\Delta r_{NPI|VDi}$. Using these parameters, we define the reduction in pandemic risk from Virus X due to VDi as follows (1):

$$E\left[Harm_{VDi}\right] = E\left[Harm_{base}\right] \times (1 - \Delta p_{BSV\,|\,VDi} \times m_{BSV}) \times (1 - \Delta p_{NPI|\,VDi} \times \Delta r_{NPI\,|\,VDi})$$

(1)

$$E\left[Benefits_{VDi}\right] = E[Harm_{base}] - E[Harm_{VDi}]$$

Due to uncertainty over the effectiveness of NPIs and levels of future investment, we make the simplifying assumption that the probability of NPI efficacy against Virus X given VDi, $\Delta p_{NPI|VDi}$ is equal to the required relative increased probability that a BSV is developed for Virus X due to VDi. $\Delta p_{BSV|VDi}$. For example, increasing the probability of a broad-spectrum vaccine from 0.37 to 0.46 requires a 14% increase in the overall likelihood of success: (0.46 – 0.37) / (1 – 0.37). We use the survey

to estimate $\Delta p_{BSV|VDi}$ and $\Delta r_{NPI|VDi}$, and estimate the value of m_{BSV}, extrapolating data based on a 2023 study by Więcek et al. evaluating the potential benefits if COVID-19 vaccines had been available earlier in the outbreak (see section 2.2.3.2 for derivation).

Uncertainty Quantification

The calculations above use the mean as point estimates for parameters based on survey data, though there is a large amount of uncertainty amongst experts reflected in the wide distributions of the various questions. To account for uncertainties in our parameter estimates, we conducted Monte Carlo simulations using Python with the NumPy and SciPy libraries. We performed 100,000 iterations for each analysis, drawing parameter values directly from survey data and from existing literature review. For each iteration, we calculated the harm reduction and deaths averted using our model equations, generating distributions of possible outcomes. From these distributions, we computed means, medians, and 90% central ranges to characterize the central tendencies and uncertainties in our results.

2.2.4.3 Pandemic Virus Identification (PVI)

To evaluate the benefits of pandemic virus identification, we evaluated the question "How would the risk of a natural Virus X pandemic decrease if we identified Virus X as a pandemic capable virus prior to spillover?". For the purposes of this model, we make a few simplifying assumptions:

1) There are a set number of *pandemic capable* viruses circulating around the world
2) Pandemic virus identification efforts conducted for a given pathogen will reveal with whether the virus is pandemic capable through characterization experiments estimating the virus's virulence and transmissibility in humans

The primary proposed benefits of PVI that are not possible only with viral discovery include targeted pathogen specific medical countermeasures, and targeted non-pharmaceutical interventions. Knowledge that *Virus X* is a pandemic-capable virus may result in prioritization of *Virus X* spillover prevention and mitigation efforts, leading to greater allocation of resources towards this specific threat, and an earlier start to development of vaccines and therapeutics.

Pandemic virus identification relies on the characterization of individual pandemic viruses, so we first evaluated the pandemic risk posed by an individual pandemic virus. Through our survey, we estimate there are on average 65 pandemic-capable viruses that are equally likely to seed a pandemic event. It should be noted that to date, no viruses identified as potential pandemic pathogens without spilling

over into humans have resulted in the development of targeted interventions. This model starts with the assumption that PVI has successfully characterized a novel zoonotic pandemic-capable virus before spillover, and first estimates the benefits of PVI per successfully identified pandemic virus. We first establish the probability PVI has successfully characterized Virus X, rather than a different pandemic-capable virus, as:

$$p(PVI) = \frac{n}{v} = \frac{n}{65}$$

where n is the number of viruses characterized and v is the number of pandemic-capable viruses that are equally likely to seed a pandemic event.

Based on the key potential benefits noted in literature about PVI, we chose the following parameters to quantify the additional reduction in risk: the likelihood a targeted MCM is funded following identification, p_{TMCM}; reduction in harm through earlier release of targeted vaccines due to PVI, Δm_{TV}; earlier release of targeted therapeutics due to PVI, Δm_{TT}, the likelihood PVI results in changes to threat-agnostic interventions and the relative reduction in pandemic risk from PVI informed non-pharmaceutical interventions $\Delta r_{NPI|PVI}$. We define the expected harm in a scenario with PVI as follows:

$$E\left[Harm_{PVI}\right] = p(PVI) \times E\left[Harm_{base}\right] \times (1 - p_{TMCM}(\Delta m_{TV} + \Delta m_{TT})) \times (1 - p_{NPI|PVI} \times \Delta r_{NPI|PVI})$$
(2)

For the parameters above in equation (2), p_{MCM} was estimated through the survey, where participants were asked to estimate the likelihood discovering Virus X would lead to sufficient funding being pooled to develop targeted medical countermeasures. To generate estimates for Δm_{TV} and Δm_{TT}, we estimated both how much earlier targeted vaccines and antivirals would be released, as well as the efficacy of the MCM itself. To estimate the shortened timeline, we asked survey participants how much earlier they anticipate a targeted vaccine and a targeted therapeutic would be released if PVI identified Virus X as a pandemic-capable virus, providing an estimate of the number of days. We also used the efficacies and distribution timelines of COVID-19 vaccines and antivirals following emergency use authorization and additional modeling literature to estimate the potential additional lives saved due to earlier release of these interventions in this scenario. For the case of a single pandemic virus successfully identified, we estimate this could result in approximately 49,000 lives saved in expectation.

Scaling Up PVI

To estimate the benefits of the entire PVI enterprise, we considered how the benefits scale for each additional identified virus. In our followup survey, we asked participants about the likelihood of sufficient funding being pooled if multiple pandemic viruses were identified. First using the median estimates of participant answers, we plotted the likelihood of acquiring sufficient funds against the number of viruses successfully identified (Fig.4)
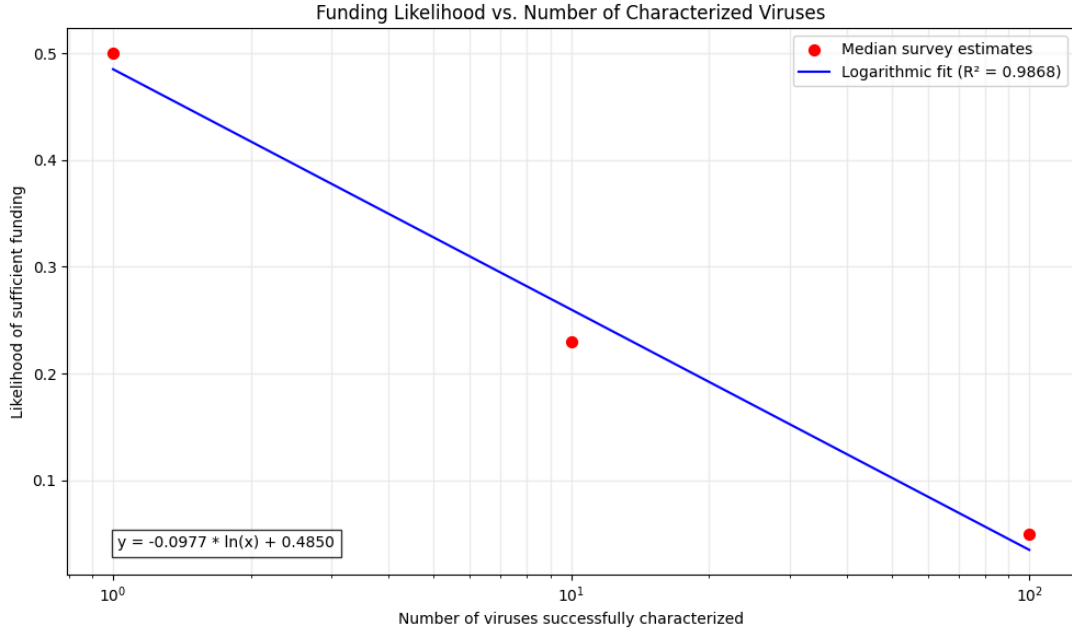
Figure 4: Relationship between the number of successfully identified pandemic-capable viruses and the likelihood of sufficient funding for countermeasure development. We fit a logarithmic curve to the data, y = -0.0977ln(x) + 0.485, where y is the likelihood of funding and x is the number of viruses.

In this scaling up version of the model, we assume the probabilities of funding for medical countermeasures and targeted non-pharmaceutical interventions are equal, such that there is a general probability targeted countermeasures will be funded p_{TCM}, where $p_{TCM} = p_{TMCM} = p_{NPI|PVI}$. We use the logarithmic curve generated by the survey data to generate estimates for p_{TCM} in the case where multiple pathogens are identified, such that:

$$p_{TCM} = -0.0977\, ln(n) + 0.485$$

We plug this into equation (2), and plot the expected harm in scenarios where PVI and multiple pathogens are identified.

$$E\left[Harm_{PVI\mid n}\right] = \frac{n}{65} \times 7,776,000 \times (1 - p_{TCM}(0.594)) \times (1 - p_{TCM} \times 0.517)$$

Uncertainty Quantification

Similar to the approach outlined in the VDi model, we assessed the uncertainty in the estimates for the benefits of identifying a single pandemic-capable virus and identifying all 65 pandemic viruses. For the case of a single pandemic virus, we ran 100,000 MC simulations drawing parameter values directly from survey estimates for p_{TMCM}, $p_{NPI\mid PVI}$ and $\Delta r_{NPI\mid PVI}$ parameters.

2.3 Results

Our model evaluates the benefits of VDi and PVI in preventing or mitigating a pandemic caused by *Virus X*, the currently unknown zoonotic virus which will otherwise cause the next pandemic to kill at least a million people. In this chapter, we typically report results using the following format: X [Y, Z]. Here X is the central estimate, and [Y, Z] represents the 90% central range, where Y is the 5th percentile value and Z is the 95th percentile value.

2.3.1 Survey of Experts

The dearth of quantitative data concerning key parameters lead us to conduct a survey of experts who have published in the fields of One Health and/or vaccine and therapeutic development (Methods). Key questions included the likelihood of a broad-spectrum vaccine effective against Virus X with and without virus discovery, the likelihood of funding for and probable acceleration of targeted vaccines and therapeutics given pandemic virus identification, and efficacies and likelihoods associated with non-pharmaceutical interventions were asked (Table 3).

Table 3: Survey questions and results from the initial survey and follow-up survey

Question	Mean [central 90% range]
Initial Survey (n = 207)	
VDi: effects on the likelihood of a broad-spectrum vaccine or therapeutic vs Virus X	
How likely are we to have at least one approved broad-spectrum vaccine or therapeutic 10 years from now that will be effective against the next high-consequence pathogen (>1 million deaths) given our current knowledge of the global virome?	0.37 [0.05 , 0.8]

How likely are we to have at least one approved broad-spectrum vaccine or therapeutic...if we discovered and sequenced 3.0x as many viruses as today?	0.46 [0.06, 0.87]
How likely are we to have at least one approved broad-spectrum vaccine or therapeutic... if we discovered and sequenced all viruses in animals?	0.55 [0.08, 0.95]
VDi: effects on preventing a Virus X pandemic	
If we sequenced 3.0x as many viruses as today, what is the relative reduction in pandemic risk from Virus X over the next 10 years relative to a world with no additional virus discovery?	0.38 [0.03, 0.86]
If we discovered and sequenced all viruses in animals, what is the relative reduction in pandemic risk from Virus X over the next 10 years relative to a world with no additional virus discovery?	0.52 [0.05, 0.96]
PVI: likelihood of characterizing Virus X	
How many distinct viruses capable of sustained human-to-human transmission, with the potential to cause at least 1 million deaths, do you estimate are currently circulating in animal reservoirs around the world?	172 [155,189]
PVI of Virus X: effects on the availability of vaccines and therapeutics targeting Virus X	
If Virus X is characterized as pandemic-capable prior to spillover through PVI, what is the probability the world will invest enough funds to develop targeted MCMs before an outbreak begins?	0.35 [0.02, 0.85]
___ days saved in Virus X vaccine availability (1B+ people) if pre-outbreak characterization occurs:	198 days [180, 200]
___ days saved in Virus X therapeutics availability (1B+ people) if pre-outbreak characterization occurs:	300 days [250, 360]
PVI of Virus X: effects on preventing a Virus X pandemic	
How much would non-medical countermeasures targeting a characterized Virus X reduce the risk of a sustained outbreak over 10 years, relative to a world where Virus X is not characterized?	0.52 [0.09, 0.94]
Follow-up Survey (n = 42)	
PVI: extrapolating effects to multiple viruses	

How is pandemic risk distributed across potential high-consequence pathogens?	The top 20% most risky pathogens contribute to ~ __% of expected mortality 70% [25%, 95%]
If ten different viruses are identified in the laboratory as suspected pandemic threats over the next ten years, what is the likelihood that the world invests enough funds to develop targeted countermeasures against all ten viruses?	0.32 [0.01, 0.81]
If one hundred different viruses are identified in the laboratory as suspected pandemic threats over the next ten years, what is the likelihood that the world invests enough funds to develop targeted countermeasures against all one hundred viruses?	0.18 [0.0, 0.55]

2.3.2 Virus Discovery Benefits

The general lack of international investment in pandemic preparedness even after COVID-19 strongly suggests that viral discovery alone is unlikely to ring alarm bells loudly enough to unlock more funding. Therefore, the primary benefits of this research will accrue from 1) ensuring that future broad-spectrum countermeasures would be effective against Virus X, and potentially 2) improved targeting of existing anti-spillover efforts. Hereafter, we use "VDi" to refer to a 3-fold increase in funding for virus discovery.

Survey participants estimated that the VDi scenario could reduce spillover risk by 38% ($\Delta r_{NPI|VDi}$ = 0.38 [0.03, 0.86]) if optimally translated to guide non-pharmaceutical interventions, as current spillover prevention efforts do not take virus density or diversity into account[71]. Achieving the full effect would require that budgets for non-pharmaceutical countermeasures undergo re-allocation based on the findings of the VDi research. Given that re-allocating cross-border funds and updating interventions based on virus discovery is presumably more challenging than for laboratories developing broad-spectrum vaccines to make use of the information, we used the estimated relative increase in likelihood of developing an approved broad-spectrum vaccine as an upper bound ($\Delta p_{NPI|VDi}$ = 14%; see Methods for higher values).

Respondents assigned a 46% [6%, 90%] probability that a broad spectrum vaccine would be effective against Virus X with VDi ($p_{BSV|VDi}$ = 0.46), as opposed to a 37% [3%, 84%] likelihood of a

broad-spectrum vaccine against Virus X without VDi ($p_{BSV|base}$ = 0.37). For the change in likelihood of a broad spectrum vaccine due to VDi, we took the difference ($\Delta p_{BSV|VDi}$ = 9% [-13%, 29%]).

To estimate the reduction in harm provided by an immediately available broad-spectrum vaccine against Virus X (m_{BSV}), we calculated the additional deaths that would have been prevented if such a vaccine had already been developed and approved at the start of the COVID-19 pandemic. A 2023 study by Więcek et al. estimated the potential additional lives saved if a targeted COVID-19 vaccine was released earlier in the U.S. and U.K.[66]. Because the Covid-19 vaccines were unusually effective at lowering the risk of death, broad-spectrum vaccines are not expected to be as effective. We consequently assumed that a BSV against Virus X would be as impactful at preventing mortality as the seasonal flu vaccine, which is approximately half as effective as the targeted Covid-19 vaccines. We therefore extrapolated the results Więcek et al. to estimate the additional reduction in global mortality we would see if a broad-spectrum vaccine for COVID-19 had been available 300 days prior to the approval and release of the targeted vaccines, occurring approximately 50 days into the outbreak. Importantly, we assumed that targeted vaccine development would have proceeded as normal, with individuals who received an early broad-spectrum vaccine in our counterfactual being vaccinated with the targeted vaccines once available.

Our point estimate suggests that VDi could save approximately 492,000 [0, 1.46 million] lives over the next decade. To account for uncertainty in parameter estimates, we conducted Monte Carlo simulations drawing parameter values directly from survey data or distributions derived from the data. These simulations resulted in a median harm reduction of approximately 9%, with a 90% central range of 0% to 19%. This wide interval and difference between the mean and median underscores the lack of consensus among experts surrounding potential impacts.

Scenario Analysis

In the initial model, we make the assumption that the change in likelihood of NPIs being targeted towards preventing Virus X, $\Delta p_{NPI|VDi}$; is equivalent to the required relative increased probability of a vaccine being approved for use against Virus X, $\Delta p_{BSV|VDi}$. Here, we relax that assumption and consider scenarios where they are not equivalent, and run MC simulations setting $\Delta p_{NPI|VDi}$ to 30% and 50%, running 100,000 simulations at each level (Table 4).

Table 4: Results of MC Simulations at Various Levels of $\Delta p_{NPI|VDi}$ (likelihood of NPIs being funded through VDi)

$\Delta p_{NPI|VDi}$	Lives Saved
30%	x̄: 1,129,123 M: 1,075,452 90% Central Range: (69,884 ; 2,235,009)
50%	x̄: 1,695,476 M: 1,555,200 90% Central Range: (161,741 ; 3,341752)

2.3.3 Pandemic Virus Identification Benefits

Given that pandemic virus identification relies on the characterization of individual viruses, we first evaluated the benefits of successful identification of a single pandemic-capable virus. Survey participants estimated that there are a total of 172 pandemic-capable viruses in nature, with 70% of the risk concentrated in the top 20% of viruses, which is statistically equivalent to 65 pandemic-capable viruses that are equally likely to seed a pandemic event. We grouped the benefits of PVI into two categories: the benefits it would provide to the production of medical countermeasures through accelerating timelines, and the benefits of targeting NPIs toward regions at high risk of identified virus spillover.

Participants were asked to estimate how much earlier a Virus X vaccine and therapeutic would be available to at least 1 billion people if the virus had been characterized and flagged as a suspected pandemic risk before the outbreak. The median response indicated that *if* sufficient funding were acquired, a targeted vaccine would be available 198 [180,200] days sooner, and a therapeutic would be available 300 [250, 360] days sooner, compared to a scenario where Virus X had not been preemptively identified. Combining these results with COVID-19 medical countermeasure data, we estimated these accelerated timelines would save an extra 50% of lives through vaccination, and an additional 9% of lives saved through access to therapeutics. Participants reported an average response of 35% [2%, 85%] when asked to estimate the probability a targeted MCM would receive sufficient funding if Virus X were characterized, and expected successful PVI of Virus X to reduce spillover risk by an additional 52% [9%, 94%] through better-targeted NPIs. Using these parameter estimates, successfully identifying a single pandemic-capable virus could save approximately 49,000 lives over the next decade.

To account for uncertainty in key parameters, we conducted Monte Carlo simulations sampling from survey data to represent the parameters p_{TMCM}, $p_{NPI\,|\,PVI}$, and $\Delta r_{NPI\,|\,PVI}$. The simulations provided a mean benefit of approximately 49,000 lives saved per virus, with a 90% central range between 10,500 and 93,600 lives saved. The wide confidence interval reflects the lack of consensus in estimates regarding the likelihood of funding as well as the efficacy of potential interventions.

To estimate the benefits of the entire PVI enterprise, we considered how the benefits scale for each additional successfully identified virus. In our follow-up survey, experts estimated the likelihood of sufficient funding being pooled for targeted countermeasures if multiple pandemic viruses were identified. For this portion of the model, we assume $p_{TMCM} = p_{NPI\,|\,PVI}$, such that the likelihood of targeted medical countermeasures and targeted non-pharmaceuticals receiving sufficient funding are equivalent. Our model predicts that identifying all 65 pandemic-capable pathogens would save 642,000 lives in expectation, reducing overall natural pandemic risk by approximately 8% over the next decade.

Similar to the VDi scenario, Monte Carlo simulations revealed significant uncertainty within these estimates, with a 90% central range of 4,600 lives to 5.1 million lives saved if all pandemic-capable viruses in nature were successfully identified (Fig. 3).

Estimated Lives Saved with Varying Levels of PVI
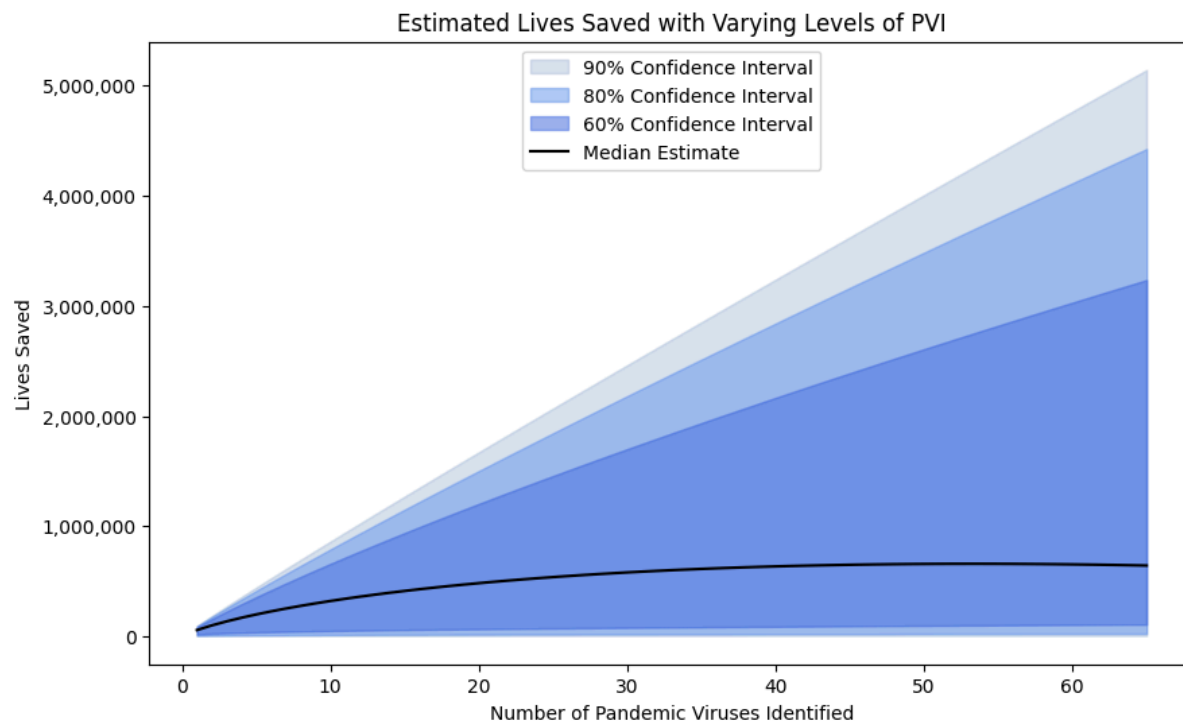
Figure 3: Estimated number of lives saved based on number of pandemic-capable viruses identified. The black line represents the model outputs using the median p_{TMCM} (and $p_{NPI|PVI}$) estimate, and the confidence bands representing model outputs using the 60%, 80% and 90% central ranges of the parameter.

Comparison of VDi and PVI

Comparing VDi and PVI, we see that a three-fold increase in VDi is expected to save approximately 492,000 [0, 1.46 million], while PVI is expected to save 642,000 lives if all 65 pandemic-capable viruses in nature are identified through characterization. Notably, this level of PVI would require a greater investment in VDi than the threefold increase we evaluate, as only a discovered virus can be characterized. It should be noted that these results are heavily dependent upon parameters from the expert surveys, which exhibited considerable variance.
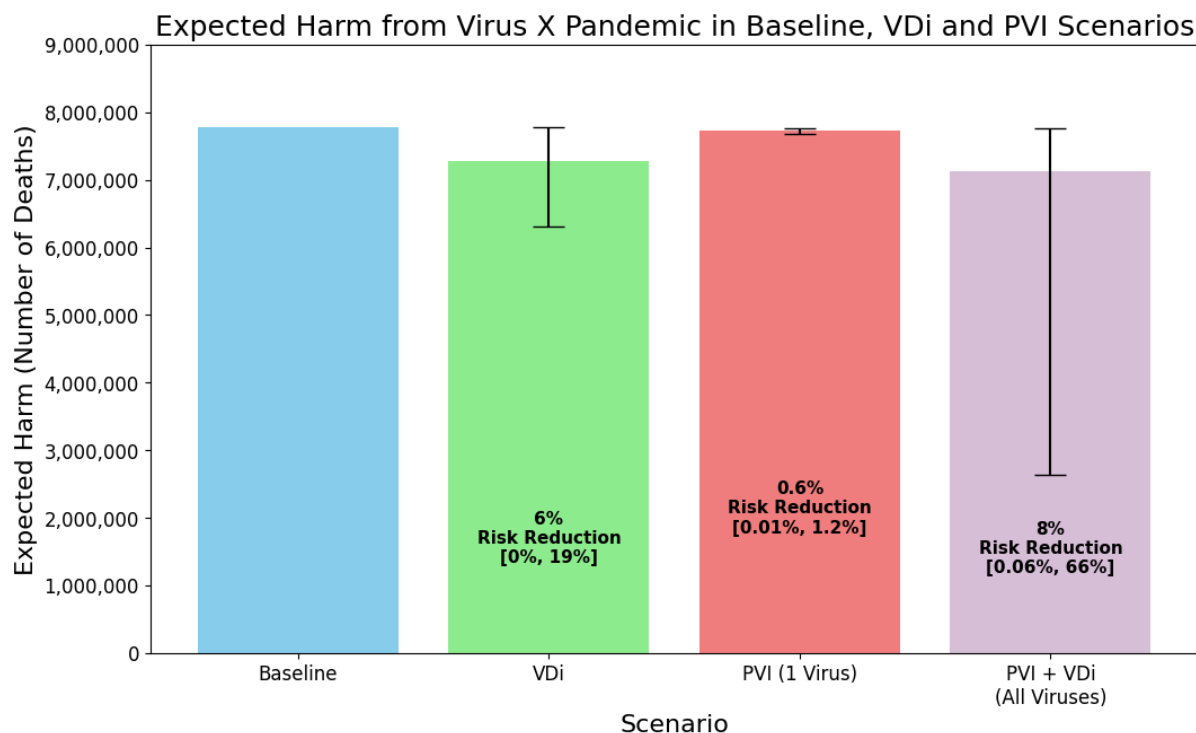
Figure 4: Comparison of expected harm from a Virus X pandemic over the next decade in the baseline scenario, VDi scenario where 3 times as many viruses are discovered and sequenced today, the PVI scenario where 1 pandemic-capable viruses is identified, and the PVI scenario where all pandemic-capable viruses are identified.

2.4 Discussion

The findings from this mathematical model and expert survey provide insights into the potential benefits of viral discovery efforts and pandemic virus identification (PVI) research in mitigating future

pandemic risks. Our results suggest tripling the amount of viral discovery and monitoring efforts at the animal-human interface could save an estimated median of 492,000 [0, 1.46 million] lives from a Virus X pandemic in the next decade. These benefits primarily arise from the hypothetical ability to better target general non-pharmaceutical interventions to regions with many discovered viruses, with a secondary contribution from the increased likelihood that broad spectrum pharmaceuticals would be effective against an emerging Virus X pandemic.

Our model estimates that pandemic virus identification could save 49,000 [10,500; 93,600] lives per virus identified in the next decade, scaling to 642,000 [4600, 5.1M] lives saved per decade given discovery and characterization of all 172 [155, 189] pandemic-capable viruses in nature, primarily from accelerated medical countermeasures.

There has been considerable debate surrounding the benefits of viral discovery efforts, with some arguments noting that these efforts are costly, and while they produce a wealth of data regarding the virome, the tangible benefits to mitigating disease emergence is small or extremely unclear [62]. A few common themes of criticism around VDi include the high cost and inefficiency, the lack of clear practical uses of this data, and the biosecurity risks. Arguments in favor of this research note developing a reference library can facilitate rapid identification and risk assessment, as well as facilitate testing of vaccines and therapies, and development of broadly targeted interventions [72]. While this lack of consensus was observed across survey participants, we observed slight differences in estimates between groups surveyed. Participants working in virus discovery and characterization programs generally provided more optimistic estimates compared to other groups (national academies, basic research and applied research cohorts). This group estimated a higher probability (58%) of having a broad-spectrum vaccine effective against Virus X without additional discovery efforts compared to the estimates from other groups (44 to 46%), and estimated a larger increase (13%) in the likelihood of an approved broad-spectrum vaccine due to viral discovery efforts compared to other groups (ranging from 8 to 10%). They also provided more optimistic estimates for the VDi's influence on risk reduction through non-pharmaceutical interventions.

Similar to VDi, the benefits of PVI are also frequently contested and debated, with some overlap in the arguments in favor and against the research efforts. Characterization work to evaluate the pandemic potential of viruses is a costly and intensive process, often requiring samples to be shipped to countries with specialized capacities. Arguments in favor of PVI research note that knowing which viruses have high pandemic potential is valuable for the development of targeted spillover prevention and earlier medical countermeasure development. Those opposed similarly point to the lack of a clear mechanism

or evidence of the practical benefits of this work, as well as the biosafety and biosecurity risks of handling potentially pandemic-capable viruses and the publication of dual-use information [72]. With regards to parameters associated with PVI, experts generally expressed hesitancy surrounding funding of targeted medical interventions following the identification of a virus with pandemic potential, with a mean response of 35% and median response of 30% when asked about the likelihood of discovering Virus X leading to sufficient funding being pooled to develop targeted medical countermeasures. These results are reflected in the funding emerging infectious disease has generally received, with the case of Nipah serving as a clear example. The Nipah virus was discovered in 1998, and since then sustained human-to-human transmission chains and outbreaks have been frequently observed in Southeast Asia [73]. Despite these observations as well as evaluations demonstrating the pandemic potential of the virus, very little funding and research effort has been dedicated towards countermeasures for the virus, including a lack of an established vaccine or therapeutic [74]. Given this lack of response for a known threat that has been demonstrated characteristics of a pandemic pathogen not only in animal models but in human populations, it is unlikely a novel threat characterized through PVI would result in a substantial increase in funding and resources for prevention interventions.

This uncertainty may reflect the speculative nature of many potential outcomes evaluated by survey participants. For example, government funding to support a vaccine for a virus that has not yet infected a human would be a priori surprising given that Nipah virus – a potential pandemic pathogen which has caused multiple outbreaks exhibiting sustained human-to-human transmission since 1998 [73] – still has no approved medical countermeasures and only recently merited candidates in clinical trials [74].

Collectively, our results indicate that despite the high uncertainty in key parameters obtained from expert surveys, quantitative modeling of the expected benefits of proposed public health research programs can provide outer-bound estimates relevant to deciding how best to allocate scarce resources.

Chapter 3: Accidental Pandemic Risks

3.1 Background

One of the primary concerns surrounding work with potential pandemic pathogens (PPPs) is the possibility of a containment failure that leads to an accidental pandemic event. Risks associated with this research include inadvertent infection of personnel during field work, laboratory procedures, or disposal and decontamination processes. There is also the risk of direct release of pathogens into the environment through containment failures or improper sterilization of effluents such as water or air [75–77].

3.1.1 How PVI Contributes to Accidental Pandemic Risk

PVI research involves handling and working with potential pandemic-capable viruses, posing various biosafety risks that can lead to accidental infection. This research generates various solid and hazardous biological wastes that personnel handling waste management may be in direct contact with during disposal. Hazardous air and water pollutants can lead to indirect community exposure if emissions are not properly filtered, as noted in the case of the 1979 anthrax leak in Sverdlovsk, where a failure to replace a clogged air filter in an exhaust vent resulted in the accidental release of *B. anthracis* spores across the city [78]. For transmissibility experiments, healthy animals are infected with the PPP, which carry their own biosafety risks. Infected animals can infect researchers during experiments, and require an incinerator to properly dispose of them following experiments. If these are improperly handled or stored, those working with them or disposing of the waste from these experiments may lead to inadvertent infection and onward transmission. Finally, extreme weather events could affect critical biosafety infrastructure that may lead to accidental release of pathogens [37]. Infections with PPPs may also only become contagious before symptoms appear or be completely asymptomatic, such that a case can go undetected for several days before the infected person is isolated.

3.1.2 Prior Literature

Laboratory acquired infections (LAI) in BSL-3 and BSL-4 facilities have been recorded for several pathogenic viruses, occurring through direct contact with animal vectors, self-inoculation, inhalation and ingestion [79]. Notable examples include the series of smallpox infectious smallpox infections in the UK in 1978 traced back to a laboratory at the University of Birmingham [80], the infection of a lab worker with SARS coronavirus in Singapore in 2003 [81], and an accidental dengue infection due to a mosquito bite in a research lab, despite the researcher donning proper PPE [82]. Other examples include

SARS-CoV-1 samples escaping containment on four separate occasions from laboratories while being stored, and multiple incidences of CDC staff accidentally exposed to dangerous pathogens such as live active anthrax spores, smallpox, and pathogenic H5N1 [83,84]. In their 2021 study, Manheim and Lewis documented 71 high-risk human-caused pathogen exposure events from 1975-2016, including both accidental and purposeful exposures to highly infectious agents [85]. Notable events include the multiple foot-and-mouth disease outbreaks linked to research facilities, and several cases of researchers being infected with Ebola, plague, and other dangerous pathogens. The paper estimated the actual number of incidents is likely much higher, as many events go unreported or undisclosed. Additionally, The American Biological Safety Association established a database of LAIs in 2016, recording all international incidents of LAI events recorded in peer-reviewed and published journal articles.[86].

Multiple studies have been carried out quantitatively estimating the likelihood of LAI events and secondary transmission events in research facilities [6,87,88]. In large part, this has been possible due to surveys and records of LAIs recorded by various agencies and groups. In the U.S., the Federal Select Agents Program (FSAP) oversees possession, use and transfer of agents on the Biological Select Agents and Toxins (BSAT) list, which is a list of biological agents that pose a severe threat to human, animal or plant health. In 2012, the FSAP also regulated nucleic acids related to BSAT, such as genomes that encoded infectious forms of viruses on the select agent list. Using publicly available data, Henkel et al. noted that between 2004 and 2010, there were 11 recorded LAIs with BSAT agents, where approximately 10,000 people had clearance to access these agents [89]. For BSL-3 labs in particular, they note there have been 4 LAIs over 2044 lab-years, suggesting a 0.2% likelihood of an LAI per BSL-3 lab-year. To obtain full incident reports regarding PPPs, multiple Freedom of Information Act (FOIA) requests were submitted to the CDC and NIH's Office of Science Policy by Dr. Lynn Klotz from the Center for Arms Control gathering information about lab accidents under the FSAP program, and lab accidents in BSL-3 and BSL-4 facilities in the U.S. The FSAP data revealed 10 to 14 undetected or unreported LAIs over 4,067 lab-years from 2003 to 2017, with 10 of the LAIs clearly undetected or unreported, while the status of the remaining 4 LAIs was uncertain regarding if they were detected. The NIH data revealed that between 2004 and 2017, 13 undetected or unreported LAIs were recorded in BSL-3 and BSL-4 labs over 458.3 lab-years. In both datasets,majority of incidents are attributed to human error, with Klotz noting 73% of NIH accidents and 79% of FSAP accidents were attributed to errors such as needlestick injuries (skin is accidentally punctured by a used needle), skin exposures, dropped objects, and improper animal handling leading to a bite or scratch from an infected animal. Looking at data from the NIH's National Institutes of Allergy and Infectious Diseases (NIAID), Lipsitch and Inglesbly noted between 1982 and 2003, NIAID recorded 3 LAIs over 634,500 person-hours of work, resulting in an average of 1 LAI per every 100 full-time person years of work at

approximately 2000 working hours a year [90]. This data has been used in numerous analyses to estimate the risk of accidental infections posed by working with various pathogens as well as the overall risk posed by the entire PPP research enterprise.

Several analyses have combined LAI data with modeling efforts around secondary transmission events to then estimate the likelihood of an accidental local or global outbreak scenario based on the characteristics of the pathogen and environmental considerations. Modeling work by Merler et al. in 2013 analyzed the likelihood of an accidental release event resulting in a global outbreak, focusing their efforts on novel transmissible flu strains released in densely populated areas. [91] They estimated there was a 5 to 15% chance of an LAI resulting in an undetected epidemic, suggesting a 5% likelihood given a 60% chance infected close contacts are detected, and 15% likelihood of the chance of close contact detection is 15%. They suggest containment using social distancing measures is likely to succeed if the $R_0 < 1.5$, but is inadequate for pathogens with greater reproduction numbers. In a 2014 study, Klotz and Sylvester estimated a 1-30% conditional probability of an LAI causing a pandemic event, based on an R_0 of between 1.4 and 3.0 [88]. In 2014, Lipsitch and Inglesby estimated that 1 lab-year of work on pandemic influenza would have an expected death toll of 2000 to 1.6 million people per BSL-3 lab year based on FSAP data, and estimate 8,000 and 10 million expected deaths per full-time worker-year in a BSL-3 lab using NIAID data listed above. They estimate the likelihood of a lab accident-induced pandemic through working with a novel, transmissible form of the flu virus between 0.01% to 0.1% per BSL-3 lab-year [90]. In their 2016 risk-benefit analysis of gain-of-function research, Gryphon Scientific estimated an accidental infection with a wild-type pandemic influenza strain could result in a global pandemic every 560 - 13,000 years, reporting a 5% likelihood of an pandemic flu LAI resulting in a local outbreak, and a 20% likelihood of an local outbreak leading to a global pandemic for a flu strain[92]. They additionally estimate the likelihood of loss of local control can increase to 30% if there is little residual immunity to the virus, and estimate this could rise to 50% for pathogens with an R_0 of 1.8 and no community mitigation measures.

While these estimates are less contested than assessments of the benefits of PVI, there has been pushback regarding the biosafety assumptions made in these estimates. Lipsitch and Inglesby's estimate was critiqued by Dr. Ron Fouchier, one of the two virologists who conducted GoF experiments on H5N1 strains in 2011. This resulted in a back and forth discussion of competing risk estimates due to disagreements in methodology, the evidence used to derive estimates for lab-acquired infections, and the types of labs PPP research was being conducted in. Fouchier estimated the likelihood of an LAI with onward transmission to be between 2.5×10^{-13} and 3×10^{-12}, occuring once every 33 billion years due to his estimate of an extremely low likelihood of a lab-acquired infection [93].

3.1.1 Research Question

Using the prior literature above, we aim to answer the following question: "What is the risk of a lab-accident pandemic over the next 10 years due to PVI research efforts?". Returning to our definition of PVI, we specifically aim to estimate the risk associated with viral characterization experiments, assessing how laboratory experiments carried out to assess the virulence and transmissibility of a virus contribute to biosafety risks.

3.2 Accident Risks Model Structure

3.2.1 Model Assumptions

Characteristics of Pathogen and Scenario

We assume the virus being characterized is a highly transmissible pandemic capable virus with an R_0 of approximately 1.5, and assume there is no to very little pre-existing population immunity as it is a novel virus that has not spilled over into human populations. Given over 75% of the 60 BSL-4 labs around the world are located in an urban center, we assume the lab(s) conducting this research are located in an urban center, and is taking place in BSL-3 or BSL-4 facilities [94]. For this model, we assume the risks from PVI research solely come from laboratory accidents that occur during characterization experiments, and do not include accidents that may arise during transport of samples to and from the facility, or accidents during the field sampling (e.g. accidental exposure and infection during animal handling or sample transport to the labs)

Duration of Experiments

To make use of the lab-accident data and statistics described above, we first need to estimate the time it takes to characterize a single virus. Below, we list the set of experiments associated with characterization of a virus's pandemic potential, and state our assumptions for the number of lab-years needed to carry them out. We assume there are a team of approximately 4-5 researchers carrying out these experiments in a lab.

PVI Experiments

1. Reconstruction and passaging of the virus (3-4 months = 0.25 - 0.33 lab-years)
   - Generate infectious viruses from cloned cDNA using reverse genetics
   - Sequence to confirm identity and assess viability of reconstructed virus

2. In vitro viral replication and pathogenicity assays (2 months = 0.17 lab-years)

- Assess viral growth kinetics in relevant cell lines (e.g., MDCK, Vero) and primary human epithelial cells
- Perform multi-step growth curves, measuring viral titers at various time points
- Evaluate cytopathic effects and plaque morphology to assess virulence

3. Receptor binding assays (1 month = 0.08 lab-years)
   - Determine receptor specificity using glycan or hemagglutination assays
   - Assess binding affinity to human-type receptors

4. Antiviral sensitivity testing (1 month = 0.08 lab-years)
   - Evaluate susceptibility to current antiviral drugs and determine efficacy of existing therapeutic antibodies

5. Animal infection studies (3 months)
   - Assess viral replication, pathogenicity, and tissue tropism in animal models
   - Evaluate morbidity, mortality and immune responses such as antibody production

6. Transmission studies in animal models (3 months)
   - Determine efficacy of direct contact and aerosol transmission between animals (e.g. ferrets, transgenic mice, hamsters)

7. Testing mutants and reassortants (4 months)
   - Generate viruses with specific mutations or reassortant viruses with gene segments from different strains
   - Compare results with wild-type virus to identify genetic determinants of pandemic potential

Based on these assumptions, we estimate it takes approximately 1.5 lab-years to carry out all the necessary experiments to characterize the pandemic potential of a virus.

3.2.2 Model Structure

This model evaluates the expected harm over a decade due to PVI research. We define this harm as follows (3):

$$E[Harm_{accident}] \ = \ p(LAI) \ \times \ p(pan\,|\,LAI) \ \times \ E[Harm\,|\,Pandemic] \qquad (3)$$

Where $E[Harm_{accident}]$ is the overall expected death toll due to PVI research, $p(LAI)$ is the probability of at least one LAI occurring during characterization experiments, $p(pan \mid LAI)$ is the conditional probability of a pandemic given an LAI has occurred, and $E[Harm \mid Pandemic]$ is the conditional expected number of deaths given a pandemic event does occur. We use the Fan et al. estimate from the benefits model for $E[Harm \mid Pandemic]$, estimating approximately 21.6M deaths for a given pandemic event.

To estimate the likelihood of at least one LAI, we take into account the likelihood of an LAI per lab year, $p(LAI_{lab-year})$ and the number of lab years it takes to characterize the pandemic potential of viruses, LY. Accordingly, we estimate $p(LAI)$ using the following equation (4):

$$p(LAI) \;=\; 1 \;-\; (1 \;-\; p(LAI_{lab-year}))^{LY} \qquad (4)$$

Using these equations, we first estimate the risks associated with the characterisation of a single pandemic-capable virus, which we then scale linearly to assess the risks posed by multiple viruses.

To establish a value for $p(LAI_{lab-year})$, we turn to literature referenced in section 3.1.2. Using FSAP and NIH data reported by Klotz, the FSAP rates result in a lower bound of 0.246% and an upper bound of 0.344% likelihood of an undetected or unreported lab-acquired infection (uuLAI) per entity year based on FSAP data, while NIH data result in a likelihood of 2.84% per lab-year. Though there is not a clear reason for why these rates are so different, Klotz hypothesizes FSAP regulations are partially enforced by the FBI which may have more stringent biosafety measures. He also notes most influenza research is not conducted under FSAP but under NIH. For the purposes of this model, we will use 0.246% as a lower bound, and 2.84% as an upper bound for $p(LAI_{lab-year})$, taking the average of the range as our central estimate at 1.543%.

3.2.2.1 Expected Harm from a Single Virus

We use n to denote the number of viruses characterized, such that $E[Harm_{accident \mid n=1}]$ represents the expected harm from characterizing a single virus, and $p(LAI_{n=1})$ denotes the probability of at least one LAI per virus characterized. We calculate $p(LAI_{n=1})$ as follows, using (4):

$$
\begin{aligned}
p(LAI_{n=1}) &= 1 \;-\; (1 \;-\; p(LAI_{lab-year}))^{LY} \\
&= 1 \;-\; (1 \;-\; 0.01543)^{1.5}
\end{aligned}
$$

$$= 0.023$$

We estimate the probability of a LAI while characterizing a single pandemic virus, $p(LAI_{n=1})$, to be 2.3%, with a lower bound of 0.36% using FSAP data, and an upper bound of 4.2% using NIH data.

To estimate the likelihood of a LAI resulting in a pandemic event, *p(pan|LAI),* we similarly turn to estimates derived in prior literature. Using Gryphon's estimates of the likelihood of a local outbreak and onward global pandemic of 5% and 20%, this would result in an overall *p(pan|LAI)* estimate of 1.5%. Lipsitch et al. noting that for a SARS pathogen with an R_0 of 1.5, the probability of onward outbreak from a single infection would range from 10% to 40%, while Merler et al. estimated the likelihood of an influenza LAI leading to extensive spread to be at least 10%, with an overall range that an influenza LAI resulting in a pandemic to be between 5% and 15% [6]. We take the median of these values to derive a central estimate of a 10% likelihood, and use [5%, 40%] as the range for $p(pan \mid LAI)$. Substituting these values into (3), we estimate the overall expected harm from characterization of a single pandemic-capable virus as follows:

$$
\begin{aligned}
E[Harm_{accident \mid n=1}] &= p(LAI_{n=1}) \times p(pan \mid LAI) \times E[Harm \mid Pandemic] \\
&= 0.023 \times 0.1 \times 21,600,000 \\
&= 0.0023 \times 21,600,000 \\
&= 49,800
\end{aligned}
$$

We estimate there is a 0.23% [0.018%, 1.68%] likelihood of an accidental pandemic event from characterization of 1 pandemic-capable virus, resulting in 49680 [3983, 365438] deaths in expectation per virus characterized.

The calculations above assume 1.5 years per virus characterized. If we relax this assumption and use the same approach, the expected harm due to accident risks posed by the characterization of a single virus is estimated to be 16,729 deaths if it takes 0.5 lab-years, 33,329 deaths if it takes 1 lab-year, and 66143 deaths if it takes 2 lab-years.

3.2.2.2 Risks from Multiple Viruses

We assume the number of lab-years scale linearly, such that the total number of lab years to characterize *n* number of pandemic viruses, $LY = 1.5n$. Accordingly, using the formulas above we estimate the accident risks from characterizing multiple pandemic-viruses in Table 5.

Table 5: Estimated death toll from characterizing between 1 and 5 pandemic-capable viruses, reporting the central estimate, lower bound and upper bound.

Number of Pandemic-Viruses Characterized	$E[Harm_{accident\,|\,n}]$
1	49,800 [3983; 365438]
2	98,452 [7951; 715420]
3	145,982 [11,904; 1,050,599]
4	192,416 [15,483, 1,371,601]
5	237,779 [19,567, 1,679,026]

3.3 Summary of Findings

We estimate the characterization of a single pandemic virus results in an expected death toll of 49,800 [3,983; 365,438] lives over the next decade with the assumption it takes approximately 1.5 lab-years to characterize a virus. Lipsitch and Inglesby estimate that laboratory work with a novel influenza virus that is airborne transmissible poses an expected death toll of 2,000 to 1.4 million fatalities per BSL3-laboratory-year.

The wide ranges in parameter estimates from previous literature are due to differences in data sources (i.e. FSAP accident reports vs NIH accident reports), varying assumptions about the characteristics of the pathogen such as the R_0 value, and the presence of superspreaders that influence the likelihood of pandemic level transmission.

Chapter 4: Deliberate Misuse Risks

One of the central concerns surrounding research with potential pandemic pathogens (PPPs) are the biosecurity risks associated with the dual-use insights the research outputs and materials produced by this research would provide. While dual-use concerns have been raised in various areas of life sciences research, PPP research has been flagged as carrying significantly high dual-use potential. It often makes the genomes of highly dangerous viruses publicly available and provides detailed insights into their replication, transmission, and pathogenicity. PVI characterization experiments publish information such as the replication efficiency of the virus in primary human cell lines, transmission potential in animal models, pathogenicity, and the virus's sensitivity to existing vaccines, antivirals and other therapeutics [34–36]. Malicious actors with the intent and capability to cause large-scale harm could exploit this information to synthesize and disseminate pathogens, deliberately inducing pandemics and causing unprecedented global disruption [42].

4.1 How PVI Contributes to Bioterrorism Risks

The primary mechanism by which PVI can influence the risk of bioterrorism is through providing information that can aid actors interested in causing large scale harm through seeding a deliberate pandemic event. By its nature, PVI research is intended to uncover novel viruses with pandemic potential, revealing the existence of previously unknown dangerous viruses. It provides the whole genome sequence, results of characterization experiments that indicate what features of the virus make it particularly concerning, and reverse genetics protocols to facilitate reconstruction of the virus in other labs [95]. Bioweapons programs tested various agents through laboratory characterization experiments to identify the ideal candidate to use as a weapon, evaluating the virulence and transmissibility. Ken Alibek, the former Deputy Director of Biopreparat, noted the scientific director of one of the facilities gained inspiration for a chimeric viral BW from research conducted in the West where scientists had inserted a gene of the Venezuelan equine encephalitis (VEE) virus into the vaccinia virus to better understand the viral genome and develop vaccines [96]. Alibek notes this research inspired the director, observing these same techniques would be used to develop a more powerful smallpox weapon, as the vaccinia virus is almost identical to the variola major virus. Biopreparat also open-air tests of anthrax, smallpox and other agents on monkeys on the Vozrozhdeniya Island in the Aral Sea to evaluate its virulence and transmissibility [96]. Well intentioned research that carries out animal transmission experiments provides similar insights regarding the potential transmissibility of the pathogen between humans, using animal models of transgenic mice with the human ACE2 receptor to assess whether a pathogen would likely be transmissible between humans[34]. Providing these insights through PVI

research obviates the need for malicious actors to carry out their own tests to identify ideal candidates, lowering the bar for identifying promising BW agents.

In addition to facilitating malicious actor choose promising candidates, reverse genetics protocols outlining the steps for de novo reconstruction of functional viruses are also published with viral characterization work, helping researchers carry out molecular biology techniques such as cDNA fragment assembly and mammalian tissue culture to help legitimate researchers to study the virus of interest. These protocols also generally increase the number of actors capable of synthetically synthesizing viruses, including the number of malicious actors [97]. This was noted in the controversy surrounding the 2005 paper reconstructing the 1918 Spanish Flu, as publication of detailed methodologies, genetic sequences, or reverse genetics protocols for such pathogens could provide a blueprint for recreating and weaponizing them [98]. This information, combined with the results from characterization experiments. , could allow bad actors with the intent and capability to cause large-scale harm to synthesize and disseminate pathogens, deliberately inducing pandemics.

4.2 History of Bioterrorism with Infectious Agents

The use of biological agents to cause harm extends far back into human history, with infectious pathogens in particular recognized for their utility as biological weapons (BWs) as far back as 600 BC, far before our understanding of pathogenesis and germ theory [99]. The most notable example during this time was the Tartar army catapulting plague-infested corpses over walls during the siege of Caffa, resulting in a plague epidemic in the besieged region [100], though other various other incidences have been recorded where biowarfare using infected human and animal carcasses was either carried out or attempted during wars [101]. Despite the passing of the Geneva protocol in 1925 which prohibited the use of biological and chemical weapons, several countries violated the conventions and continued or established bioweapons programs. Japan's Unit 731 carried out intensive and tortuous bioweapons research before and during World War two, conducted horrifying experiments and attacks using anthrax, plague, and other deadly pathogens on Chinese civilians and prisoners of war.

In the 20th century, BWs became more sophisticated due to advances in microbiology, resulting in the stockpiling and use of *B. anthracis*, smallpox virus, *C. botulinum*, *Y. pestis*, *F. tularensis,* and other agents during the World Wars and Cold War by several countries [102]. The largest and most established was the Biopreparat BW program, employing between 30,000 to 40,000 people, and stockpiling approximately 300 tonnes of weaponized anthrax spores alongside other natural and engineered strains of bacterial and viral pathogens. Biopreparat used the guise of legitimate medical or agricultural research to not only

mislead the public, but their own researchers through operating with high levels of secrecy and compartmentalization. Civilian scientists might have believed they were working on improving vaccine production for biodefense or studying pathogens for disease control, not realizing that their work was being used to weaponize these pathogens [96,103]. As The USSR was a signatory of the 1972 Biological Weapons Convention (BWC), Biopreparat covertly operated in violation of the treaty for 20 years, maintaining high levels of secrecy throughout its operation. Around the late 1970s, Iraq also violated the BWC and began development of its BW program under Saddam Hussein's dictatorship, conducting many of its operations under the guise of well intentioned medical research or food production [104].

In addition to state BW programs, several non-state actors including lone individuals and extremist groups expressed interest in the use of bioweapons in the late 20th and early 21st century. The most notable was the apocalyptic cult Aum Shinrikyo, who aimed to wipe out humanity through the use of BWs between 1987 and 1995 [105] This program was run by Seiichi Endo who was a molecular biologist training in virus and genetic engineering. With a team of between 3 and 20 people, they explored and tested the use of *C. botulinum*, *C. burnetti*, and Ebola virus, finally settling on anthrax spores as their weapon of choice. They were unsuccessful in their BW attempts due to mistakenly using a non-lethal strain, but successfully carried out a chemical weapons attack using sarin gas in the Tokyo subway station, killing 13 people. Due to the destruction of most evidence and execution and deaths of key members of the BW development program, it is hard to assess why it failed as it was a well-financed operation with dedicated workers. Some hypothesize the failures were due to the technical challenges of working with and disseminating biological agents, Endo's lack of experience with bacteria compared to viruses, and organizational failures. Other groups such as Daesh/ISIS, the Gaia Liberation Front and R.I.S.E have expressed interest in the use of BWs to kill most or all humans on Earth [106].

4.3 Current and Emerging Synthetic Biology Capabilities

In addition to intent, a notable current barrier to bioterrorism is the specialized capabilities required to assemble BWs [107]. Even with access to reverse genetics protocols, an actor would still need to be able to carry out complex molecular biology techniques, have access to specialized equipment and reagents, and possess the tacit knowledge to troubleshoot inevitable experimental failures. However, this is subject to change with emerging technologies that make it easier to perform reverse genetics. Improvements in gene assembly techniques, such as Gibson Assembly and Golden Gate cloning, simplify the process of constructing large DNA fragments (Ellis et al. 2011). Laboratory automation such as that offered by liquid handling robots and automated research facilities such as cloud labs make complex biological experiments more accessible to non-experts [108].

Chapter 5: Pandemic Bioterrorism Risk Assessment

Despite the qualitative discussions around the changing risk landscape of pandemic bioterrorism, there is a lack of public literature quantitatively estimating the risk of such events. These probabilities are particularly challenging to estimate given there has yet to be a case of successful pandemic bioterrorism, that is, using a high consequence biological agent to deliberately seed a pandemic. In addition to the lack of a baseline risk estimate, there is also a lack of literature estimating the relative biosecurity risk various forms of dual-use research pose, making direct assessments of dual-use trade offs challenging to conduct. To assess the biosecurity risks of PVI, this chapter estimates the question "What is the probability a pathogen identified by PVI is used biological weapon by a malicious actor to successfully seed a pandemic in the next 10 years, if the genome sequence and reverse genetics protocol for the virus are publicly available?". The initial model estimates the risk posed by the successful identification of a single pandemic-capable virus, and is then scaled to assess the risks posed by the identification of multiple pandemic viruses.

5.1 Prior Literature

Literature specifically exploring pandemic bioterrorism is quite sparse, though more general probabilistic risk assessment of terrorism, particularly bioterrorism, has been an area of significant government focus since the early 2000s. The development and implementation of these assessments have primarily been carried out by government agencies and their contractors. The most notable effort is the Bioterrorism Risk Assessment (BTRA) developed by the Department of Homeland Security (DHS) 2001 in response to the Amerithrax attacks [109]. This initial version received significant criticism, with the National Academies noting this framework should assume terrorist attacks are from intelligent adversaries making strategic decisions rather than random events, as well as noting the framework could be simplified through directly assessing probabilities instead of probability distributions for various event outcomes. In addition to the BTRA, other researchers have proposed alternative approaches to bioterrorism risk assessment. Radosavljevic and Belojevic (2009) introduced a model categorizing bioterrorist attacks into strategic, operational, and tactical levels, aiming to provide a more nuanced understanding of potential threats [110]. Their model incorporates qualitative and quantitative parameters for four key components: perpetrators, agents, means of delivery, and targets. Ezell et al. (2010) emphasized the value of using probabilities to quantify terrorism risks and advocated for the use of event trees to decompose terrorism scenarios, while also discussing the potential of other tools such as fault trees, Bayesian network analysis, and game theory in terrorism risk analysis, noting there remains a lack of consensus regarding approaches and assumptions made surrounding a terrorist's intentions, and where they make optional decisions to maximize consequences based on their goals [111].

These prior modeling efforts have predominantly focused on bioterrorism involving non-transmissible agents, such as anthrax. Historically, most bioterrorism incidents have used agents that do not efficiently spread between people, requiring actors to produce and disseminate large quantities of an agent to cause harm. This is evident in the Soviet and Iraqi biological weapons programs, where substantial stockpiles were accumulated [104]. However, pandemic bioterrorism presents unique risks compared to these historical cases. Unlike attacks with non-transmissible agents, pandemic agents have the potential to cause widespread harm with a much smaller initial release. Inglesby and Relman note that a novel pandemic-capable virus could potentially seed a pandemic with relatively few initial infections, after which human-to-human transmission would drive further spread [112]. This lower barrier to initiating an attack, combined with the potential for global spread, suggests that the risks and impacts of pandemic bioterrorism may be significantly greater than those modeled in previous bioterrorism risk assessments.

One of the challenges of risk assessment surrounding pandemic bioterrorist is the lack of empirical data surrounding key parameters, such as the number of actors with the intention to release a pandemic agent as a biological weapon, the number of actors with the capabilities to obtain a pathogen or synthesize one from published data, and the likelihood a malicious and capable actor would successfully disseminate a biological weapon are not available, making it difficult to evaluate the risk. In response to this challenge, Inglesby and Relman [112] propose estimating these parameters by using available information and exploring the various motivations of malicious actors, development of emerging technologies, amount of dual-use information that could aid an actor, and historical examples of biological warfare or intent to use bioweapons. Some approaches have directly used expert elicitation, synthesizing the judgments of experts to identify where there are areas of consensus and where there is significant disagreement. A survey of experts influential in life science policy (n= 62) was asked to estimate the likelihood of a large-scale BW attack within the next 10 years, resulting in a mean of 57.5%, and 95% CI (49.4%, 65.7%) [113]. The same survey found experts estimated only a 27.7% chance that intelligence agencies would provide actionable warning before an attack. The same survey found experts estimated likelihood that intelligence agencies would provide an actionable indication or warning before an attack. In Gryphon Scientific's Risk and Benefit Analysis (RBA) of Gain of Function (GoF) research, they carried out qualitative and semi-quantitative risk assessments evaluating the biosecurity risks from malicious actors obtaining a GoF influenza pathogen from the lab itself through theft, and the risks posed by information produced by this research. They do not make any absolute quantitative risk assessments noting the lack of available data, though they generally note that if events such as theft of animals, materials or stocks by an insider lead to an initial infection, the risk

this leads to a global pandemic is much greater than an accidental LAI given the covert nature of intentional infections caused by malicious actors, estimating there is an 11% likelihood an initial deliberate infection would lead to a pandemic [92]. While the RBA contains valuable information, one key limitation noted by the U.S. National Science Advisory Board for Biosecurity was the lack of an established baseline risks and fully quantitative assessments [114].

Using this prior literature, the section below outlines a complete quantitative assessment of the biosecurity risks associated with PVI, estimating the likelihood and expected harm of a deliberate pandemic event over the next decade if PVI were to identify a pandemic-capable virus.

5.2 Model Structure

Risk assessments are often broken down into three key components: the scenario, likelihood, and consequences of the scenario. [115]. In this model, we first consider the following scenario:

> Threat Scenario: PVI research publicly publishes the whole genome sequence, results of characterization experiments, and a reverse genetics protocol for a novel pandemic-capable virus. An actor (person, group, country) constructs the virus using published data and synthetic DNA (or obtains the virus through other means) and uses it to successfully deliberately seed a pandemic event.

The general framework of threat assessment then takes the product of two elements: the likelihood of the attack and the consequences of the attack [116]. Each of these elements can be decomposed further, where likelihood of an attack is a function of the motivation and capability of the attacker(s), and the consequences are determined by the pathogen used, the number of potential victims, and the vulnerability of these victims (Fig. 5).
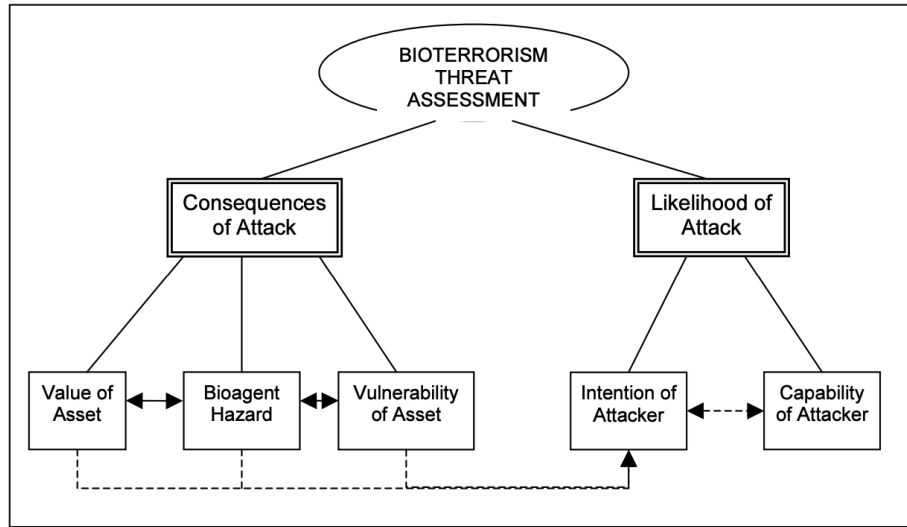
Figure 5: Bioterrorism Threat Assessment Framework. This diagram shows how the threat of a bioterrorist attack is evaluated by the consequences and the likelihood of a threat. Arrows indicate interdependencies among these factors. [116]

In Ackerman and Moran's framework, the consequences of an attack are defined by the biological agent used, the number of people at risk, and the vulnerability of the people affected. A novel pandemic-capable virus identified through PVI efforts would have the capability of global efficient transmission amongst humans, typically with an $R_0 > 1$, as well as considerable virulence, with case fatality rates for COVID-19 ranging from 0.7% to 4% for the various strains [117], and an average CFR of approximately 0.78% for historical pandemic events [60]. The general population is also particularly vulnerable to a PVI identified pathogen as there is a lack of pre-existing population immunity to novel pathogens. Due to these factors, as a conservative estimate we assume the conditional expected harm of a deliberate pandemic event is equivalent to that of a natural pandemic event from a novel pandemic-capable virus, and use the Fan et al. estimate of 21.6 million deaths for a pandemic event.

This model primarily focuses on evaluating the likelihood of a deliberate pandemic event over the next decade, focusing on the two key elements highlighted in the framework above: capability and intent. The technical skill and expertise surrounding synthetic virus assembly varies greatly between malicious actors, with some able to synthesize or obtain live samples of a virus within months, while others would need several years to develop the relevant skills and knowledge. Similarly, malicious actors also have varying degrees of commitment and persistence to causing harm. Some may give up after a year of trying or be dissuaded from using biological agents due to the time needed to develop skills, while others may be motivated enough to dedicate years of concerted effort. To capture this, we consider three different

categories of actors that have the capability to synthesize or acquire viruses: highly skilled actors, generally skilled actors, and capable unskilled actors. We define them each below:

- Highly Skilled Actors: Proficient synthetic virologists and researchers who already have the specific skill set associated with synthetic virology (e.g. viral assembly, purification, mammalian cell culture, etc.). We assume these actors could synthesize and release novel agents within a year.

- Generally Skilled Actors: Those with some experience with wet lab biology such as general life scientists, but would need to invest some time to familiarize themselves with the specific techniques necessary for viral assembly. We assume this category of actors would not be able to immediately assemble a virus given the sequence and reverse genetics protocol, but would be able to with between 1 to 4 years of training.

- Unskilled Capable Actors: Those with no current relevant biology or wet-lab experience, though have sufficient competency to acquire the necessary skills if they dedicate between 4 to 10 years of effort.

We generally conceptualize bioterrorism likelihood using the formula:

$$P_{attack} = p(capability) \times p(intent \,|\, capability) \times p(success)$$

Where capability represents whether an actor has the technical skills and access to resources necessary to obtain the virus within X number of years (either through financial resources or through access to an established facility), and intent represents whether an actor would attempt to release a pandemic virus if it required X years of effort to do so.

5.3 Misuse Risks Academic Surveys

While there is some literature discussing methodologies surrounding bioterrorism risk assessments, this research primarily focuses on exploring approaches to threat assessments and how to structure models and frameworks, rather than complete assessments with quantitative conclusions regarding the risk. Due to this, there is a lack of publicly available estimates of key parameters surrounding reverse genetics capabilities and intentions to use biological agents to cause large scale harm. To address this gap, we carried out two types of surveys: a synthetic virology capabilities survey and two surveys regarding intentions to cause large scale harm using pandemic agents.

Similar to the benefits survey expert selection, we selected a list of academic journals to obtain a list of experts for the capabilities survey and terrorism intention surveys. We then used Scopus Sources search to gather a list of all authors who had published at least twice in any of the selected journals between Jan 1, 2019 through June 1, 2023, inclusive. The two publications did not necessarily need to be in the same journal, though the publications had to be either in the 'Article' or 'Review' category, and could not have more than 30 authors.

For the synthetic virology capabilities survey, the following journals were selected for participant selection: *Current Opinion in Virology*, *Annual Review of Virology*, *Journal of General Virology*, *ACS Synthetic Biology*, *Nature Biotechnology*, *Gene Therapy*, *Molecular Systems Biology*, *Current Opinion in Biotechnology*, *G3: Genes, Genomes, Genetics*. This provided us with a list of 3449 authors, where we were able to obtain the email addresses of 2432 people and receive responses from 24 participants. This survey asked experts to estimate the duration it would take a skilled virologist to perform reverse genetics for a newly discovered virus with and without a detailed protocol, the number of individuals who would be able to perform reverse genetics within a year to obtain a functional live virus with and without a protocol, and the number of individuals who would be able to carry out these experiments immediately, but would able to perform them within 1-4 years, with and without a protocol. Each question asked participants to estimate values assuming the virus was an influenza virus, coronavirus, and paramyxovirus. The complete survey questions and results can be found in Appendix B.

For the terrorist intention surveys, we selected participants using the criteria above from the following journals: *Terrorism and Political Violence*, *Perspectives on Terrorism*, *Studies in Conflict & Terrorism*, *Journal of Policing, Intelligence and Counter Terrorism*, *Journal for Deradicalization*, *International Security*. This yielded 1650 authors. Using web searches, we were able to obtain email addresses for 1,407 (85%) of them. For our initial terrorism survey, this resulted in 111 responses. Similar to our benefits assessment we sent a follow-up survey aiming to clarify participants' answers to the same group of experts, and received 115 responses. The complete survey questions and results can be found in Appendix C.

In our initial terrorism survey, we first ask participants to estimate the overall likelihood of a successful pandemic bioterrorist attack over the next decade conditional on a novel pandemic virus being publicly identified. This served as a basis to compare the outputs of our model based on their estimates for specific actor groups in subsequent questions. We then compared participants' overall estimates to the estimates from our model.

5.4 Model Framework and Calculations

5.4.1 Initial Approach: Per-Actor Probability Estimates

We started with N_{total}, which is the total number of individuals who either already have or could reasonably gain the capability for viral assembly. We define this by the number of people who have acquired doctoral degrees over the past 20 years around the world, and estimate it to be approximately 5.6 million individuals, using NSF and OECD data.

In our synthetic virology capabilities survey, participants estimated an average of 25,277 people would be able to perform reverse genetics with an established protocol within 1 year, with a median of 10,000. We use the median value to estimate the number of highly skilled actors in the world. Participants also estimated an average and median number of 150486 and 55000 people respectively for the number of people that would be able to synthesize a virus between 1 and 4 years. We use the median value to estimate the number of generally skilled actors.

To estimate the number of capable unskilled actors in the world, we estimated the number of people with doctorate degrees globally over the past 20 years. We used a combination of NSF regarding the number of doctorates awarded in the U.S. [118]and OECD data regarding the proportion of global doctorates awarded in the U.S. [119]. Using this data, we estimated there are approximately 4.27 million who have received a doctorate degree in the past 20 years around the world.

In our initial terrorism survey, we asked participants to estimate the conditional probability an actor would have the attempt to cause global harm using a synthesized virus given they had the capability to do so. We asked three versions of this question, asking about the likelihood if it required up to a year of effort, 1 to 4 years, and 4 to 10 years. See Appendix B for complete survey questions and results of the synthetic virology capabilities survey. We used these values to estimate the conditional probability an actor had the intent to cause large scale harm given they had the necessary capabilities, $p(i \mid c_{Tx})$. We asked survey participants to estimate this value for each group, where they estimated a median of 0.1% for all three groups, means of 5.16%, 3.4% and 3.6% for T1, T2 and T3 respectively. Given the strong skew we use the median estimate of 0.1% for all $p(i \mid c_{Tx})$. The average response for likelihood of counterterrorism successfully preventing an attack was 59.7%. We took the complement to estimate the likelihood an attack is successful.

$$p(attack_{Tx}) = p(c) \times p(i \mid c_{Tx}) \times p(s)$$

$$P_{Tx} = 1 - (1 - p(attack_{Tx}))^{N_{total}}$$

Where $x = \{1, 2, 3\}$, representing the three groups of actors. We then define the total probability that at least one attack occurs in the next decade as:

$$P_{Total} = 1 - (1 - P_{T1}) \times (1 - P_{T2}) \times (1 - P_{T3})$$

We calculated the probabilities each group would successfully carry out an attack over the next decade, (P_{T1}, P_{T2}, P_{T3}); as well as the total probability of a successful attack P_{Total} for each participant. Using this approach, this resulted in an average of 71% for highly skilled actors (P_{T1}), 84% for generally skilled actors (P_{T2}), and 87% for unskilled capable actors (P_{T3}), with an overall average of 91% likelihood there would be a successful attack in the next decade (P_{Total}).

Their overall estimates of bioterrorism given by participants differed significantly from the estimates generated by the model, where model outputs were on average 66% higher than their overall estimate. This is likely due to the high estimates participants provided for the likelihood a capable actor has the intent to seed a deliberate pandemic, $p(i|c_{Tx})$. With a median of 0.1% across all three groups, this would suggest 1 in every 1000 individuals with sufficient capability have the intent to cause global indiscriminate harm. Using the Global Terrorism database as a reference, the global average base rate for terrorist attacks is approximately one attack per 700,000 people [120]. Assuming each individual carries out a single attack, we see the rates estimated through the survey for bioterrorism are far greater than the base rate of global terrorist attacks of all types. To address this, a followup survey was developed to clarify participants' estimates.

5.4.2 Revised Approach: Overall Probability Estimates

In our followup survey to terrorism experts, we directly asked participants to estimate a 90% confidence interval for the likelihood that at least one highly skilled actor attempts to carry out an attack, P_{HS}; at least one generally skilled actor attempts to carry out an attack, P_{GS}, and at least one capable unskilled actor attempts to carry out an attack, P_{US}, within the next decade. Participants were also asked to estimate the likelihood counterterrorism efforts would prevent an attack.

$$P_{total} = (1 - (1 - P_{HS}) \times (1 - P_{GS}) \times (1 - P_{US})) \times P(success)$$

Where P_{HS} represents the probability at least one highly skilled actor will successfully carry out an attack within the next decade, P_{HS} represents the probability at least one generally skilled actor will successfully carry out an attack within the next decade, and P_{US} represents the probability at least one capable unskilled actor will successfully carry out an attack within the next decade.

We estimated P_{total} for each participant and aggregated the results. The median of the lower bound estimates was 2.5% and the median of the upper bound estimates was 22.4%. The averages were slightly higher, at 7.5% and 31.2% respectively. Given the lower and upper bound estimates were skewed, we took the average of the median of the lower and upper bounds to derive a central estimate of a 12.45% likelihood of a deliberate pandemic event in the next decade. Using the Fan et al. estimate of the conditional expected harm of a pandemic event of 21.6 million deaths per pandemic event, this would result in a median central estimate of 2,689,200 deaths per decade per virus identified, with a lower bound 540,000 deaths and upper bound of 4,838,400 deaths We use these values in our overall risk-benefit assessment of PVI.

5.5 Discussion

In the model described in section 5.4.2 above, we quantitatively estimate the biosecurity risks posed by the identification of a single pandemic capable virus. Specifically, we estimated that if a PVI successfully identifies a pandemic-virus, and publishes the genome sequence, results of characterization experiments and verified reverse genetics protocol, there is a 12.45% [2.5%, 22.4%] likelihood of a malicious actor seeding a deliberate pandemic using the identified virus as a biological weapon, resulting in an expected harm of 2,689,200 [540,000 ; 4,838,400] over the next decade. This estimate is slightly lower than the overall estimate survey participants provided, where they estimate a median likelihood of 19% that the identified virus would be used to deliberately cause a pandemic.

Survey participants assigned the highest probability of an attempt to highly skilled actors, with median lower and upper bounds of 10% to 50%, followed closely by generally skilled actors at 10% to 37.5%, with capable unskilled at the lowest at 5% to 30%. Technical knowledge and skill barriers is often cited as a significant barrier to the development of BWs, where tacit knowledge developed through lab experience is often challenging to gather outside the lab [107].

In addition to quantitative estimates, our survey asked terrorism experts about the perceived motivations and actors driving bioterrorism risks. Participants rated religious-based ideologies and political extremism as the highest risk factors, scoring 3.53 and 3.24 out of 5 respectively. This aligns

with historical examples of bioterrorism attempts, such as the Aum Shinrikyo cult's efforts. Additionally, respondents expressed the highest level of concern about state or state-sponsored groups engaging in bioterrorism, ranking lone actors as the least concerning group.

Our findings heavily relied on survey data from terrorism and synthetic virology capability experts, which exhibited wide distributions and significant uncertainties. The broad 90% confidence interval (2.5% to 22.4%) reflects this uncertainty and the inherent challenges associated with predicting unprecedented high-consequence events. As such, the findings from this analysis should not be interpreted as definitive probabilities, but rather as exploratory estimates that may help bound potential outcomes.

5.5.1 Future Directions

While this study provides initial quantitative estimates of biosecurity risks associated with PVI, several key areas warrant further investigation. Firstly, additional research to develop more robust estimates of relevant parameters regarding synthetic virology capabilities and malicious intentions would allow for a better understanding of the current and emerging risk landscape. Additionally, future models should consider introducing complexity through considering additional parameters associated with the release scenario. This can include modeling the simultaneous release of multiple pathogens, estimate the conditional expected harm from attacks with multiple sites of release, and characteristics of the pathogen (e.g. R_0, incubation period).

Chapter 6: Expected Value of PVI Efforts: Findings and Analysis

6.1 Expected Value of PVI Research

The chapters above modeled the expected benefits, accident risks, and deliberate misuse risks of PVI, estimating the expected benefits and risks over the next decade. Our findings suggest that the identification of a single pandemic-capable virus through PVI efforts could potentially save 49,000 [10,500; 93,600] lives over the next decade by contributing to the reduction of natural pandemic risks.

However, this benefit is outweighed by the estimated anthropogenic risks. We calculated an expected harm of 49,800 [3,983; 365,438] deaths from potential accidental pandemic events and 2,689,200 [540,000; 4,838,400] deaths from deliberate misuse over the next decade. Using equation (5), we calculated the overall expected value (EV) of Pandemic Virus Identification (PVI) efforts by taking the difference between the benefits (defined as expected lives saved over a decade) and the risks (expected lives lost over a decade), which are the sum of accident risks and deliberate misuse risks.

$$E[PVI] \ = \ E\left[Benefits_{PVI \mid n=1}\right] \ - \ E[Accidents_{n=1}] \ - \ E[Misuse_{n=1}] \quad (5)$$

Through mathematical modeling, we estimate the EV of identifying a single pandemic-capable pathogen to be –2,690,000 lives, indicating that in expectation, the public identification of a single pandemic-capable virus will result in 2.69 million deaths over the next decade. For a conservative estimate of the EV, we take the upper bound of the benefit range and minimum of the accidental and deliberate misuse risks, resulting in an EV of –450,383 lives.

For an even more conservative estimate, if we assume the likelihood of a deliberate pandemic event to be 1% per decade or once every thousand years , the expected value remains negative, at –170,983 lives over the course of a decade using the upper bound of benefits and lower bound of accident risks. For the risks to outweigh the benefits in this scenario, the risk of a deliberate pandemic event over the next decade would need to be 0.208% or lower. This suggests that the EV of PVI remains strongly negative even with significant uncertainties and wide confidence intervals in both the benefits and risks assessments.

6.1.1 Expected Value Comparison of VDi and PVI

This project evaluated the benefits of VDi, but did not assess the biosafety or biosecurity risks associated with this work. VDi itself has garnered a fair amount of attention for the biosafety risks it poses to researchers and the surrounding community. This work often involves gathering bodily fluid samples such as saliva, blood and urine, requiring field teams to catch and handle live animals while samples are collected. Repeated accidents have been reported by numerous research groups, with multiple groups reporting team members have been bitten by bats, cases of airborne exposure due to equipment failures, and skin exposures to blood and urine [75,121]. Most of these projects occur in zoonotic hotspots and other animal reservoirs within low to middle income countries, with USAID's PREDICT program particularly focusing on regions in Africa and Southeast Asia [122]. These regions often have inadequate infrastructure to properly maintain high-containment labs and lax biosafety policies, with many countries having no reporting requirements at all [123,124]. This lack of systematic reporting of accidents

makes quantitative analysis of the biosafety risks challenging. Independently, some researchers have stepped back from VDi due to the risks the work posed to themselves and their team, suggesting the risks paired with and the lack of discernible benefits make them question whether they wanted to continue conducting the research [75].

If VDi research is critically important for prevention efforts, several improvements can be made to reduce the risks of large-scale harm. Funders and institutions can mandate that field teams undergo routine diagnostic surveillance, as well as quarantine following exposure events to prevent further spread of potential infections. From a biosecurity perspective, VDi on its own does not reveal which of the discovered viruses are particularly harmful, such that we assume the information produced by these efforts alone would not significantly influence the risk of a deliberate pandemic event.

6.2 Discussion

With PVI research and other forms of research involving potential pandemic pathogens such as specific forms of GoF research with PPPs, the qualitative approach risk-benefit assessments have made it challenging to consistently implement policy and guidance, as qualitative assessments make direct assessments of tradeoffs challenging and often impossible. Despite the inherent challenges and uncertainties within quantitative modeling, the NSABB notes that in addition to qualitative descriptions and assessments of risk, quantitative assessments are crucial for providing a more robust and objective framework for evaluating risks and benefits [114]. As some researchers suggest the benefits of characterization and GoF work are necessary for the development of novel prophylactic and therapeutic interventions [125], it is crucial to evaluate the likelihood these benefits will be realized, and if so, how much of a benefit they provide. The lack of qualitative consensus regarding the benefits of this work and the negative expected value of this work assessed in our model above suggests alternatives to PVI should be prioritized for funding. Interventions that empower communities in zoonotic hotspots to prevent spillover and suppress epidemics before they can spread such as behavioral interventions, active surveillance of animals and humans at these hotspots do not require PVI to carry out. Thanks to recent advances in biotechnology, these efforts could be much more effective than when they were initially implemented decades ago. Local communities with access to nanopore sequencing technology can obtain and share the sequence of a novel pathogen within a day of recognizing an outbreak[126], potentially enabling the development and manufacturing of CRISPR-based rapid diagnostics systems and targeted nucleic acid vaccines within weeks or even days [127]. Deploying these in a combined Phase 1/2 vaccination trial, including ring vaccination surrounding anyone ill who tests positive, could maximize the likelihood of containing the outbreak[128].

6.2.1 Current Governance & Policy Recommendations

In the *United States Government Policy for Oversight of Dual Use Research of Concern and Pathogens with Enhanced Pandemic Potential* issued on May 2024, two categories of research are defined, category 1 and category 2. Category 1 encompasses research involving a select set of high-consequence pathogens and toxins that could likely produce dual-use information or products through increasing the transmissibility, virulence, resistance to MCMs, and other characteristics that could make the agent more dangerous. Principal Investigators and Institutional Review Entities (IRE) review the research to determine if it meets the criteria for category 1, and if so, a risk-benefit assessment and risk mitigation plans submitted to and approved by the federal funding agency before the research can proceed. Category 2 is defined by research that uses or will reasonably produce a PPP that poses a significant threat to public health and national security. Category 2 has more stringent oversight, including a detailed review process by the IRE and multidisciplinary review entities at the federal funding agency level. In the policy and accompanying implementation guidance, some aspects of Pandemic Virus Identification (PVI) fall under Category 1 research, such as experiments that assess whether the virus can enter human cells (susceptibility) and replicate within them (permissiveness). Transmissibility experiments such as serial passaging experiments in primary human cells, human organoid systems or animal models fall under category 2 if they are reasonably anticipated to select for increased virulence or transmissibility in humans. While these experiments do not deliberately select for strains that can cause increased harm, Sandbrink et al. note even assessing for the replication potential of zoonotic viruses in human cell lines and animal models can select for viruses with greater transmissibility. These categorizations broadly inform the level of oversight, but do not make clear what types of research would be prohibited, or what specific risk mitigation measures would be taken to manage both the biosafety and biosecurity risks of PVI research.

To manage biosafety risks, alternatives to PVI such as the use of *in silico* experiments, sequence analysis, and computational modeling tools have been suggested [6]. This would reduce the amount of time researchers are physically handling PPPs and transporting them within facilities, reducing the likelihood of an LAI. Additionally, instead of working with the entire virus, some researchers suggest studying individual viral components such as studying the hemagglutinin (HA) protein for its binding affinity to human sialic acids and the pH stability of fusion proteins to understand viral adaptation mechanism [6]. If PVI research is being conducted with the full, infectious strain of the virus, efforts should focus on addressing human error, which accounts for the majority of lab-acquired infections. This could include improving personal protective equipment and its proper usage, and developing better systems that incentivize reporting and learning from near-miss incidents. Routine diagnostic surveillance of laboratory personnel may also increase the likelihood infections are quickly identified and contained.

Management of biosecurity risks requires involvement from stakeholders in all stages of the research life cycle. In the project development and grant proposal stages, funders and oversight bodies could require that risk-assessments be submitted in proposals, as well as seek out domain experts to evaluate proposals since scientists are not always aware of the potential misuse risks of their research [129]. Researchers and institutional review boards may design experiments with risk mitigation in mind and develop risk mitigation plans. Journal editors and publishers have a strong influence over the dissemination of scientific information, and can implement more robust screening mechanisms around research to minimize information risks surrounding this work. Mitigation measures might include only selectively sharing results of characterization experiments to relevant stakeholders to use in surveillance, publishing the partial genome instead of the whole genome, or broadly prohibit the sharing of characterization experiments that reveal a novel pathogen has pandemic capabilities [42,128].

While the current policy framework provides a foundation for oversight, there is a need for more detailed, actionable guidelines that specifically address the unique challenges posed by PVI research. Future policy development should focus on bridging the gap between these broad categories and the specific risk management strategies needed for different types of PVI experiments.

6.3 Limitations

6.3.1 Reliance on Expert Surveys

A key limitation of this modeling work is the heavy reliance on academic surveys and expert elicitation to derive estimates for various aspects of this model. Research by Tetlock et al. finds making accurate predictions about the future is quite difficult even for domain experts, with this made even more difficult with complex topics and unprecedented events such as pandemic bioterrorism and potential benefits of identification of a novel pandemic virus [130]. Therefore, while our mathematical model provides a structured approach to estimating potential benefits and risks, the results of this study should be viewed as exploratory estimates that may be useful for bounding potential outcomes, not as definitive assessment of the benefits and risks of this research. We hope future research is conducted to gather more robust estimates of key parameters to evaluate the benefits and risks associated with PVI research and other work with pandemic pathogens.

6.3.2 Rapid Developments in Emerging Biotechnologies

Additionally, in both the benefits and risk assessments, this work simplifies the complex systems and decision-making processes, and accordingly does not fully capture them in the frameworks. On the benefits side, the model assumes a relatively straightforward relationship between VDi/PVI and medical countermeasure development, which may not capture the full complexity of this process. Similarly, our biosecurity risk model does not fully account for the nuanced decision-making processes of potential malicious actors. One significant complexity that has not been incorporated into this model is the rapidly evolving landscape of biotechnology and synthetic biology. Capabilities in these fields are not only improving but also becoming more accessible, potentially significantly influencing the risk landscape over time. The increasing democratization of information and technology associated with PVI research may alter both the benefits and risks in ways that are difficult to predict. Our model provides a snapshot based on current understanding and capabilities, but it does not fully capture future developments that could shift the risk-benefit balance. This limitation underscores the need for ongoing reassessment of the PVI risk-benefit equation as technologies and capabilities evolve.

Chapter 7: Conclusion

In this thesis, we presented a comprehensive mathematical framework for assessing the risks and benefits of Pandemic Virus Identification (PVI) research. Through mathematical modeling and expert surveys, we evaluated the potential benefits of PVI in reducing natural pandemic risks, as well as the biosafety and biosecurity risks associated with this research. The overall expected value of identifying a single pandemic-capable pathogen was estimated to be –2.69 million, indicating that in expectation, the public identification of a single pandemic-capable virus could result in 2.69 million deaths over the next decade. Even with conservative estimates favoring benefits and minimizing risks, the expected value remains strongly negative. While PVI aims to improve pandemic preparedness, our analysis suggests that the risks of this approach may outweigh its benefits.

These findings suggest attention and resources should prioritize other forms of pandemic preparedness interventions that carry fewer dual-use risks compared to PVI. These could include improving surveillance systems at known spillover hotspots, and empowering local communities to detect, sequence, and suppress nascent epidemics. On the governance side, the scientific community and policymakers should continue to work together to develop robust governance frameworks that can effectively manage the risks associated with dual-use research while ensuring the beneficial research is not hampered. If PVI research continues to be carried out, biosafety and biosecurity measures should be significantly improved to mitigate the risks of accidental release and deliberate misuse.

There are several areas where future research could enhance our understanding and refine our estimates. Most notably, many of our model parameters rely heavily on expert elicitation. Future studies could gather empirical data to validate and refine these estimates, as well as conduct more extensive expert elicitation such as Delphi studies. Additionally, longitudinal studies tracking the impact of virus discovery on medical countermeasure development could provide valuable insights. To improve estimates of parameters for the benefits assessment, longitudinal studies could track the impact of virus discovery and PVI on medical countermeasure development and non-pharmaceutical interventions. To capture the changing risk landscape due to emerging biotechnology, consistent evaluations of various actors' capabilities and access to various tools could improve risk assessments.

While the pursuit of knowledge around pandemic pathogens to prevent future pandemics is well intentioned, our analysis suggests that the current approach to PVI may be creating more risk than it mitigates. As we continue to face global health challenges, it is important for us to critically evaluate our research strategies, ensuring we are optimally using the limited resources available and truly improving global health security.

Citations

1. Douglas, T. The dual-use problem, scientific isolationism and the division of moral labour. *Monash Bioeth. Rev.* **32**, 86–105 (2014).

2. National Academies of Sciences, Engineering, and Medicine, Policy and Global Affairs, Committee on Science, Technology, and Law & Committee on Dual Use Research of Concern: Options for Future Management. *Dual Use Research of Concern in the Life Sciences: Current Issues and Controversies*. (National Academies Press, 2017). doi:10.17226/24761.

3. Rosenwald, M. S. The invention of sarin was an accident. A German scientist was trying to kill bugs. *Washingtonpost. com* (2017).

4. Hama, S., Al-Jaff, B. & Mahmud, B. The effect of chemical warfare agents on the immune system of survivors in Halabja. *J. Zankoy Sulaimani - A* **11**, 41–52 (2007).

5. Research involving enhanced potential pandemic pathogens. *National Institutes of Health (NIH)* https://www.nih.gov/news-events/research-involving-potential-pandemic-pathogens (2021).

6. Lipsitch, M. & Galvani, A. P. Ethical alternatives to experiments with novel potential pandemic pathogens. *PLoS Med.* **11**, e1001646 (2014).

7. Gilbertson, B. & Subbarao, K. What Have We Learned by Resurrecting the 1918 Influenza Virus? *Annu Rev Virol* **10**, 25–47 (2023).

8. Plowright, R. K. & Hudson, P. J. From Protein to Pandemic: The Transdisciplinary Approach Needed to Prevent Spillover and the Next Pandemic. *Viruses* **13**, (2021).

9. Lipsitch, M. & Bloom, B. R. Rethinking biosafety in research on potential pandemic pathogens. *MBio* **3**, (2012).

10. Casadevall, A. & Imperiale, M. J. Risks and benefits of gain-of-function experiments with pathogens of pandemic potential, such as influenza virus: a call for a science-based discussion. *MBio* **5**, e01730–14 (2014).

11. Offit, P. A. The Cutter incident, 50 years later. *N. Engl. J. Med.* **352**, 1411–1412 (2005).

12. Nathanson, N. & Langmuir, A. D. The Cutter incident. Poliomyelitis following formaldehyde-inactivated poliovirus vaccination in the United States during the Spring of 1955. II. Relationship of poliomyelitis to Cutter vaccine. 1963. *Am. J. Epidemiol.* **142**, 109–40;

discussion 107–8 (1995).

13. Barquet, N. & Domingo, P. Smallpox: the triumph over the most terrible of the ministers of death. *Ann. Intern. Med.* **127**, 635–642 (1997).

14. Koblentz, G. D. *Living Weapons: Biological Warfare and International Security*. (Cornell University Press, 2011).

15. Frinking, E., Sinning, P., Bontje, E., della Frattina, C. F. & Abdalla, M. *The Increasing Threat of Biological Weapons: Handle with Sufficient and Proportionate Care*. (The Hague Centre for Strategic Studies, 2017).

16. Olson, K. B. Aum Shinrikyo: once and future threat? *Emerg. Infect. Dis.* **5**, 513–516 (1999).

17. Tucker, J. B. Breaking the deadlock over destruction of the smallpox virus stocks. *Biosecur. Bioterror.* **9**, 55–67 (2011).

18. Imperiale, M. J., Howard, D. & Casadevall, A. The Silver Lining in Gain-of-Function Experiments with Pathogens of Pandemic Potential. *Methods Mol. Biol.* **1836**, 575–587 (2018).

19. Institute of Medicine (US) Committee to Study Decision Making. *Biomedical Politics*. (National Academies Press (US), Washington (DC), 1991). doi:10.17226/1793.

20. Racaniello, V. R. & Baltimore, D. Cloned poliovirus complementary DNA is infectious in mammalian cells. *Science* **214**, 916–919 (1981).

21. Neumann, G. *et al.* Generation of influenza A viruses entirely from cloned cDNAs. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9345–9350 (1999).

22. Jackson, R. J. *et al.* Expression of mouse interleukin-4 by a recombinant ectromelia virus suppresses cytolytic lymphocyte responses and overcomes genetic resistance to mousepox. *J. Virol.* **75**, 1205–1210 (2001).

23. Cello, J., Paul, A. V. & Wimmer, E. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science* **297**, 1016–1018 (2002).

24. National Research Council, Policy and Global Affairs, Development, Security, and Cooperation & Committee on Research Standards and Practices to Prevent the Destructive Application of Biotechnology. *Biotechnology Research in an Age of Terrorism*. (National Academies Press, 2004). doi:10.17226/10827.

25.   Taubenberger, J. K. *et al.* Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889–893 (2005).

26.   van Aken, J. When risk outweighs benefit. Dual-use research needs a scientifically sound risk-benefit analysis and legally binding biosecurity measures. *EMBO Rep.* **7 Spec No**, S10–3 (2006).

27.   van Aken, J. Risks of resurrecting 1918 flu virus outweigh benefits. *Nature* vol. 439 266 (2006).

28.   National Research Council *et al. Globalization, Biosecurity, and the Future of the Life Sciences.* (National Academies Press, 2006). doi:10.17226/11567.

29.   Herfst, S. *et al.* Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* **336**, 1534–1541 (2012).

30.   Imai, M. *et al.* Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* **486**, 420–428 (2012).

31.   Lipsitch, M. Why Do Exceptionally Dangerous Gain-of-Function Experiments in Influenza? in *Influenza Virus: Methods and Protocols* (ed. Yamauchi, Y.) 589–608 (Springer New York, New York, NY, 2018). doi:10.1007/978-1-4939-8678-1_29.

32.   Administration for Strategic Preparedness and Response. *Screening Framework Guidance for Providers and Users of Synthetic Nucleic Acids.* https://aspr.hhs.gov/legal/synna/Documents/SynNA-Guidance-2023.pdf (2023).

33.   Office of Science Policy. *NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules (NIH Guidelines).* https://osp.od.nih.gov/wp-content/uploads/NIH_Guidelines.pdf (2024).

34.   Hou, Y. J. *et al.* Host range, transmissibility and antigenicity of a pangolin coronavirus. *Nat Microbiol* **8**, 1820–1833 (2023).

35.   Wan, H. *et al.* Replication and transmission of H9N2 influenza viruses in ferrets: evaluation of pandemic potential. *PLoS One* **3**, e2923 (2008).

36.   Menachery, V. D. *et al.* SARS-like WIV1-CoV poised for human emergence. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 3048–3053 (2016).

37.   U.S. Agency for International Development. *Notice of Funding Opportunity (NOFO) Discovery & Exploration of Emerging Pathogens – Viral Zoonoses (DEEP VZN).*

https://www.grants.gov/search-results-detail/329847 (2021).

38. Kelly, T. R. *et al.* Implementing One Health approaches to confront emerging and re-emerging zoonotic disease threats: lessons from PREDICT. *One Health Outlook* **2**, 1 (2020).

39. Wille, M., Geoghegan, J. L. & Holmes, E. C. How accurately can we assess zoonotic risk? *PLoS Biol.* **19**, e3001135 (2021).

40. Carlson, C. J. From PREDICT to prevention, one pandemic later. *Lancet Microbe* **1**, e6–e7 (2020).

41. Willman, D. The US quietly terminates a controversial $125m wildlife virus hunting programme amid safety fears. *BMJ* **382**, 2002 (2023).

42. Sandbrink, J., Ahuja, J., Swett, J., Koblentz, G. & Standley, C. Mitigating biosecurity challenges associated with zoonotic risk prediction. *SSRN Electron. J.* (2022) doi:10.2139/ssrn.4035760.

43. MacIntyre, C. R. Re-thinking the ethics of dual-use research of concern on transmissible pathogens. *Environment Systems and Decisions* **35**, 129–132 (2015).

44. Ellwanger, J. H. & Chies, J. A. B. Zoonotic spillover: Understanding basic aspects for better prevention. *Genet. Mol. Biol.* **44**, e20200355 (2021).

45. Plowright, R. K. *et al.* Pathways to zoonotic spillover. *Nat. Rev. Microbiol.* **15**, 502–510 (2017).

46. Pandit, P. S. *et al.* Predicting the potential for zoonotic transmission and host associations for novel viruses. *Commun Biol* **5**, 844 (2022).

47. Simpson, S., Kaufmann, M. C., Glozman, V. & Chakrabarti, A. Disease X: accelerating the development of medical countermeasures for the next pandemic. *Lancet Infect. Dis.* **20**, e108–e115 (2020).

48. Brende, B. *et al.* CEPI-a new global R&D organisation for epidemic preparedness and response. *Lancet* **389**, 233–235 (2017).

49. CEPI. Disease X. https://cepi.net/disease-x.

50. Lu, L. *et al.* Temporal Dynamics, Discovery, and Emergence of Human-Transmissible RNA Viruses. *Mol. Biol. Evol.* **41**, (2024).

51. Adalja, A. A., Watson, M., Toner, E. S., Cicero, A. & Inglesby, T. V. Characteristics of Microbes Most Likely to Cause Pandemics and Global Catastrophes. *Curr. Top. Microbiol. Immunol.* **424**,

1–20 (2019).

52. Mipatrini, D. *et al.* 'Disease X'-time to act now and prepare for the next pandemic threat. *Eur. J. Public Health* **32**, 841–842 (2022).

53. Becker, D. J., Washburne, A. D., Faust, C. L., Mordecai, E. A. & Plowright, R. K. The problem of scale in the prediction and management of pathogen spillover. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **374**, 20190224 (2019).

54. Berry, K., Horan, R. D., Finnoff, D., Pompa, R. & Daszak, P. Investing to Both Prevent and Prepare for COVID-XX. *Ecohealth* **19**, 114–123 (2022).

55. Rodrigue, V., Gravagna, K., Yao, J., Nafade, V. & Basta, N. E. Current progress towards prevention of Nipah and Hendra disease in humans: A scoping review of vaccine and monoclonal antibody candidates being evaluated in clinical trials. *Trop. Med. Int. Health* **29**, 354–364 (2024).

56. van der Mensbrugghe, H. T. A. B. D. *Evaluating the Economic Consequences of Avian Influenza*. https://documents1.worldbank.org/curated/en/977141468158986545/pdf/474170WP0Evalu10 1PUBLIC10Box334133B.pdf (2008).

57. Dobson, A. P. *et al.* Ecology and economics for pandemic prevention. *Science* **369**, 379–381 (2020).

58. Bernstein, A. S. *et al.* The costs and benefits of primary prevention of zoonotic pandemics. *Sci Adv* **8**, eabl4183 (2022).

59. Carroll, D. *et al.* Building a global atlas of zoonotic viruses. *Bull. World Health Organ.* **96**, 292–294 (2018).

60. Madhav, N. *et al.* Pandemics: Risks, Impacts, and Mitigation. in *Disease Control Priorities: Improving Health and Reducing Poverty* (eds. Jamison, D. T. et al.) (The International Bank for Reconstruction and Development / The World Bank, Washington (DC), 2017). doi:10.1596/978-1-4648-0527-1.

61. Hinchliffe, S. *et al.* Understanding the roles of economy and society in the relative risks of zoonosis emergence from livestock. *R Soc Open Sci* **11**, 231709 (2024).

62. Holmes, E. C., Rambaut, A. & Andersen, K. G. Pandemics: spend on surveillance, not prediction. *Nature* **558**, 180–182 (2018).

63. Jonas, O. & Seifman, R. Do we need a Global Virome Project? *Lancet Glob Health* **7**,

e1314–e1316 (2019).

64. Vora, N. M. *et al.* Interventions to reduce risk for pathogen spillover and early disease spread to prevent outbreaks, epidemics, and pandemics. *Emerg. Infect. Dis.* **29**, 1–9 (2023).

65. Office of the Commissioner. FDA Approves First Oral Antiviral for Treatment of COVID-19 in Adults. *U.S. Food and Drug Administration* https://www.fda.gov/news-events/press-announcements/fda-approves-first-oral-antiviral-treatment-covid-19-adults (2024).

66. Więcek, W., Johnston, D., Dulka, T., Toomey, D. & Lewis, E. Vaccines at velocity: Evaluating potential lives saved by earlier vaccination in the COVID-19 pandemic. *bioRxiv* (2023) doi:10.1101/2023.06.16.23291442.

67. Office of the Commissioner. FDA Approves First COVID-19 Vaccine. *U.S. Food and Drug Administration* https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine (2024).

68. Roser, M., Ritchie, H., Ortiz-Ospina, E. & Hasell, J. Coronavirus Pandemic (COVID-19). *Our World in Data* (2020).

69. Shah, M. M. *et al.* Paxlovid Associated with Decreased Hospitalization Rate Among Adults with COVID-19 - United States, April-September 2022. *MMWR Morb. Mortal. Wkly. Rep.* **71**, 1531–1537 (2022).

70. Fan, V. Y., Jamison, D. T. & Summers, L. H. The Loss from Pandemic Influenza Risk. in *Disease Control Priorities: Improving Health and Reducing Poverty. 3rd edition* (The International Bank for Reconstruction and Development / The World Bank, 2017). doi:10.1596/978-1-4648-0527-1_ch18.

71. Authored by the members of the One Health High-Level Expert Panel (OHHLEP) *et al.* Prevention of zoonotic spillover: From relying on response to reducing the risk at source. *PLoS Pathog.* **19**, e1011504 (2023).

72. Fonseca, B. de P. & Morel, C. M. Relevance of national, regional and global virome projects on pandemics prediction, prevention, and control: a social network analysis of GVP-citing articles. *Mem. Inst. Oswaldo Cruz* **118**, e230116 (2023).

73. Luby, S. P. The pandemic potential of Nipah virus. *Antiviral Res.* **100**, 38–43 (2013).

74. Johnson, T., Jamrozik, E., Hurst, T., Cheah, P. Y. & Parker, M. J. Ethical issues in Nipah virus control and research: addressing a neglected disease. *J. Med. Ethics* (2023) doi:10.1136/jme-2023-109469.

75. David Willman, J. W. Research with exotic viruses risks a deadly outbreak, scientists warn. *The Washington Post*.

76. Monrad, J. & Katz, R. Biosecurity, biosafety, and the management of dangerous pathogens for public health research. *Viral sovereignty and* 100–119 (2020) doi:10.1017/9781108676076.008.

77. Artika, I. M. & Ma'roef, C. N. Laboratory biosafety for handling emerging viruses. *Asian Pac. J. Trop. Biomed.* **7**, 483–491 (2017).

78. Nikolakakis, I., Michaleas, S. N., Panayiotakopoulos, G., Papaioannou, T. G. & Karamanou, M. The History of Anthrax Weaponization in the Soviet Union. *Cureus* **15**, e36800 (2023).

79. Pedrosa, P. B. S. & Cardoso, T. A. O. Viral infections in workers in hospital and research laboratory settings: a comparative review of infection modes and respective biosafety aspects. *Int. J. Infect. Dis.* **15**, e366–76 (2011).

80. Geddes, A. M. The history of smallpox. *Clin. Dermatol.* **24**, 152–157 (2006).

81. Lim Poh Lian *et al.* Laboratory-Acquired Severe Acute Respiratory Syndrome. *N. Engl. J. Med.* **350**, 1740–1745.

82. Britton, S. *et al.* Laboratory-acquired dengue virus infection--a case report. *PLoS Negl. Trop. Dis.* **5**, e1324 (2011).

83. Demaneuf, G. The good, the bad and the ugly: A review of SARS lab escapes. in (Zenodo, 2020). doi:10.5281/ZENODO.4293257.

84. Owens, B. Anthrax and smallpox errors highlight gaps in US biosafety. *Lancet* **384**, 294 (2014).

85. Manheim, D. & Lewis, G. High-risk human-caused pathogen exposure events from 1975-2016. *F1000Res.* **10**, 752 (2021).

86. Gillum, D., Krishnan, P. & Byers, K. A Searchable Laboratory-Acquired Infection Database. *Appl. Biosaf.* **21**, 203–207 (2016).

87. Wedum, A. G. *Assessment of Risk of Human Infection in the Microbiological Laboratory*. (Department of the Army, Fort Detrick, 1966).

88. Klotz, L. C. & Sylvester, E. J. The consequences of a lab escape of a potential pandemic pathogen. *Front Public Health* **2**, 116 (2014).

89. Henkel, R. D., Miller, T. & Weyant, R. S. Monitoring Select Agent Theft, Loss and Release Reports in the United States—2004–2010. *Appl. Biosaf.* **17**, 171–180 (2012).

90. Lipsitch, M. & Inglesby, T. V. Moratorium on research intended to create novel potential pandemic pathogens. *MBio* **5**, (2014).

91. Merler, S., Ajelli, M., Fumanelli, L. & Vespignani, A. Containing the accidental laboratory escape of potential pandemic influenza viruses. *BMC Med.* **11**, 252 (2013).

92. Casagrande, R. *et al. Risk and Benefit Analysis of Gain of Function Research.* https://gryphonsci.wpengine.com/wp-content/uploads/2018/12/Risk-and-Benefit-Analysis-of-Gain-of-Function-Research-Final-Report-1.pdf (2016).

93. Fouchier, R. A. M. Studies on influenza virus transmission between ferrets: the public health risks revisited. *mBio* vol. 6 (2015).

94. Goad, M. A new interactive map reveals where the deadliest germs are studied. *Schar School of Policy and Government* https://schar.gmu.edu/news/2021-07/new-interactive-map-reveals-where-deadliest-germs-are-studied.

95. Evans, N. G. Dual-Use and Infectious Disease Research. in *Infectious Diseases in the New Millennium: Legal and Ethical Challenges* (eds. Eccleston-Turner, M. & Brassington, I.) 193–215 (Springer International Publishing, Cham, 2020). doi:10.1007/978-3-030-39819-4_9.

96. Alibek, K. & Handelman, S. *Biohazard: The Chilling True Story of the Largest Covert Biological Weapons Program in the World--Told from the Inside by the Man Who Ran It.* (Random House Publishing Group, 2014).

97. Musunuri, S., Sandbrink, J. B., Monrad, J. T., Palmer, M. J. & Koblentz, G. D. Rapid Proliferation of Pandemic Research: Implications for Dual-Use Risks. *MBio* **12**, e0186421 (2021).

98. Miller, S. & Selgelid, M. J. Ethical and philosophical consideration of the dual-use dilemma in the biological sciences. *Sci. Eng. Ethics* **13**, 523–580 (2007).

99. Riedel, S. Biological warfare and bioterrorism: a historical review. *Proc.* **17**, 400–406 (2004).

100. Wheelis, M. Biological warfare at the 1346 siege of Caffa. *Emerg. Infect. Dis.* **8**, 971–975 (2002).

101. Poupard, J. A. & Miller, L. A. History of biological warfare: catapults to capsomeres. *Ann. N. Y. Acad. Sci.* **666**, 9–20 (1992).

102. Frischknecht, F. The history of biological warfare. Human experimentation, modern nightmares and lone madmen in the twentieth century. *EMBO Rep.* **4 Spec No**, S47–52 (2003).

103. Leitenberg, M., Zilinskas, R. A. & Kuhn, J. H. *The Soviet Biological Weapons Program*. (Harvard University Press, 2012). doi:10.4159/harvard.9780674065260.

104. Zilinskas, R. A. Iraq's biological weapons. The past as future? *JAMA* **278**, 418–424 (1997).

105. Yuki, H. *et al. Aum Shinrikyo: Insights into How Terrorists Develop Biological and Chemical Weapons*. https://www.cnas.org/publications/reports/aum-shinrikyo-insights-into-how-terrorists-develop-biological-and-chemical-weapons (2011).

106. Tucker, J. B. Historical trends related to bioterrorism: An empirical analysis. *Emerg. Infect. Dis.* **5**, 498–504 (1999).

107. Ouagrham-Gormley, S. B. Barriers to bioweapons: Intangible obstacles to proliferation. *Int. Secur.* **36**, 80–114 (2012).

108. O'Donoghue, O. *et al.* BioPlanner: Automatic Evaluation of LLMs on Protocol Planning in Biology. *arXiv [cs.CL]* (2023) doi:10.48550/ARXIV.2310.10632.

109. National Research Council *et al. Department of Homeland Security Bioterrorism Risk Assessment: A Call for Change*. (National Academies Press, 2009).

110. Radosavljevic, V. & Belojevic, G. A new model of bioterrorism risk assessment. *Biosecur. Bioterror.* **7**, 443–451 (2009).

111. Ezell, B. C., Bennett, S. P., von Winterfeldt, D., Sokolowski, J. & Collins, A. J. Probabilistic risk analysis and terrorism risk. *Risk Anal.* **30**, 575–589 (2010).

112. Inglesby, T. V. & Relman, D. A. How likely is it that biological agents will be used deliberately to cause widespread harm? Policymakers and scientists need to take seriously the possibility that potential pandemic pathogens will be misused. *EMBO Rep.* **17**, 127–130 (2016).

113. Boddie, C., Watson, M., Ackerman, G. & Gronvall, G. K. BIOSECURITY. Assessing the bioweapons threat. *Science* **349**, 792–793 (2015).

114. Science Advisory Board for Biosecurity, N. Recommendations for the evaluation and oversight of proposed gain-of-function research. *Office of Science Policy*.

115. Ezell, B. C. & von Winterfeldt, D. Probabilistic risk analysis and bioterrorism risk. *Biosecur. Bioterror.* **7**, 108–10; discussion 111–2 (2009).

116. Ackerman, G. & Moran, K. *Bioterrorism and Threat Assessment*. https://www.wmdcommission.org/files/No22.pdf (2005).

117. Xia, Q. *et al.* Case fatality rates of COVID-19 during epidemic periods of variants of concern: A meta-analysis by continents. *Int. J. Infect. Dis.* **141**, 106950 (2024).

118. Falkenheim, J. & Kang, K. Doctorate Recipients from U.S. Universities: 2019. https://ncses.nsf.gov/pubs/nsf21308/data-tables.

119. OECD. *OECD Science, Technology and Innovation Outlook 2016*. (OECD, 2016). doi:10.1787/sti_in_outlook-2016-en.

120. LaFree, G. & Dugan, L. Introducing the Global Terrorism Database. *Terrorism and Political Violence* **19**, 181–204 (2007).

121. Grossman, D. The sentinels. *Science* **372**, 450–455 (2021).

122. Zawati, M. H. *et al.* Country Reports. *J. Law Med. Ethics* **47**, 582–704 (2019).

123. Callaway, E. Biosafety concerns for labs in the developing world. *Nature* **485**, 425 (2012).

124. Ross, E. & Harper, D. Laboratory accidents and biocontainment breaches. (2023).

125. Kilianski, A., Nuzzo, J. B. & Modjarrad, K. Gain-of-Function Research and the Relevance to Clinical Practice. *J. Infect. Dis.* **213**, 1364–1369 (2016).

126. Cao, Y. *et al.* Nanopore sequencing: a rapid solution for infectious disease epidemics. *Sci. China Life Sci.* **62**, 1101–1103 (2019).

127. Kellner, M. J., Koob, J. G., Gootenberg, J. S., Abudayyeh, O. O. & Zhang, F. SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat. Protoc.* **14**, 2986–3012 (2019).

128. Esvelt, K. *Delay, Detect, Defend: Preparing for a Future in Which Thousands Can Release New Pandemics*. (2022).

129. National Academies of Sciences, Engineering, and Medicine, Division on Earth and Life Studies &

Board on Life Sciences. *Governance of Dual Use Research in the Life Sciences: Advancing Global Consensus on Research Oversight: Proceedings of a Workshop*. (National Academies Press, 2018). doi:10.17226/25154.

130. Tetlock, P. E. *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton University Press, 2005).

Appendix A: Benefits of Virus Discovery and Pandemic Virus Identification Surveys

Initial Survey (n = 207)

1. Participant Research Areas and Affiliations
   National Authorities - 126
   - US CDC - 74
   - China CDC - 8
   - EU CDC (or member state equivalent) - 25
   - CDC equivalent - other countries - 52

   Virus Detection & Characterization Programs - 37
   - USAID Predict - 20
   - USAID DEEP VZN - 5
   - EcoHealth Alliance - 21
   - Global Virome Project - 12

   Basic Research - 118
   - Pathogens: Molecular or Cellular Biology - 90
   - Pathogens: Ecology or Evolution - 75

   Applied Research - 174
   - Vaccine Research - 54
   - Diagnostics Development - 78
   - Therapeutics Development - 38
   - Community Engagement and Training - 80
   - Policy or other preventative countermeasures - 108

2. Participant Role
   - Academia - 138
   - Industry - 6
   - Government - 70
   - Principal Investigator - 114

- ○ Staff Scientist or Postdoctoral fellow - 32
- ○ Graduate student -2
- ○ Predoctoral researcher/undergraduate student - 0

Question 1 - Discovering Virus X: Effects on Broad-Spectrum Medical Countermeasures

Question: How likely are we to have at least one approved broad-spectrum vaccine or therapeutic 10 years from now that will be effective against the next high-consequence pathogen (>1 million deaths)?

1 a) "Likelihood given our current knowledge of the global virome"
Parameter: $p_{BSV|\neg VDi}$
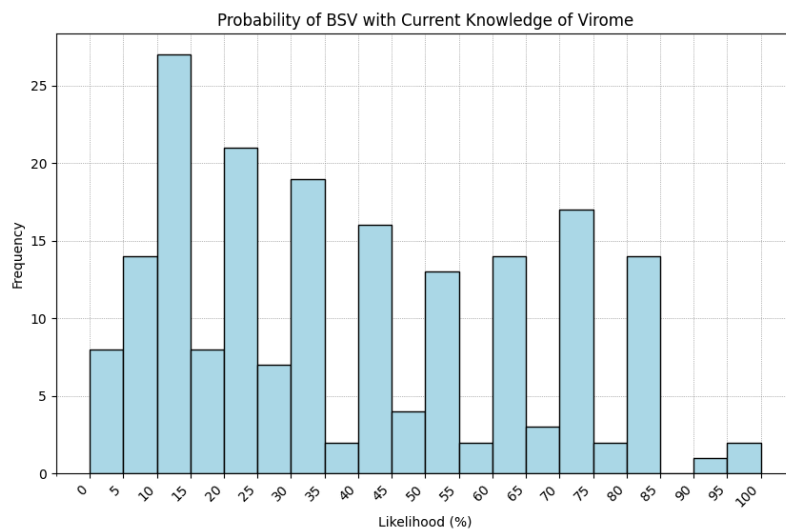
Summary Statistics:
Mean: 0.37 (90% CI [0.34, 0.4])
90% central range: [0.05 , 0.8]
Distribution: Beta (0.96, 1.63)
St. dev: 0.26
Median: 0.3

1 b) "If we discovered and sequenced 3.0x as many viruses as today"
Parameter: $p_{BSV|3x\,dis}$
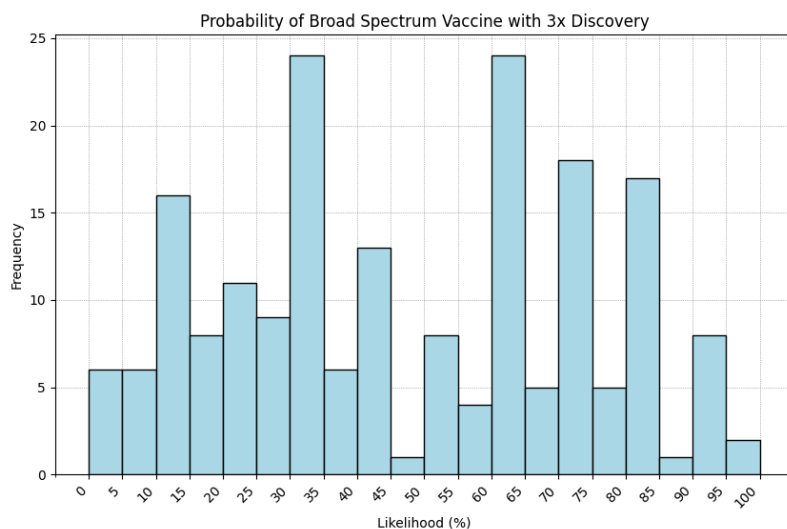
Summary Statistics:
Mean: 0.46 (90% CI [0.43, 0.49])
90% Central Range: [0.06, 0.87]
Distribution: Beta (1.17, 1.40)
St.dev: 0.26
Median: 0.4

1c) "What if we discovered and sequenced all viruses in animals?"

Parameter: $p_{BSV \mid full\ virome}$
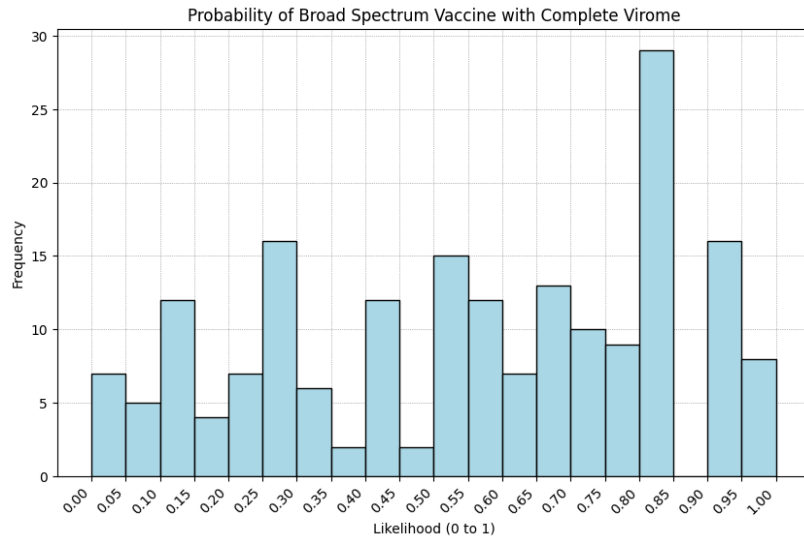
Summary Statistics:

Mean:0.55 (90% CI [0.51, 0.58])

90% Central Range: [0.08, 0.95]

Distribution: Beta (1.14,0.93)

Median: 0.6

St. dev: 0.28

Change in Probabilities

Difference between 3x discovery and baseline (current knowledge of virome)
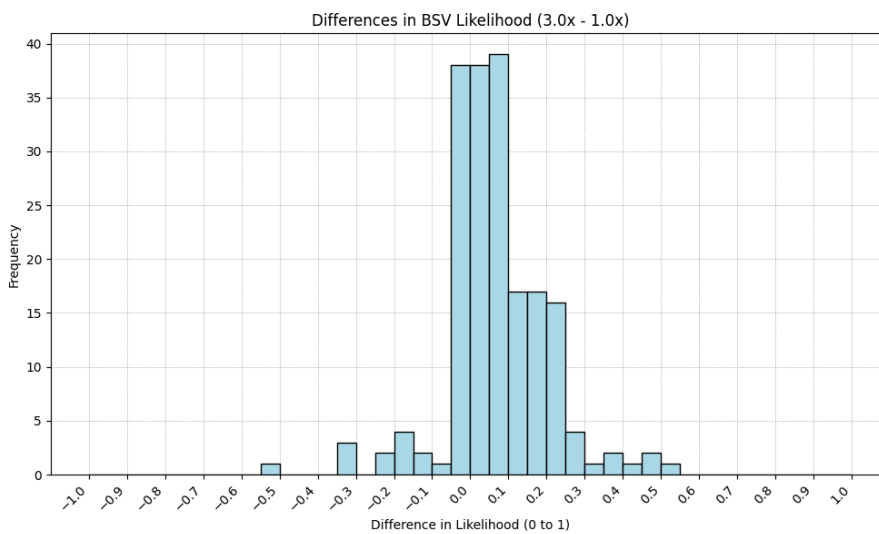
Parameter: $\Delta\,p_{BSV \mid VDi}$

Summary Statistics:

Mean:0.08 (90%CI [0.07, 0.1]

90% Central range: [-0.13, 0.29]

Median: 0.09

St. dev: 0.13

Difference between complete virome and baseline (current knowledge of virome):

Parameter: $\Delta\, p_{BSV\,|\,\text{full } VDi}$
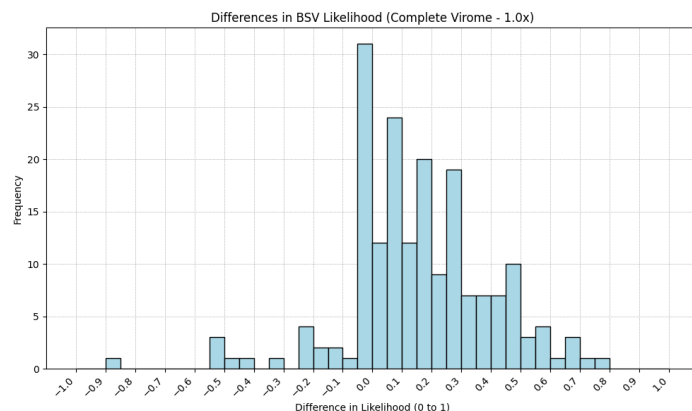
Summary Statistics:

Mean: 0.17 (90% CI [0.14, 0.19])

90% central range: [-0.2, 0.58]

Median: 0.09

St. dev: 0.13

Question 2 - Discovering Virus X: Effects on Preventative (Non-Medical) Interventions

Suppose we pursue virus discovery and use this information to improve our broad-spectrum non-medical interventions.

3x Discovery

Question: " If we sequenced 3.0x as many viruses as today, estimate the relative reduction in pandemic risk from Virus X over the next 10 years relative to a world with no additional virus discovery (almost fully effective = "99%", no help at all = "0%")."

Parameter: $\Delta p_{NPI\,|\,VDi}$
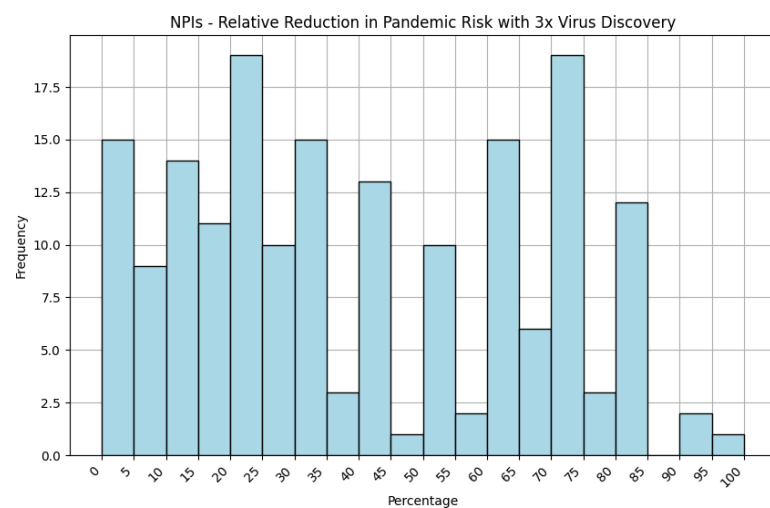
Summary Statistics:

Mean: 0.38 (90% CI [0.35,0.41])

90% central range: [0.03, 0.86]

Distribution: Beta (0.95, 1.52)

St.dev: 0.26

Median: 0.3

IQR: 0.43

Complete Virome

2b) "What if we discovered and sequenced all viruses in animals?"
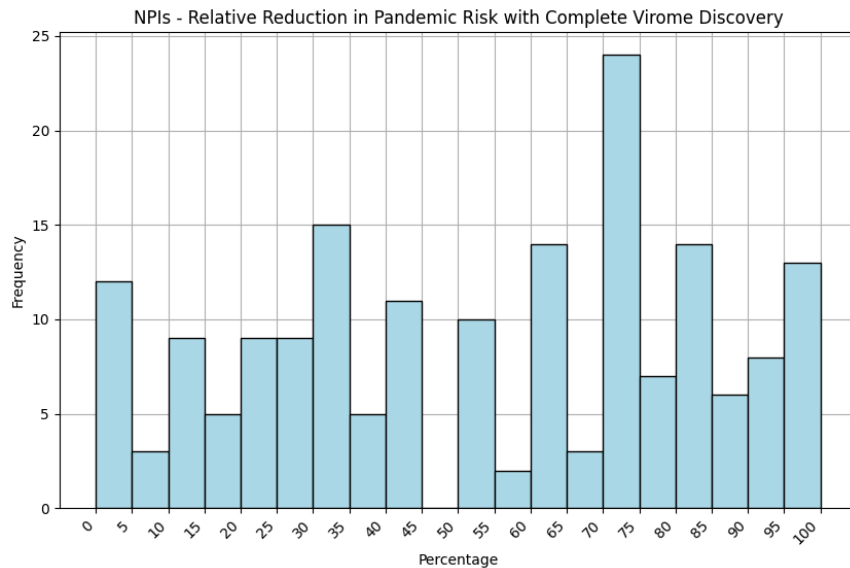
Summary Statistics:

Mean: 0.52 (90% CI [0.48, 0.55])

90% central range: [0.05, 0.95]

Distribution: Beta(0.97,0.91)

Median: 0.57

St.dev: 0.3

IQR: 0.46

Question 3 - High-Consequence Pathogens Circulating in Animal Reservoirs

Question: "How many distinct viruses capable of sustained human-to-human transmission, with the potential to cause at least 1 million deaths, do you estimate are currently circulating in animal reservoirs around the world?"

- Less than 10 viruses
- 10 - 30 viruses
- 30 - 100 viruses
- 100 - 300 viruses
- 300 - 1k viruses
- 1k - 3k viruses
- 3k - 10k viruses
- 10k - 30k viruses
- More than 30k viruses
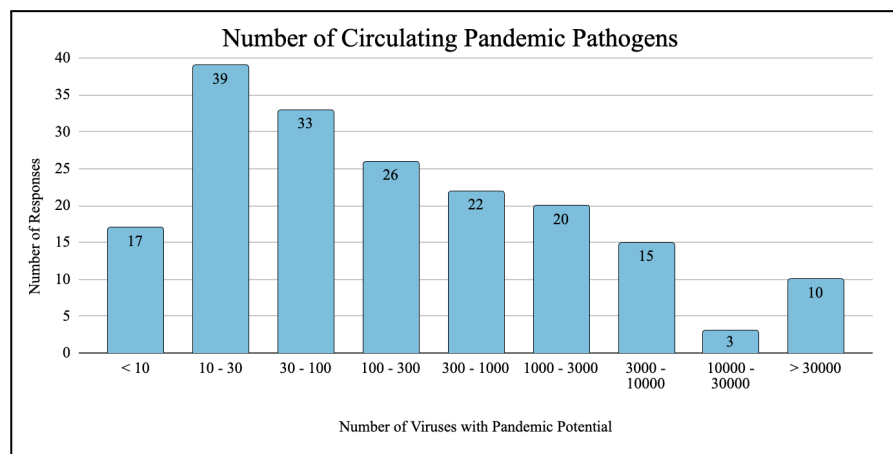
Summary Statistics:

Point estimate average: 172 90%CI [155,189]

Distribution: LogN(5.1291,0.1924)

*90% CI calculated through bootstrapping method

Question 4 - Pandemic Virus Identification: Effects on Targeted Medical Countermeasures

Question: " If Virus X is characterized as pandemic-capable prior to spillover through PVI, what is the probability the world will invest enough funds to develop targeted MCMs before an outbreak begins?"

Parameter: p_{TMCM}

Probability of Pre-Outbreak Funding for Targeted MCMs for Virus X
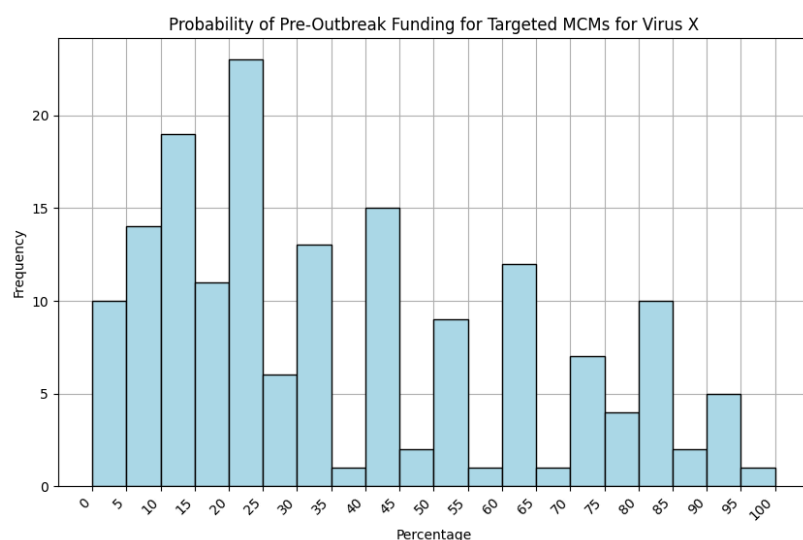
Summary Statistics:

Mean: 0.35 (90% CI [0.32,0.39])

90% central range: [0.04, 0.83]

Median: 0.30

St.dev: 0.26

IQR: 0.47

Question 5 - Acceleration of Medical Countermeasure Timelines due to PVI

Vaccines

Question: "If the world does invest sufficient funds, a Virus X vaccine would be widely available (to at least 1 billion people) _____ days sooner relative to a world in which Virus X had not been characterized and flagged as a suspected pandemic risk in advance of the outbreak".

Vaccine Timeline Acceleration due to PVI
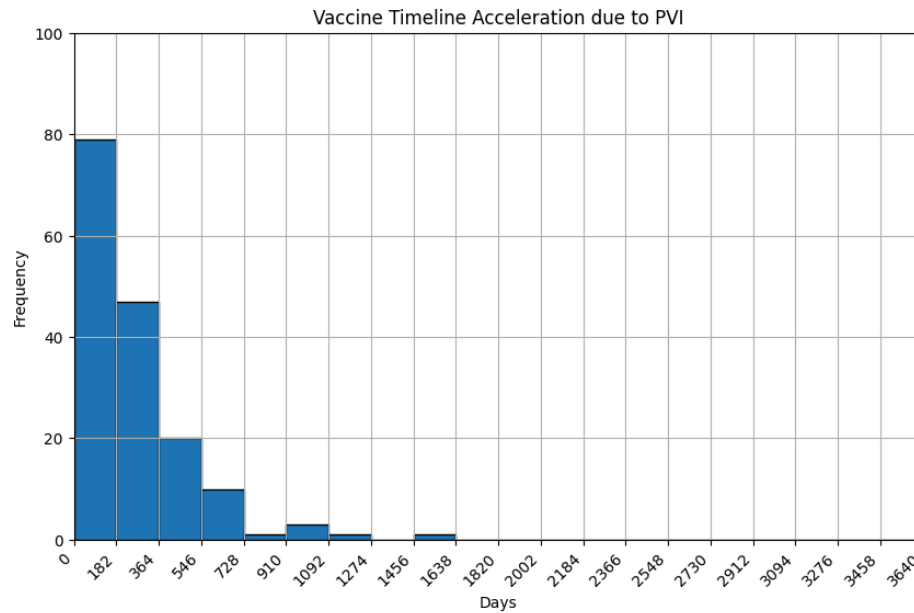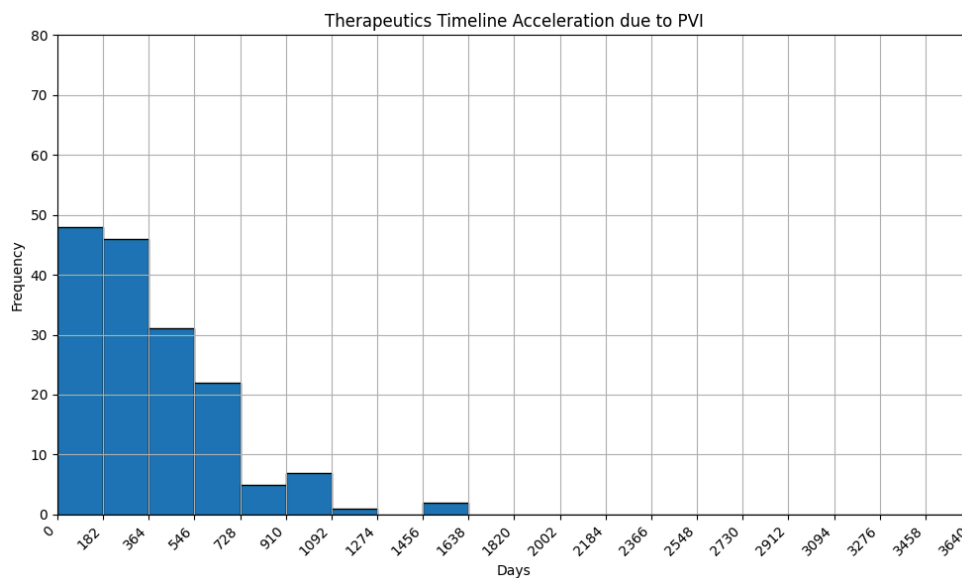
Summary Statistics:

Mean: 363 days

St.dev: 1550 days

Median: 198 days (90% CI* [180, 200])

*used bootstrapping method for CI

Therapeutics

Question: "If the world does invest sufficient funds, a Virus X therapeutic would be widely available _____ days sooner, relative to a world in which Virus X had not been characterized and flagged as a suspected pandemic risk in advance of the outbreak".

Therapeutics Timeline Acceleration due to PVI

Summary Statistics:

Mean: 421 days

St.dev: 807 days

Median: 300 days (90% CI* [250, 360])

*used bootstrapping method for CI

Question 6 - Pandemic Virus Identification: Effects on Non-Medical Interventions

Question: "How much would non-medical countermeasures targeting a characterized Virus X reduce the risk of a sustained outbreak over 10 years? Please input the reduction in risk relative to a world where Virus X is not characterized (almost fully effective = "99%", won't help at all = "0%")

Parameter: $\Delta r_{NPI\,|\,PVI}$

PVI Risk Reduction: Non Medical Interventions
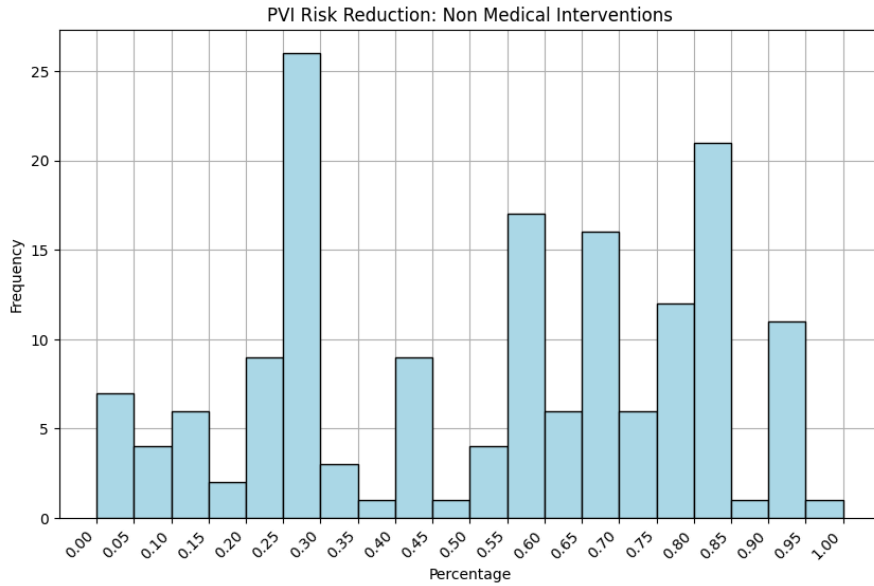
Summary Statistics:

Mean: 0.52 (90% CI [0.49, 0.56])

90% central range: [0.09, 0.90]

Distribution: Beta (1.30,1.9)

St.dev: 0.27

Median: 0.6

IQR: 0.45

Breakdown of Key Parameter Estimates based on Professional Background of Participants

x̄: mean, M: median

	National Academies	Virus Discovery and Characterization Programs	Basic Research	Applied Research
$\Delta p_{BSV\,|\,VDi}$	x̄ = 0.08, M = 0.07	x̄ = 0.13, M = 0.10	x̄ = 0.10, M = 0.09	x̄ = 0.08, M = 0.07
$\Delta r_{NPI\,|\,VDi}$	x̄ = 0.37, M = 0.3	x̄ = 0.43, M = 0.4	x̄ = 0.4, M = 0.38	x̄ = 0.38, M = 0.3
Δp_{TMCM}	x̄ = 0.35, M = 0.29	x̄ = 0.43, M = 0.4	x̄ = 0.36, M = 0.3	x̄ = 0.35, M = 0.29
$\Delta r_{NPI\,|\,PVI}$	x̄ = 0.53, M = 0.6	x̄ = 0.54, M = 0.61	x̄ = 0.54, M = 0.6	x̄ = 0.53, M = 0.6

Vaccine Timeline Acceleration	x̄ = 240, M = 180	x̄ = 224, M = 180	x̄ = 277, M = 200	x̄ = 245, M = 196
Therapeutics Timeline Acceleration	x̄ = 451, M = 300	x̄ = 318, M = 205	x̄ = 394, M = 300	x̄ = 427, M = 300

Follow-up Survey (n = 42)

Question 1 - Risk Distribution

Question: "Which of the following best match with your assessment of how pandemic risk is distributed across potential high-consequence pathogens? Note that all percentages below are as a fraction of the potentially high-consequence pathogens only; pathogens without pandemic potential are not included here.

1. All pathogens in this category are equally risky: the top 20% most risky pathogens contribute 20-30% of expected mortality

2. Narrow-tailed distribution: there is some difference between pathogens, but only a minority of the risk is concentrated in the 20% most risky pathogens (30-50% of expected mortality)

3. Fat-tailed distribution: most, but not all, of the risk is concentrated in the top 20% most risky pathogens (50-90% of expected mortality)

4. Very fat-tailed distribution: nearly all risk is concentrated in the top 20% most risky pathogens (defined as 90-100% of expected mortality)"

Distribution of Risk of Pandemic Viruses
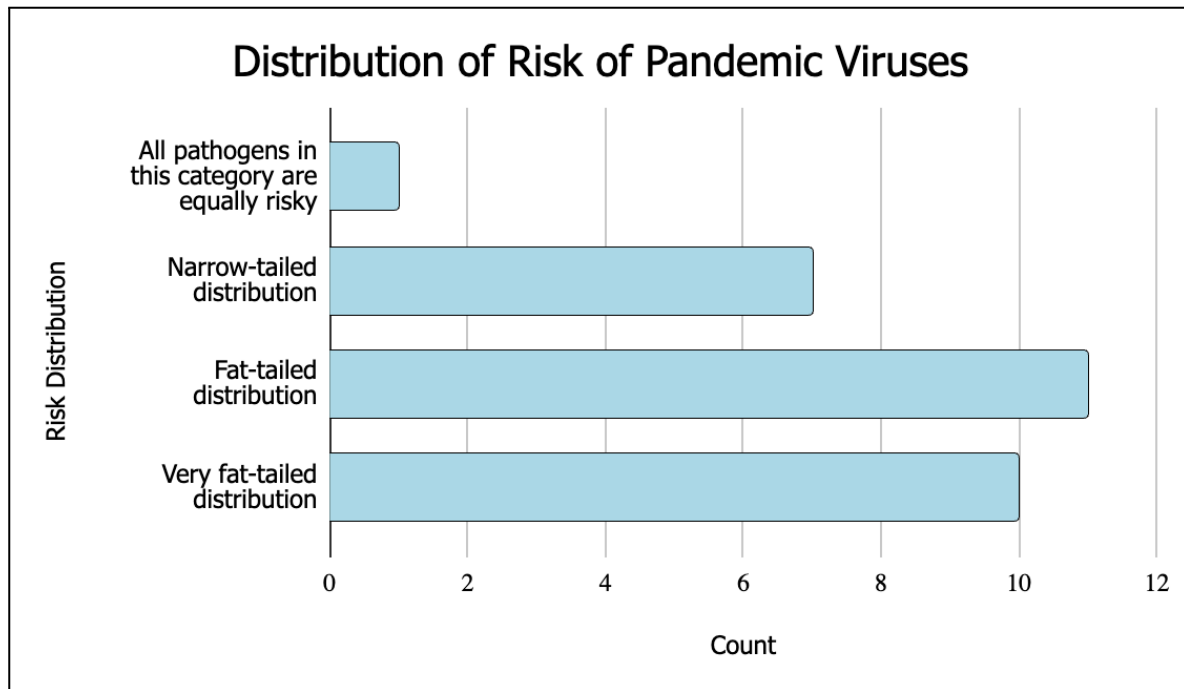
Weighted Average of Risk Distribution Results:

Weighted average was calculated using the midpoints of the expected mortality ranges and the number of respondents as weights.

$$\text{Weighted Average} = \frac{(9.5 + 7.7 + 2.8 + 0.25)}{29} = 0.70$$

The top 20% of most risky pathogens contribute to 70% of the expected mortality.

Question 2a - Vaccine Development without PVI

Question: "What is your best estimate (50th percentile) of how long it will take to develop vaccines against Virus X in a world without pandemic characterization?"
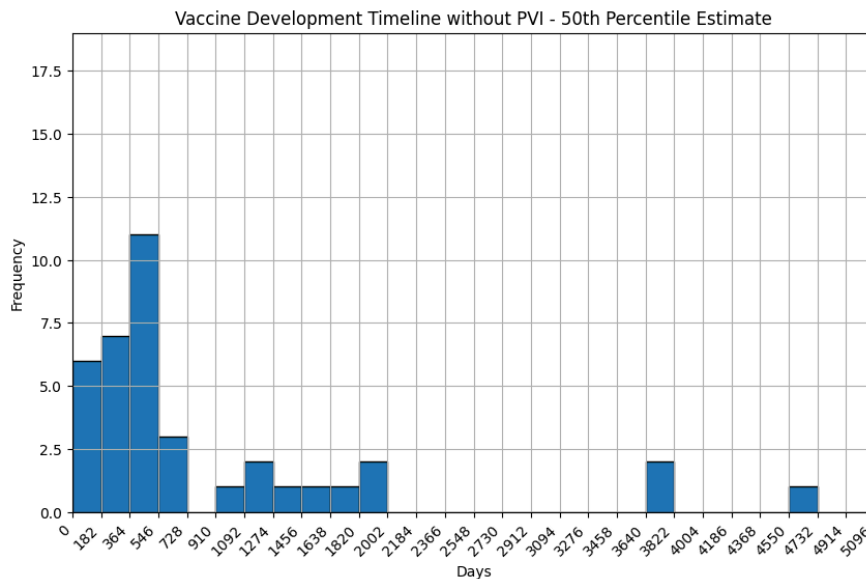
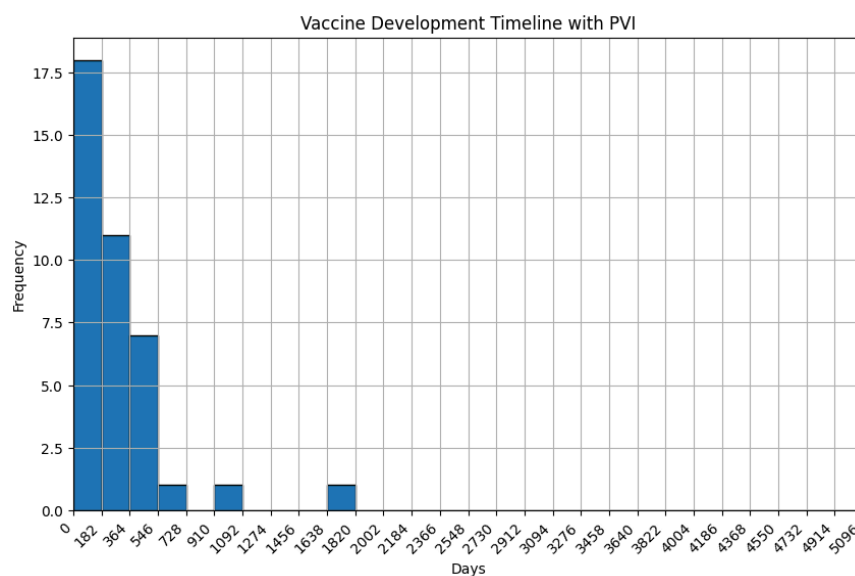**Summary Statistics:**

Mean: 858 days

St.dev: 1050 days

Median: 382 days

IQR: 756 days

*90K estimate outlier removed

Question 2b - Vaccine Development with PVI

Question: "What is your best estimate (50th percentile) of how long it will take to develop vaccines against Virus X in a world with pandemic characterization?"

Summary Statistics:

Mean: 287 days

St.dev: 303 days

Median: 200 days

IQR: 213 days

Note: Median differences are reported in followup surveys regarding timelines.

Question 3 - Funding

Parameter: p_{TCM}

Question:

1. "If only one (1) virus is identified in the laboratory as a suspected pandemic threat over the next ten years, what is the likelihood that the world invests enough funds to develop targeted medical countermeasures against this virus?"

2. "If ten (10) different viruses are identified in the laboratory as suspected pandemic threats over the next ten years, what is the likelihood that the world invests enough funds to develop targeted countermeasures against at least one (1) of these viruses? "

3. "If ten (10) different viruses are identified in the laboratory as suspected pandemic threats over the next ten years, what is the likelihood that the world invests enough funds to develop targeted countermeasures against all ten (10) viruses?"

4. "If a hundred (100) different viruses are identified in the laboratory as suspected pandemic threats over the next ten years, what is the likelihood that the world invests enough funds to develop targeted countermeasures against at least one (1) of these viruses?

5. "If a hundred (100) different viruses are identified in the laboratory as suspected pandemic threats over the next ten years, what is the likelihood that the world invests enough funds to develop targeted countermeasures against all one hundred (100) viruses?"

Likelihood of Investment into 1 Targeted Countermeasure given PVI
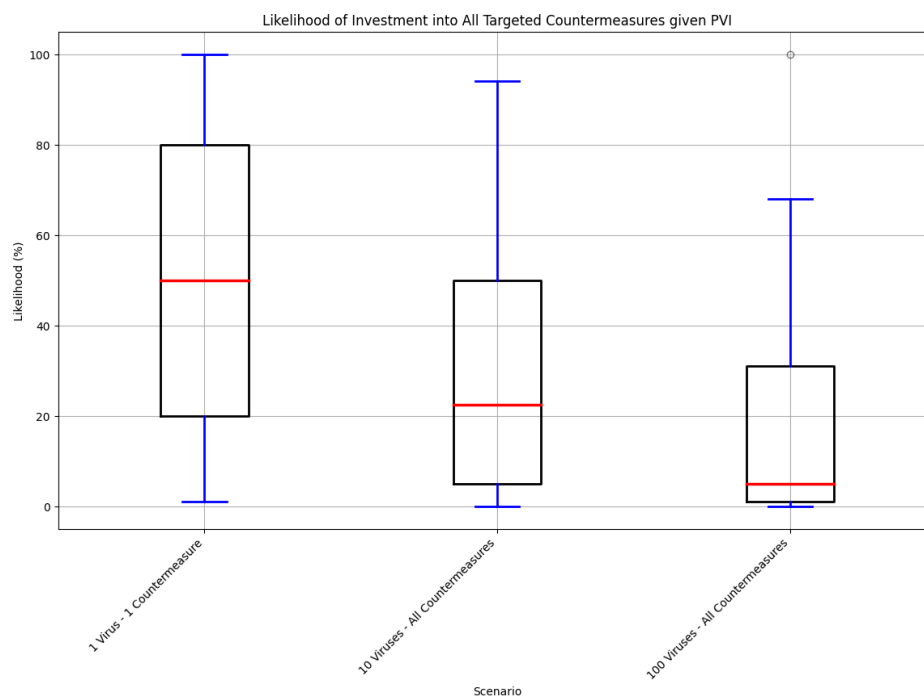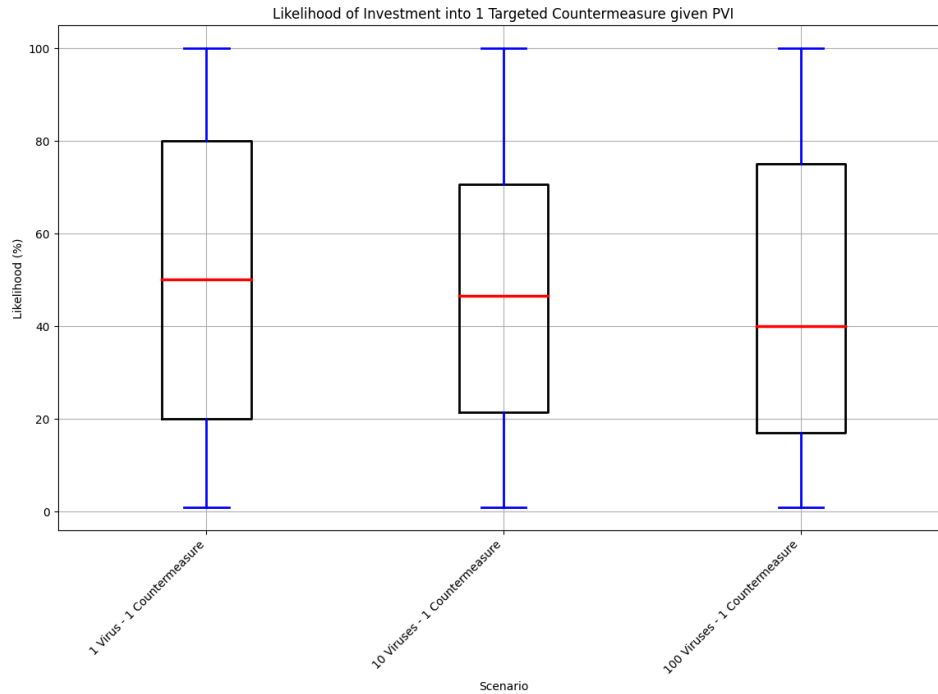
Summary Statistics

	1 Virus - 1 MCM (%)	10 Viruses - 10 MCMs (%)	100 Viruses - 100 MCMs (%)
Mean	0.46	0.32	0.18
90% Central Range	(0.02, 0.94)	(0.01, 0.81)	(0.0, 0.55)
St.dev	0.32	0.29	0.25
Median	0.5	0.2	0.05
Distribution	Beta (0.72,0.84)	Beta (0.53,1.18)	Beta (0.27,1.20)

Question 4 - PVI

Question: "Considering only the risk of a pandemic of zoonotic origin, do you think pandemic characterization on net decreases the risk from Virus X, increases the risk, or does not change the level of risk?

A. Pandemic characterization decreases the risk from pandemics of zoonotic origin, on net
B. Pandemic characterization does not make a difference to the risk from pandemics of zoonotic origin
C. Pandemic characterization increases the risk from pandemics of zoonotic origin, on net"
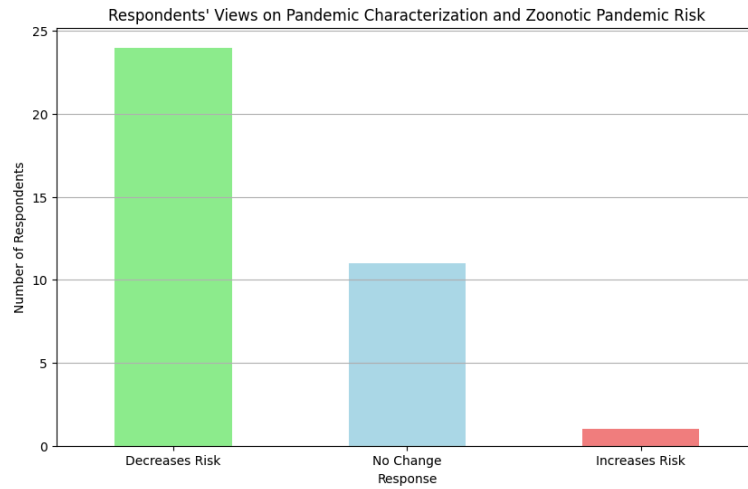
Model Parameters

The table below presents all the parameters of our model, noting the parameter name, point estimate, distribution for parameters estimated with survey data, as well as the description and source of the data for each parameter.

Table S1: Model parameters and relevant information

Parameter	Point Estimate	Description	Source
$E[Harm_{base}]$	119,631 deaths	Baseline expected number of deaths from natural pandemics per virus per decade	Model
$E[Harm_{VDi}]$	96,901 deaths	Expected number of deaths from natural pandemics per virus per decade given viral discovery efforts triple the number of viruses we know about	Model
$E[Harm_{PVI}]$	62,208 deaths	Expected number of deaths from natural pandemics per virus per decade given we characterize the pandemic potential of the virus	Model
v_{total}	172	Total number of pandemic-capable viruses circulating in the world	Survey

$p_{BSV \mid \neg VDi}$	37.1%	Baseline probability at least one approved broad-spectrum vaccine for use against Virus X prior to spillover without any additional VDi	Survey
$p_{BSV \mid 3x\,dis}$	46%	Probability discovering 3 times as many viruses as today results in at least one approved broad-spectrum vaccine for use against Virus X prior to spillover	Survey
$p_{BSV \mid full\,virome}$	55%	Probability discovering all viruses in animals results in at least one approved broad-spectrum vaccine for use against Virus X prior to spillover	Survey
$\Delta p_{BSV \mid VDi}$	9%	Relative increase in likelihood of an approved broad-spectrum vaccine prior to Virus X spillover given 3x virus discovery	$p_{BSV \mid 3x\,dis} - p_{BSV \mid \neg VDi}$
$\Delta p_{BSV \mid full\,VDi}$	17.9%	Relative increase in likelihood of an approved broad-spectrum vaccine prior to Virus X spillover given complete virome discovery	$p_{BSV \mid full\,VDi} - p_{BSV \mid \neg VDi}$
$\Delta p_{NPI \mid VDi}$	14%	Relative change in probability non-pharmaceutical interventions would be directed towards Virus X hotspot due to 3x viral discovery	$(p_{BSV \mid 3x\,dis} - p_{BSV \mid \neg VDi})/\,1 - p_{BSV \mid \neg VDi})$
$\Delta p_{NPI \mid full\,VDi}$	55%	Relative change in probability non-pharmaceutical interventions would be directed towards Virus X hotspot due to complete virome discovery	Survey (assumed to be equal to $p_{BSV \mid full\,virome}$)
m_{BSV}	41.6%	reduction in harm due to the availability of a broad-spectrum vaccine for Virus X prior to spillover	Literature
$\Delta r_{NPI \mid VDi}$	38%	Relative reduction in harm due to increased targeted of non-pharmaceutical interventions towards Virus X hotspot due to 3x viral discovery	Survey
$\Delta r_{NPI \mid full\,VDi}$	55%	Relative reduction in harm due to increased targeted of non-pharmaceutical interventions towards Virus X hotspot due to complete virome	Survey
p_{TMCM}	35%	Probability of targeted Virus X medical countermeasures receiving sufficient funds for development prior to spillover and outbreak	Survey
$p_{NPI \mid PVI}$	48%	Probability PVI results in targeted non-pharmaceutical interventions at the Virus X hotspot	Survey
Δm_{TV}	54%	Relative reduction in harm due to earlier release of targeted vaccines	Literature + Survey
Δm_{TT}	9%	Relative reduction in harm due to earlier release of targeted therapeutics	Literature + Survey
$\Delta r_{NPI \mid PVI}$	52%	Relative reduction in risk due to improved non-pharmaceutical interventions from PVI	Survey

Appendix B: Synthetic Virology Capabilities Survey

Capabilities Survey (n = 24)

1. Current Role
   - Academia - 19
   - Industry - 3
   - Principal Investigator - 12
   - Staff Scientist or Postdoctoral Fellow - 4
   - Graduate Student - 2
   - Predoctoral researcher/undergraduate student - 0
2. Field of Expertise
   - Virology - 5
   - Synthetic Biology - 14
   - Molecular Biology - 11
   - Cellular Biology - 2
   - Other - 6
     - Quantitative Genetics and Plant Breeding
     - Immunology
     - Computational Biology
     - Genetics and Evolutionary Biology
     - Genomics and Biotechnology
     - Microbiology - biochemistry

Question 1 - Reverse Genetics Timelines

a) How long would it take a researcher who is skilled in synthetic virology to successfully perform reverse genetics for a newly discovered virus of the following families if they need to conduct research to **adapt a protocol developed** for a different virus from the same family?

b) How long would it take a researcher who is skilled in synthetic virology to successfully perform reverse genetics for a newly discovered virus of the following families if they had a **validated, step-by-step protocol** for the virus in question?

	Reverse Genetics: Adapt a Protocol (Months)	Reverse Genetics: Validated Protocol Available (Months)
Influenza Virus	x̄ =3.34 , M = 3	x̄ =2.95 , M = 2
Coronavirus	x̄ =3.88 , M = 3	x̄ =2.07 , M = 2
Paramyxovirus	x̄ =4.18 , M = 3	x̄ =2.23 , M = 2
Across All Viruses	x̄ =3.80 , M = 3	x̄ =2.43 , M = 2

*excluded outliers ≥ 60

Question 2 - Evaluating Current Scope of Capabilities (n = 14)

We hope to estimate how many people possess this level of skill and tacit knowledge. In the past 20 years, approximately 15,000 people earned a PhD in virology. Another approximately 250,000* earned doctorates in fields such as molecular biology, microbiology, bioengineering, cell biology, and biotechnology over the past 20 years (NSF, 2019; OECD, 2016).

a) How many researchers do you think can successfully perform reverse genetics for a new virus in less than a year if they **need to conduct research to adapt a protocol**?

b) How many researchers do you think can successfully perform reverse genetics in less than a year when using a **validated, step-by-step protocol** for the virus in question?

	Adapt a Protocol (# of People)	Validated Protocol Available (# of People)
Influenza Virus	x̄ =20197 , M = 4000	x̄ =25861 , M = 10000
Coronavirus	x̄ =20234 , M = 4000	x̄ =25790 , M = 10000
Paramyxovirus	x̄ =15877 , M = 2000	x̄ =24181, M = 7500
Across All Viruses	x̄ =18769 , M = 3000	x̄ =25277 , M = 10000

*Note: Averages quite skewed by 1 participant who estimated 120,000 people could adapt a protocol to perform reverse genetics within a year, and 200,000 people could follow a validated protocol within a year. This affected individual virus means much more than overall means

Question 3 - Acquiring Capabilities (n = 14)

Another way to assess the importance of detailed step-by-step protocols is to estimate how many people could successfully perform a technique with additional practice or specialized training. For example, a bacteriologist might first need to master mammalian tissue culture and viral passaging before they can perform reverse genetics, but the requisite level of mastery might be lower if a protocol is sufficiently detailed. Approximately 1.26 million individuals have received a doctorate in the life sciences in the past 20 years. Approximately 66,000 of those doctorates are in molecular biology.

a) How many scientists could not reliably perform reverse genetics today, but could succeed given 1 to 4 years of training (e.g. a postdoctoral fellowship) **with a detailed protocol**?

b) How many scientists could not reliably perform reverse genetics today, but could succeed given 1 to 4 years of training (e.g. a postdoctoral fellowship) **without a detailed protocol**?

	Adapt a Protocol (# of People)	Validated Protocol Available (# of People)
Influenza Virus	$\bar{x}$ =86075 , M = 25000	$\bar{x}$ =150007 , M = 55000
Coronavirus	$\bar{x}$ =86432 , M = 25000	$\bar{x}$ =150728 , M = 55000
Paramyxovirus	$\bar{x}$ =84501 , M = 20000	$\bar{x}$ =150722 , M = 55000
Across All Viruses	$\bar{x}$ =85670 , M = 25000	$\bar{x}$ =150486 , M = 55000

*Note: Averages were quite skewed by 1 participant who estimated 700,000 people could adapt a protocol to perform reverse genetics within 1-4 years, and 1,260,000 people could follow a validated protocol to perform reverse genetics within 1-4 years.

# Appendix C: Bioterrorism Intention Surveys

Terrorism Survey (n = 111)

Question 1 - Overall Risk Estimates

Imagine that a novel pandemic capable virus is publicly identified. The genetic blueprint (the whole genome sequence) and detailed step-by-step assembly instructions (a reverse genetics protocol) for this pathogen are available online. This means that someone with enough expertise and resources could make live infectious samples of the virus.

Suppose that approximately 10,000 people have the ability to make enough virus to infect ten people, which is enough to reliably cause a pandemic (Lipsitch et al., 2003).

a) What is the likelihood that this virus will be used to deliberately cause a pandemic within the **next 10 years (2023-2033)**?
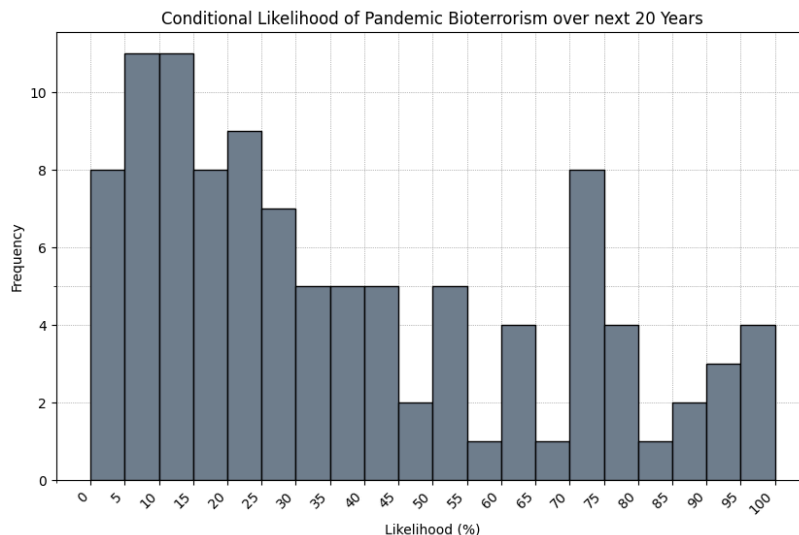

Conditional Likelihood of Pandemic Bioterrorism over next 10 Years
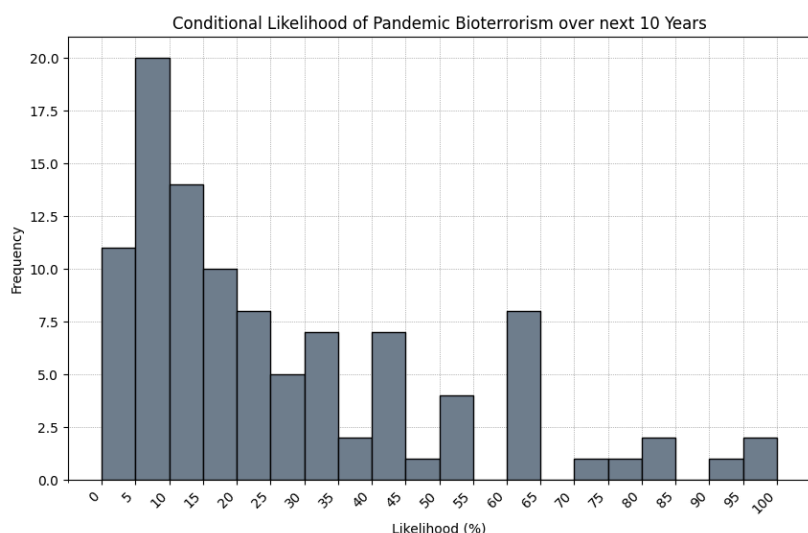
Summary Statistics:

Mean: 25.32% 90%CI (21.41, 29.22)

Median: 19.0%

Standard Deviation: 24.00%

b) What is the likelihood that this virus will be used to deliberately cause a pandemic within the **next 20 years (2023-2043)?**

Summary Statistics:
Mean: 35.88%, 90% CI (31.17, 40.60)
Median: 25.0%
Standard Deviation: 28.97%

Question 2 - Highly Skilled Actors

Background: Individuals with expert-level skill in molecular biology and virology can already perform the techniques and access materials necessary to synthesize live viruses. We characterize this group as "highly skilled actors", who would be able to synthesize live samples of a pandemic-capable virus within a year, given the genome sequence and reverse genetics protocol.

Historical example: Seiichi Endo was a graduate-trained virologist from Kyoto University with a specialty in genetic engineering. He joined the Aum Shinrikyo apocalyptic cult, sought samples of Ebolavirus for use against civilians, and was executed for committing mass murder with chemical weapons. A modern terrorist with his educational background could assemble most viruses from synthetic DNA.

Question: If there are 10,000 highly skilled actors globally, how many would spend up to a year of effort to cause global harm using such a synthesized virus?

Highly Skilled Actors: Likelihood of Intent for Large Scale Harm
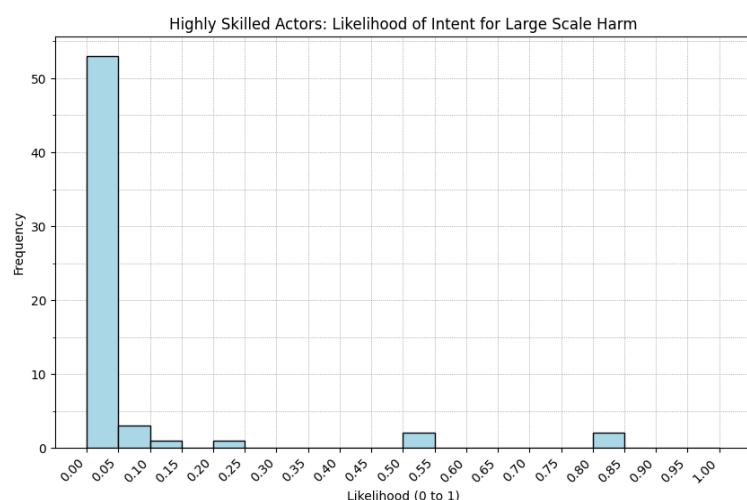
Summary Statistics:

Mean: 5.16%
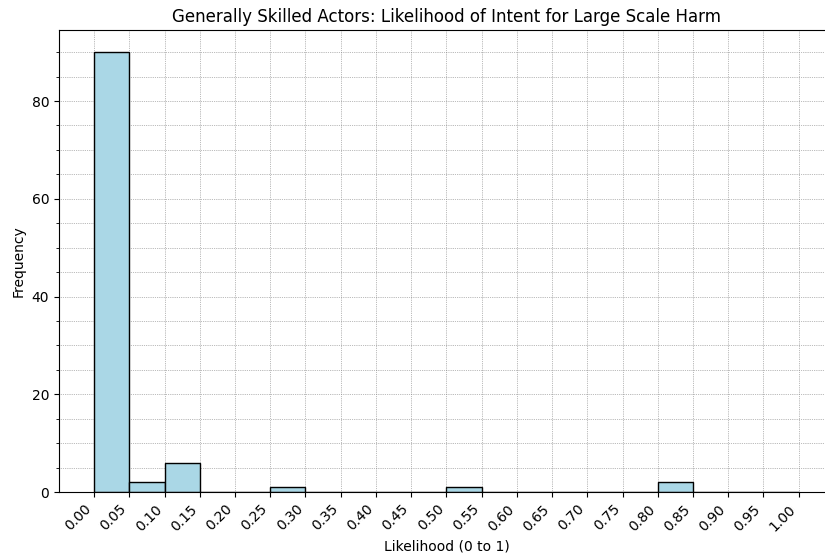
90%CI (0%, 48%)

Median: 0.1%

Standard Deviation: 16.56%

*Converted to probabilities by dividing answers by 10,000

Question 3 - Generally Skilled Actors

Background: Individuals who have some technical competency in a different area of the life sciences would need more specialized training to acquire the specific capabilities needed to synthesize viruses. This would require them to undertake the equivalent of postdoctoral training or a new job in a different specialty. We characterize this group as "generally skilled actors", who would be able to synthesize live samples of a pandemic-capable virus within 1 to 4 years given the genome sequence and reverse genetics protocol.

Historical example: Abdur Rauf Ahmed was a mid-level Pakistani government microbiologist who was recruited by Al-Qaeda to develop a biological weapons program. Documents captured in Afghanistan in 2001 revealed his mission to obtain anthrax spores and equipment for biological weapons. He was not specifically trained to work with viruses.

Question: For every 10,000 generally skilled actors, how many do you think would spend 1-4 years (e.g. a postdoc-equivalent) in order to cause indiscriminate global harm?

Generally Skilled Actors: Likelihood of Intent for Large Scale Harm

Summary Statistics:

Mean: 3.4%

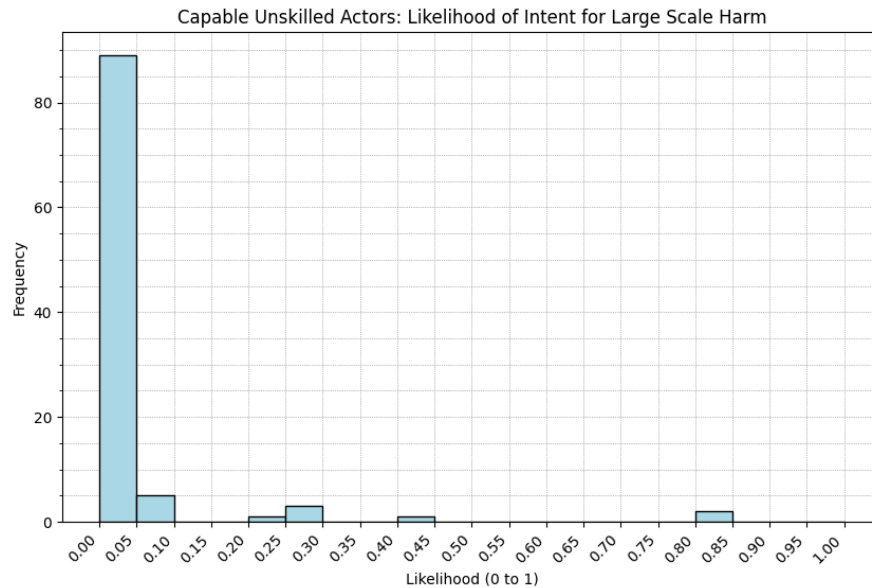90% central range (0%, 10%)

Median: 0.1%

Standard Deviation: 12.42%

*Converted to probabilities by dividing answers by 10,000

Question 4 - Capable Unskilled Actors

Background: The vast majority of individuals have none of the relevant skills or knowledge necessary to assemble viruses. However, with dedicated effort, actors with the intellectual acumen and personal discipline to obtain advanced degrees in non-biological disciplines could presumably acquire the necessary skills with several years of dedicated effort. We characterize this group as "capable unskilled actors", who would be able to synthesize live samples of a pandemic-capable virus between 4 to 10 years given the genome sequence and reverse genetics protocol.

Historical example: Ted Kaczynski, the Unabomber, was a Berkeley mathematics professor who spent over a decade engaging in a mail-bombing campaign with the goal of bringing down the industrial system.

Question: For every 10,000 capable unskilled actors, how many do you think are motivated enough to spend 4 to 10 years – approximately equivalent to earning a new PhD in a different field – in order to acquire the necessary skills and cause indiscriminate global harm?

Capable Unskilled Actors: Likelihood of Intent for Large Scale Harm

Summary Statistics:

Mean: 3.6%

90% central range (0%, 25%)

Median: 0.1%

Standard Deviation: 12.42%

*Converted to probabilities by dividing answers by 10,000

Question 4 - Efficacy of Counterterrorism Efforts

Background: To successfully carry out a bioterrorist attack using a pandemic agent, an actor would need to acquire the necessary materials, assemble the virus or obtain and isolate live samples, and successfully disseminate the virus - without being reported by coworkers or noticed by intelligence services.

Question: If someone with all the necessary resources and skills and resources attempts to make and release a pandemic virus, how likely is it that they will be detected and prevented by current counterterrorism efforts: whistleblowers, intelligence services, law enforcement or any combination of similar factors?

Likelihood of Prevention through Counterterrorism
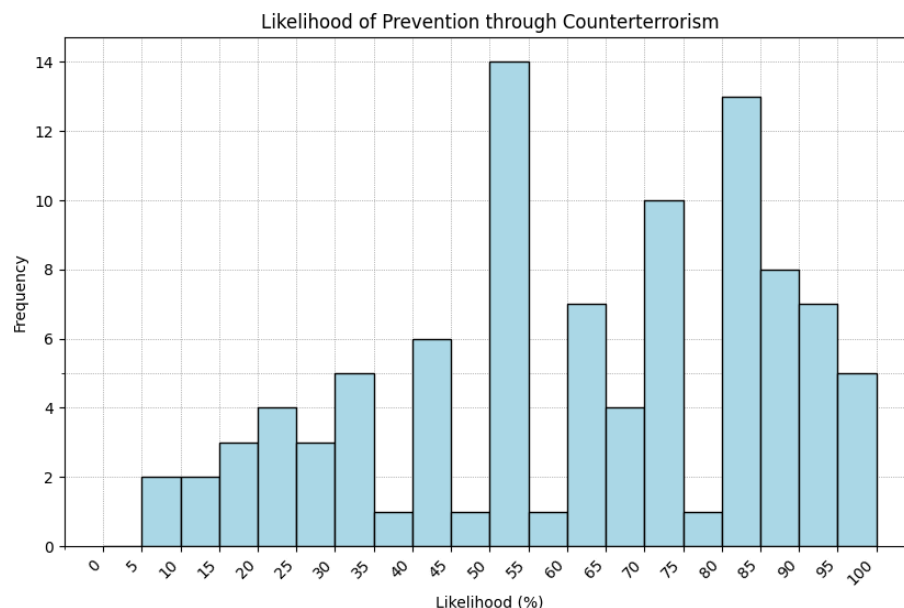
Summary Statistics:

Mean: 59.7%

90% central range (19%, 93%)

Median: 64%

Standard Deviation: 24.9%

Question 5 - Organized Groups

Background: In the past fifty years, several organized groups demonstrated a clear willingness to use biological weapons. Notable examples include Aum Shinrikyo, which attempted to use non-transmissible biological weapons to commit mass murder and sought to obtain samples of ebolavirus; the Soviet Union, which supported a large-scale biological weapons program that allegedly developed vaccine-resistant smallpox; and more recent groups such as ISIS and Al Qaeda.

Question: In your estimation, how many current organized and active groups (both non-state and state-sponsored) would attempt to deliberately release a pandemic agent, if they thought it feasible?

Assume they have access to the genetic blueprint (the whole genome sequence) and detailed step-by-step assembly instructions (a reverse genetics protocol) for the agent.


Summary Statistics:
Mean: 65
90%CI (1, 332)
Median: 5
Standard Deviation: 210
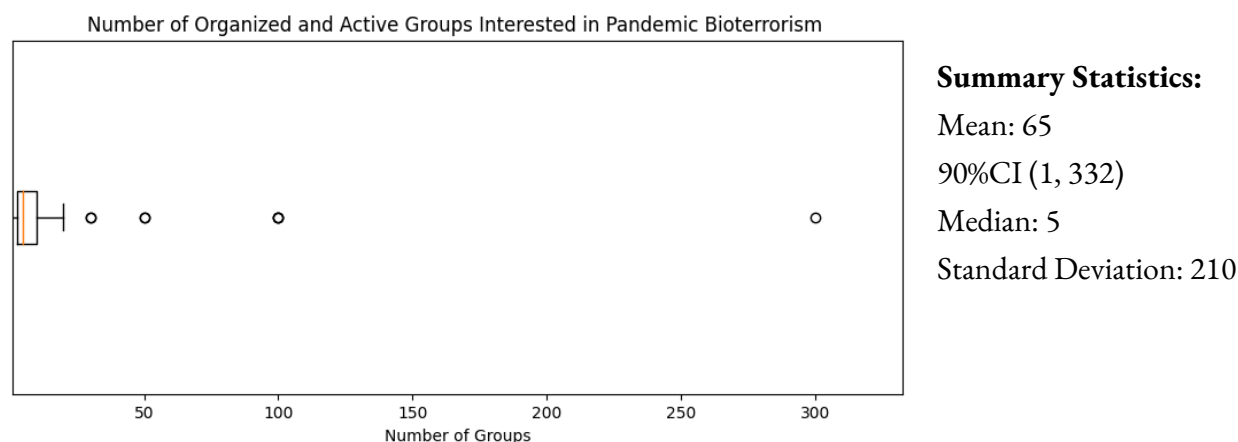
Followup Terrorism Survey (n = 106)

Participant Research Areas and Affiliations

SUMMARY STATISTICS	
Political Science and International Relations	41
Terrorism and Counterterrorism Studies	79

|	
Intelligence and Security Studies	15
Criminology	12
Other (specify below)	20

Question 1 - Types of Actors and Ideologies Driving Bioterrorism Risks

a) Which actors are you most concerned about engaging in pandemic bioterrorism? (1= least concerned, 5 = most concerned)
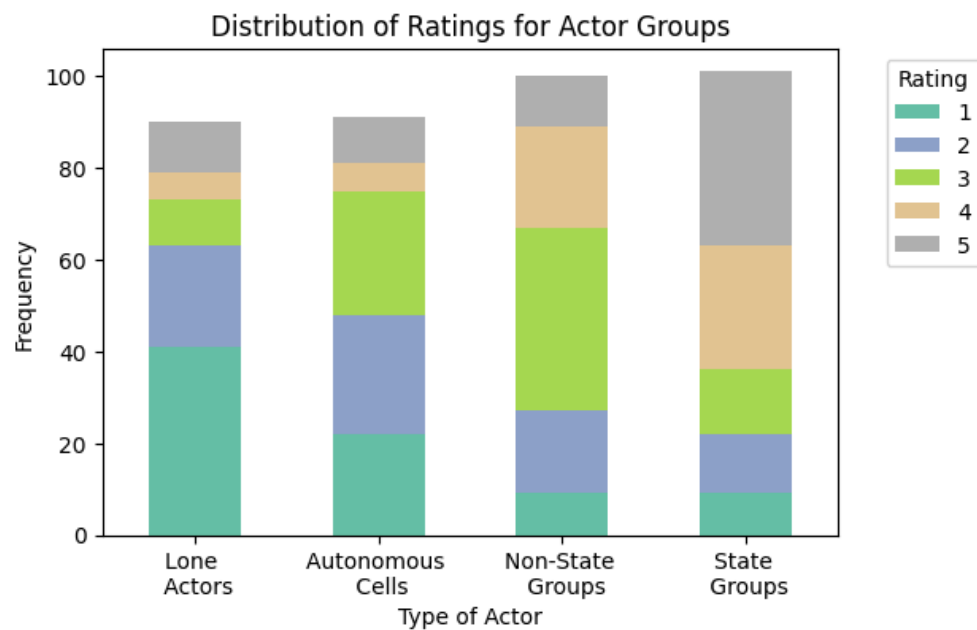
	Lone Actors	Autonomous Cells	Non-State Groups	States; State Sponsored Groups
Average	2.13	2.52	3.08	3.71
Median	2	2	3	4

b) Which particular motivations or ideologies are of concern? Rate the following categories according to your perception of the risk they pose as drivers of bioterrorism. Please assign 1 star to indicate the lowest level of risk and 5 stars to indicate the highest level of risk.


Distribution of Ratings for Bio-terrorism Motivations
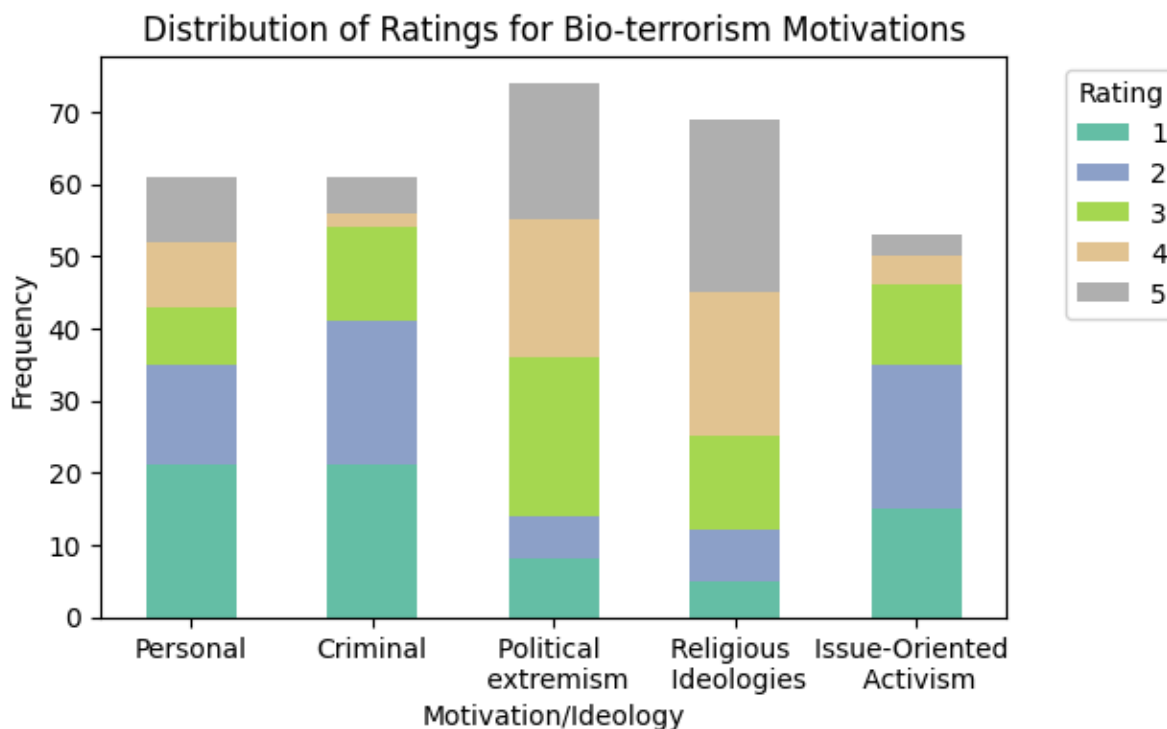
	Personal	Criminal	Political Extremism	Religious-based Ideologies	Issue-Oriented Activism
Average	2.15	1.86	3.24	3.53	2.24
Median	1.75	1.5	3.25	4	2

Question 2 - Risks from Highly Skilled Actors

Background: Synthetic biology experts estimate there are on average approximately 25,000 individuals who have the ability to reliably assemble the pandemic-capable virus within a year unless they are detected and stopped. It costs approximately $50K-$100K USD to acquire the necessary materials and equipment.

Historical example: Seiichi Endo was a graduate-trained virologist from Kyoto University with a specialty in genetic engineering. He joined the Aum Shinrikyo apocalyptic cult, sought samples of Ebola virus for use against civilians, and was executed for committing mass murder with chemical weapons. A modern terrorist with his educational background could likely assemble most viruses from synthetic DNA within months.

Question: What is the probability that at least one of these highly skilled actors will attempt to cause a pandemic within the next 10 years (2024-2034)?


Probability of a Highly Skilled Actor Attempting to Cause a Pandemic Within the Next 10 Years
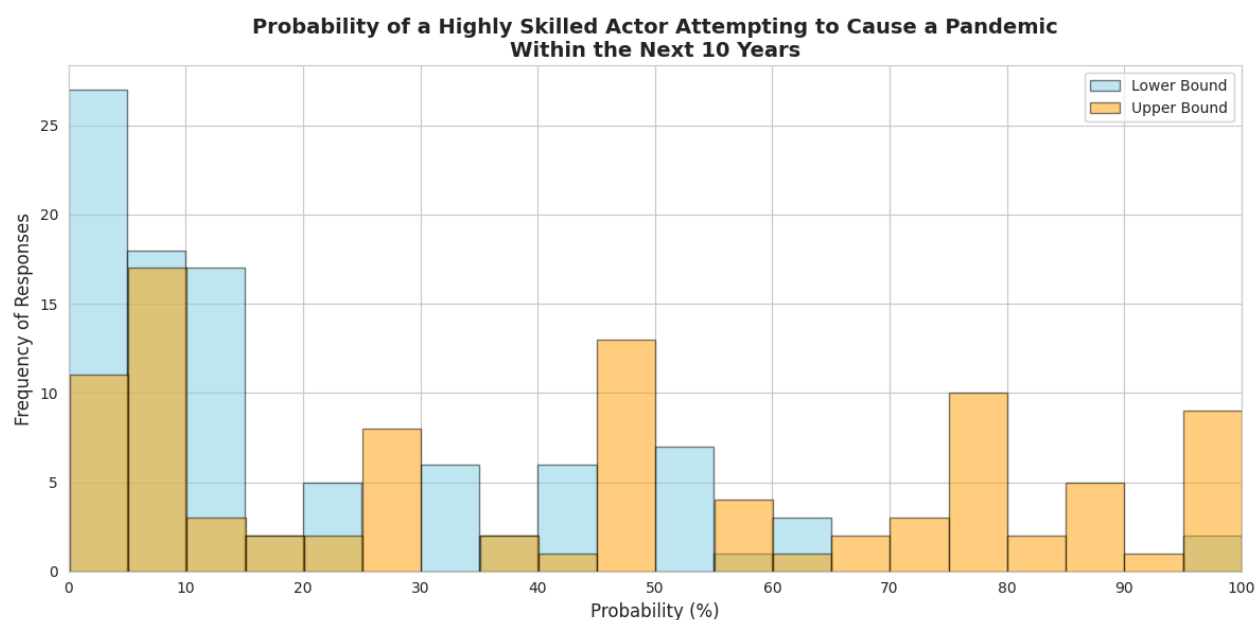
x̄: mean , M: median

Lower Bound (5th percentile)	x̄ =17% , M =10%
Upper Bound (95th percentile)	x̄ = 46% , M = 50%

Question 3 - Risks from Generally Skilled Actors

Background: Many more people with training in the life sciences lack the skills to assemble a virus, but could learn to do so if they were willing to devote 1-4 years. Synthetic biologists estimate there are 160,000 semi-skilled actors who have the ability to reliably assemble a pandemic-capable virus given 1-4 years of training, equivalent to a postdoctoral fellowship.

Historical example: Abdur Rauf Ahmed was a mid-level Pakistani government microbiologist who was recruited by Al-Qaeda to develop a biological weapons program. Documents captured in Afghanistan in 2001 revealed his mission to obtain anthrax spores and equipment for biological weapons. He was not specifically trained to work with viruses.

Question: What is the probability that at least one of these 160,000 generally skilled actors will attempt to cause a pandemic within the next 10 years (2024-2034)?

Probability of a Generally Skilled Actor Attempting to Cause a Pandemic Within the Next 10 Years
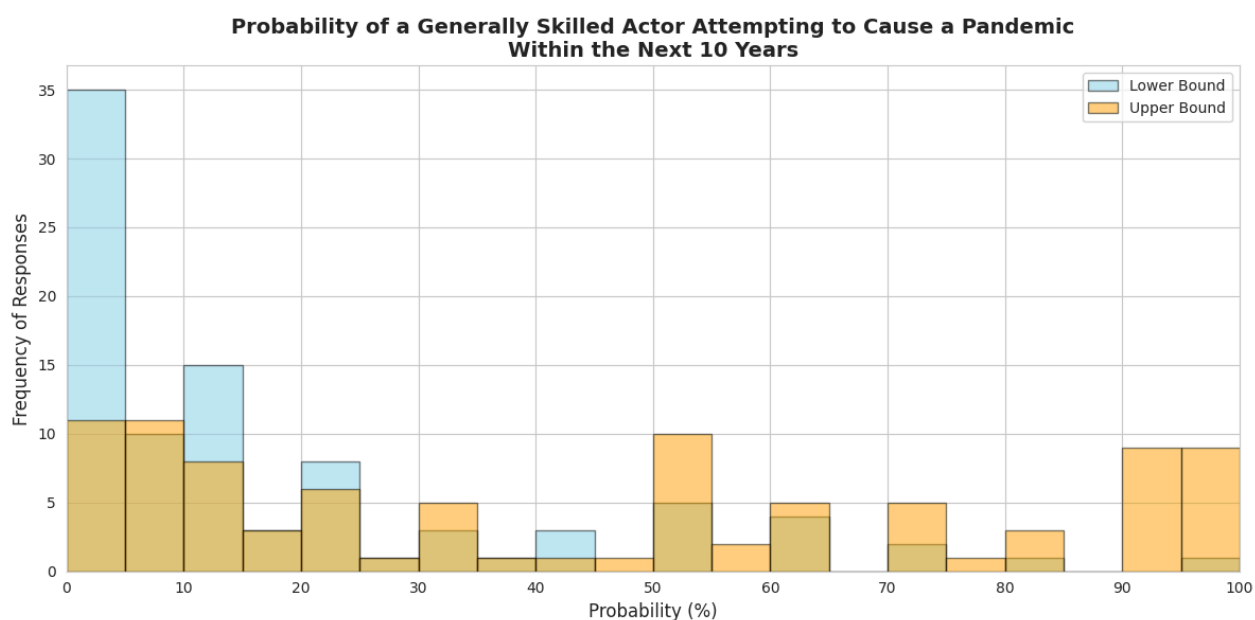
x̄: mean , M: median

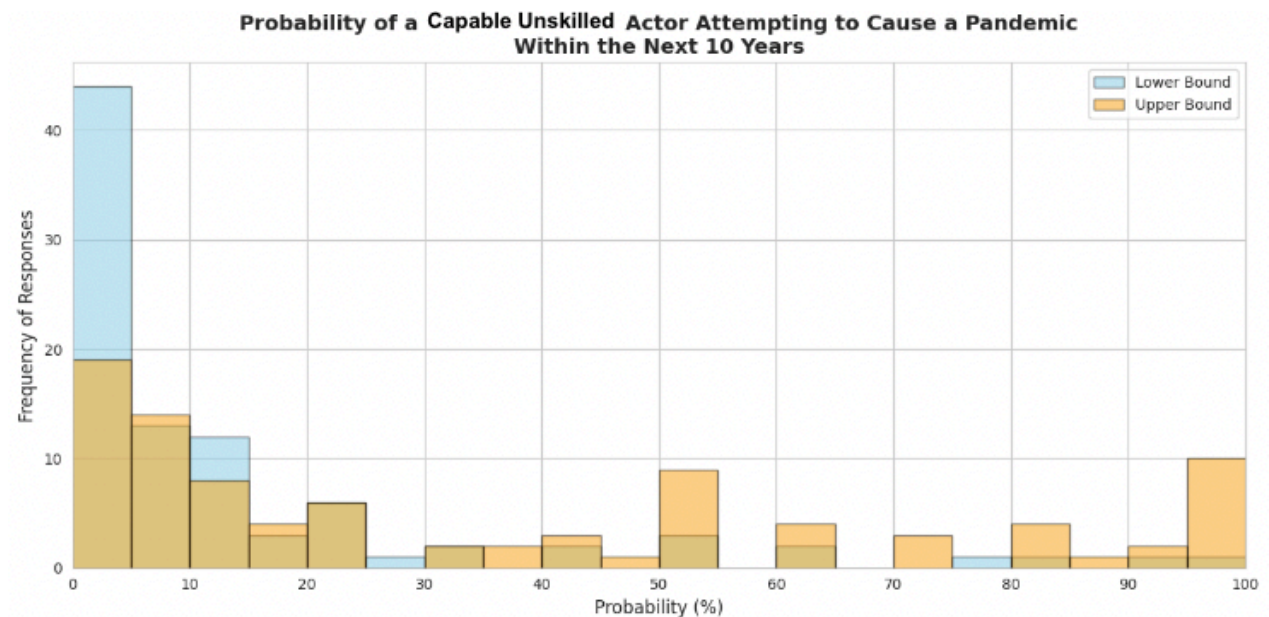| Lower Bound (5th percentile)	x̄ =16% , M =10%
Upper Bound (95th percentile)	x̄ = 42% , M = 37.5%

Question 4 - Risks from Capable Unskilled Actors

Background: Still more people have the talent and dedication to obtain a PhD in virology, but have never previously trained in the life sciences. They could potentially acquire the necessary skills to assemble a virus if they were to change careers and devote 4-10 years of effort. Globally, there are approximately 4.3 million people who have been awarded doctorate degrees in non-life sciences

disciplines over the past 20 years. These individuals are untrained, but are likely to be capable of acquiring the necessary skills given 4-10 years of training, equivalent to earning a PhD.

Historical example: Ted Kaczynski, the Unabomber, was a Berkeley mathematics professor who spent over a decade engaging in a mail-bombing campaign with the goal of bringing down the industrial system to keep it from depriving people of dignity and autonomy. In the Unabomber Manifesto, he asserted that "the temptation presented by the immense power of biotechnology would be irresistible".

Question: What is the probability that at least one of these 4.3 million capable untrained actors will attempt to cause a pandemic within the next 10 years?


Probability of a Capable Unskilled Actor Attempting to Cause a Pandemic Within the Next 10 Years

x̄: mean , M: median

Lower Bound (5th percentile)	x̄ =12% , M =5%
Upper Bound (95th percentile)	x̄ = 34% , M = 30.5%

Question 4 - Efficacy of Counterterrorism Efforts

Background: To successfully carry out a bioterrorist attack using a pandemic agent, an actor would need to acquire the necessary materials, assemble the virus or obtain and isolate live samples, and successfully disseminate the virus - without being reported by coworkers or noticed by intelligence services.

Question: Assume someone with all the necessary resources and skills attempts to make and release a pandemic virus. How likely is it that they will be detected and prevented by current counterterrorism efforts: whistleblowers, intelligence services, law enforcement or any combination of similar factors?
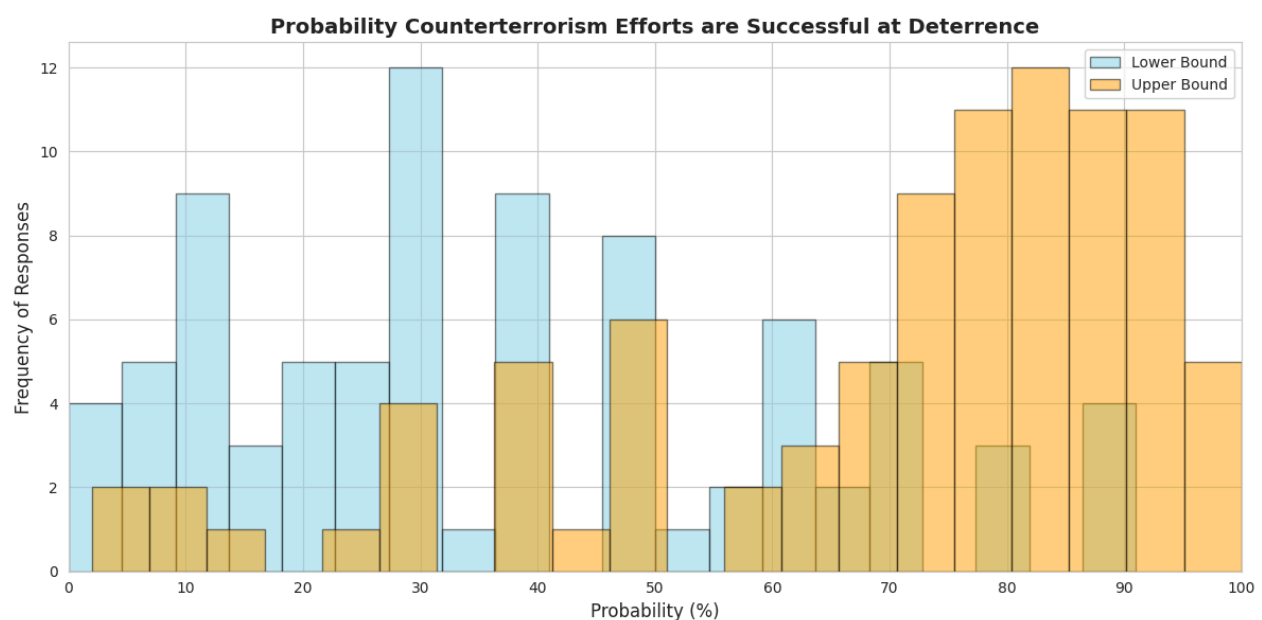
x̄: mean , M: median

Lower Bound (5th percentile)	x̄ =37% , M =30.5%
Upper Bound (95th percentile)	x̄ = 70.5% , M = 80%